IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-6, Issue-6, June 2018 E-ISSN: 2347-2693

A Survey on Cloud Service Scheduling Using Genetic Algorithm

M. Durairaj[1*], C. Dhanavel[2]

[1] Computer Science, School of Computer Science and Engineering, Bharathidasan University, Tiruchirappalli, India
[2] Computer Science, School of Computer Science and Engineering, Bharathidasan University, Tiruchirappalli, India

[*]*Corresponding Author: durairaj.bdu@gmail.com, Mobile.: +91 9487542202*

Available online at: www.ijcseonline.org

Accepted: 12/Jun/2018, Published: 30/Jun/2018

***Abstract*—** Cloud services are widely used around the world since the cloud services are playing a key role in many industries such as Supply Chain, Networking, Storages, etc. Different task scheduling algorithms have been used to handle cloud service applications, but none of the algorithms contain all the constraints such as load balancing, makespan time, cost and the time of execution. The scheduling technique considers well when it efficiently performs utilizing resources of the cloud. The heuristic scheduling algorithm provides the optimal solution, thereby increasing the efficiency of the overall system. Heuristic methods such as Genetic Algorithm (GA) are deals with the natural selection of solutions from the all possible solutions. Genetic algorithms schedule the cloud tasks according to the computational power of the system, memory resources and requirements of the tasks. The aim of this survey is to propose a technique to minimize the completion time and cost of tasks and maximize resource utilization using Genetic Algorithm (GA). This work also presents the comparative analysis of different task scheduled applications proposed by the researchers during the last five years.

***Keywords*—**Cloud Computing, Scheduling, Genetic Algorithm, Optimization, Scheduling Algorithms.

I. INTRODUCTION

Cloud computing is the delivery of services such as servers, storage capacities, databases, networking, Programming's and analysis over the internet. Cloud Computing can also be defined as a type of Parallel, Grid and Distributed computing that consist of interconnected computers [1,2]. The Information Technology (IT) industries are growing day-by-day accordingly the demand for computing and storage are increasing in a rapid manner. In other terms, the cloud computing is based on PAY (Pay-as-You) use financial models. The cloud computing is a multi-level distribution of virtualization. There are two different classes of clouds: deployment model and service model [3]. The deployment model has something to say about where the cloud is found, and what the purpose of the cloud is the categorizations of the cloud are as public cloud, private cloud, community cloud and hybrid cloud.

Cloud computing services are given by CSS (Cloud Service Supplier) as per client requirements. Different user's requests are fulfilled as if they are given a different quality of services in a cloud computing environment. Cloud services are dynamic in behaviour and resources with multiple services are provided to users [3,4]. The cloud best-known service models are as Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) [5]. The IaaS is a service where a user to give the updated and powerful mechanism of infrastructure through the internet (e.g. Amazon EC2, GoGrid, Linode, etc.). The PaaS is providing some platform and applications to the clients or users (e.g. Email services, Google AppEngine). The SaaS is a software mode access through the internet (e. g. Google Apps, etc.) [6]. The example of Cloud computing service model is as shown in Figure 1.

Scheduling is one of the most attractive activities performed in cloud computing environments to increase the productivity in terms of workload. The cloud scheduling algorithm in cloud computing is designed for the proper allocation of resources in minimum execution time [3]. Scheduling is being performed on the basis of different parameters to increase cloud computing performance. Overall, the scheduling algorithm always improves the usages of Central Processing Unit (CPU), performance time, reducing cost and total throughput [4].

Genetic Algorithms (GAs) are a part of evolutionary processes, which is one of the rapidly developing Area of Artificial Intelligence (AI). Genetic algorithms are search algorithms based on the mechanics of natural selection and natural genetics [ ]. In this algorithm, evolution is mostly determined by natural selection or different individuals competing for resources in the environment. Those individuals that are better are more likely to survive and spread their genetic material. To use of genetic algorithm you must represent the genome or chromosome. Solutions which are selected to form new solutions (offspring) are selected according to their fitness level. The basic GA is very generic and there are many aspects that can be

implemented differently according to the problem. For instance, the problems such as the representation of the solution or chromosomes, type of encoding, selection strategy, type of crossover and mutation operators, etc [7].

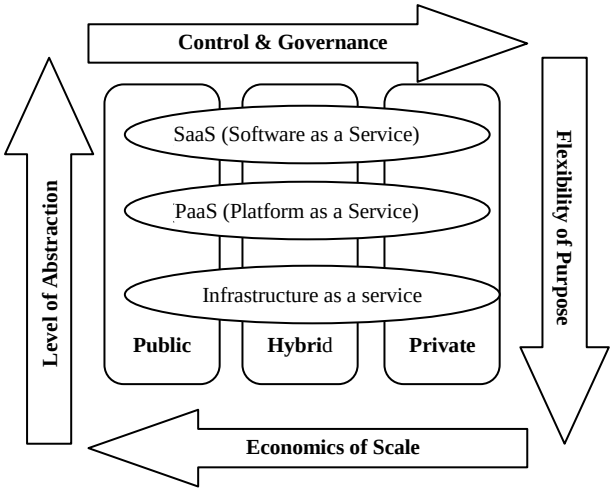

Figure 1. Cloud Computing Services Model

Rest of the paper is organized as follows, Section I contains related work with brief introduction. This section also discusses elaborately on the need of scheduling algorithms and different processes of genetic algorithm. Section II contains the description on role of scheduling algorithms in cloud. In this, the research outcome of various researchers on this subject is listed. Section III contains methodology of cloud applications using scheduling algorithms. Section IV briefly discusses the results obtained by different researchers during last five years. Section V briefs the proposed methodology and recommendations. Section VI concludes with the study outcomes and future directions.

II. RELATED WORK

In this paper, different categories of cloud service scheduling algorithms using Genetic process presented. The research papers published on this concept during the last five years are reviewed for this study and presented pros and cons with comparative analysis. Discussion and review of research articles are presented under the subheadings such as Need of scheduling in Cloud, Genetic algorithms' processes and Scheduling Algorithms.

A. Need of Scheduling in Cloud

There are seven major factors in Scheduling that used to measure the performance [8].

1) *Throughput Time:* Throughput time refers to the performance of tasks by a computing service or device over a particular period. It quantity the amount of finished work against time used and might be used to measure the performance of a processor.
2) *Turnaround Time:* In computing, turnaround time is the amount of time to execute a particular task.
3) *Waiting Time:* Amount of time a process has been spent on the particular task.
4) *Response Time:* A measure of time takes from when a request was presented until the first response is released, not output.
5) *Utilization of CPU Time:* CPU (Central Processing Unit) utilization refers to the computers usages of processing tasks or the number of tasks handled by a CPU.
6) *Computational Complexity (job length, processing power):* The complexity of every processor job length is the number of resources required for running it on the computing tasks.
7) *Computing Cost (Processor cost):* The processor cost or computing cost is choosing manipulates the cloud resources.

B. Genetic Algorithm Process

Genetic algorithm (GA) is based on the organic concept of generating the population. The genetic process is the natural selection of evolution as a means of advance of the optimal solution [9]. A Genetic algorithm is used to find out the optimal sequence to allocate the processor. The use case diagram for a Genetic algorithm is as shown in Figure 2. The steps involved in the Genetic process are as follows [10,11].

1) Initial Population: The initial population is the set of all individuals that are used in the GA to find out the optimal solution. In every solution, a population is called as an individual. A solution set is taken as possible in the initial population or chromosome, each and every individual represents the chromosome for making suitable genetic operations.

2) Fitness Function: The efficiency of any individual population depends on the fitness value. The genetic algorithm evaluates each chromosome by a fitness function in every population. The fitness function is propelled to the Genetic algorithm.

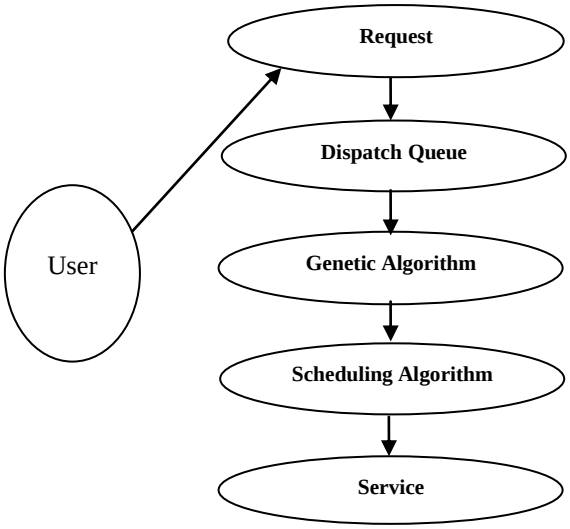

Figure 2. Use case diagram for Genetic Algorithm

3) Selection: Selection is the way towards deciding the circumstances of a specific individual for reproduction and hence the quantity of offspring of an individual will be created. In this approach, the choice directs the genetic search toward promising regions within the search space and that will enhance the performance of Genetic algorithms. There are various selection methods available in the GA to select the best individual, such as Roulette wheel selection, Rank selection, Tournament selection, Steady state selection and Elitism.

4) Crossover: Crossover operation is performed on all the chromosomes. Two chromosomes are combined to create a new one in the crossover mechanism. The crossover techniques are used in single-point, multi-point, and uniform crossover.

5) Mutation: Once the crossover is completed, mutation takes place in the GA. Mutation is a common operator used to help preserve diversity in the population by finding new points in the search phase to evaluate. It provides new gene value added to the gene pool. Two types of mutation techniques are single-point mutation and multi-point mutation. A commonly used method for mutation is called single point mutation. The genetic algorithm is able to produce the best solution process. The general procedure for GA is as illustrated in Figure 3.

Algorithm:

Initial population generation;
Evolution;
While (stop condition) {
Selection;
Crossover;
Mutation;
Evolution;
}
Output best solution;

Figure 3. General Procedures for a Genetic Algorithm

C. Scheduling Algorithms in Cloud

Durairaj et al. proposed a scheduling technique which uses the principle of time measurement but simple in nature. This technique divides the time into multiple time slices and assigns specific time slice or case interval to each node using time scheduling rule. Each node has its unique measure with its operation. Round Robin method used for queue management system as per literature, the Round Robin is the best method for queue management [12].

Athokpam Bikramjit Singh, et al., used Round Robin scheduling algorithm and processes are dispatched in a FIFO (First in First Out) manner but are given a restricted amount of CPU time called a Time-aggregate or a Slice. The process does not complete before its CPU-time terminates and CPU is acquired more time in order to follow the procedure of holding up in a queue. The waiting processes are then placed on the back to the ready state [13].

Pardeep Kumar, et al., Proposed a methodology where minimum completion time for all tasks is found. The proposed algorithm chooses small tasks to be executed firstly, and then turn postpones to the larger tasks for a long time. This starts with a set of every single unassigned task. At a point of time, base time and base value are selected which is considered as the minimum time among all the tasks on any assets or resources. Then, according to that base time, the task is scheduled on the related machine [14].

Tarun Goyal et al., Proposed Min-min start with a set Meta task, task deals with another task, of every unassigned task and possess two phases. In the first phase, a set of the least expected finish time for each task with other task is found. In the second phase, the task with a general minimum expected finish time from Meta task is picked and assigned to the relating machine. At that point, the task is removed from Meta task and the procedure is repeated to the point that all tasks in the Meta task mapped. However, the min-min algorithm first completes the shorter assignments and after that executes the longer assignments [2].

In Min-min, one of the heuristic scheduling algorithms, select the smallest task first from all the available or accessible tasks and assign to a machine to provide a base time (quickest device) for the tasks. It multiplies total completion time of every task and thus increases the makespan. However, the device assigns the smaller task on faster machines before scheduling. Here, the normal finishing time and execution time for a task is thought to be a

relatively same value or close value. The long tasks have to stay for finishing the execution of the little ones. Hence, the proposed techniques enhance the device overall throughput [15].

Sonali Jain, et al. proposed an algorithm which fixes the task with maximum execution time and distributing to the resources with minimum computation time. The algorithm worked with smaller rang but fails with a large range of tasks worked. The max-min algorithm is almost similar to the min-min algorithm. This algorithm executes large task first, then the smaller tasks are delayed for a long time. According to the finding, the minimum execution time, the maximum value is selected which is the maximum time centre of all the tasks or any resources [14,16].

In [13], RASA (Resource-Aware-Scheduling-Algorithm) is proposed by integrating Min-min and Max-min algorithms. RASA uses a point of benefits in min-min and max-min and also catches the drawbacks. In fact, that the deadline for each task, arriving rate of allocated tasks, amount of task computation of the each of the resources are computed but communication costs are not compared. Finally, the RASA test is performed on the large scale of distributed systems. It achieves the lower Makespan and Good QoS (Quality of Service) [11,17].

Jeni Patel et al. proposed an improved the Max-min algorithm known as enhanced max-min algorithm. The Max-min algorithm and Enhanced Max-min algorithm are based on execution time rather than consumption time. By comparing the two algorithms, the only differences is that the largest task should be assigned then produce the resources at a lower level, improved max-min assign tasks with an average completion time to Resource produced and process at a minimum completion time. Rather than choosing a large task, the algorithm selects an average task, then resulted in decreased total makespan and also balanced crosswise of resources. The improved max-min is based on expected execution time rather than finishing time and selection of tasks [18].

Shekhar Singh, et al. proposed the independent task schedule which is based on adapting the different calculation and memory necessity. Though the Genetic algorithm indented to explain the bundling of optimization problems, its contribution is little for global enhancement. The fitness function regulates to the deadline of the cost of delay. In this technique, the issues of execution time are the better [19].

In [1,19], a new scheduling algorithm is proposed which has an improved version of a genetic algorithm. In the proposed scheduling method the min-min and max-min algorithm companies with a genetic algorithm. The initial population is generated by using the max-min and min-min techniques. This algorithm is performed to reduce the makespan by making use of resources in the virtual machine techniques.

Sourabh Budhiraja, et al. proposed Meta heuristic scheduling techniques minimize the execution time and execution cost. Each and every task contains computational complexity and computing capacity of processing elements. Those individuals are selected for generating the population values of makespan and execution cost is lesser than the standard Genetic algorithm [1].

III. METHODOLOGY

Scheduling algorithms focused on to enhance the proper utilization of cloud resources with the reduced task completion time. The purpose of task scheduling is to allocate tasks to indented resources in an instance of time. There are various scheduling algorithms with different techniques proposed to solve a variety of problems in task scheduling. The comparison of the different application on cloud services schedule is presented in Table 1.

IV. RESULTS AND DISCUSSION

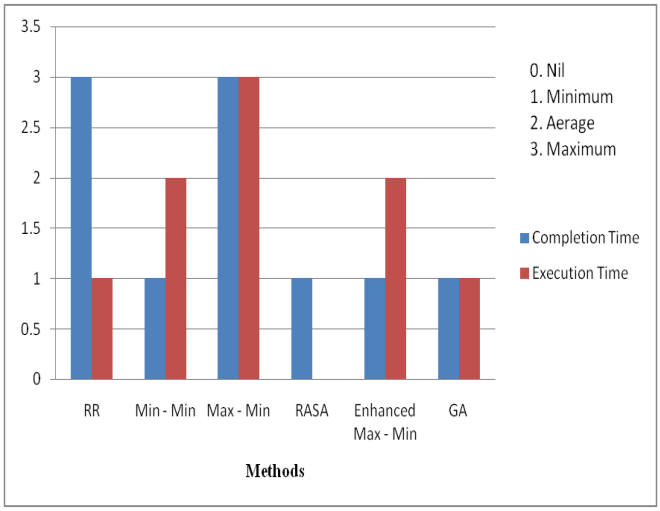

Figure 4. Comparisons of Completion and Execution Time

Figure 4 shows the graphical representation of comparison of completion and execution time of different scheduling algorithms. The graph illustrates the completion time and execution time of scheduling algorithms. The completion time is denoted by three different categories, where 1 denoted by minimum time completion, 2 denoted by average time completion and 3 denoted by maximum time completion. Similarly wherein execution time, as mentioned in the completion time, 1 denoted by minimum time completion, 2 denoted by average time completion and 3 denoted by maximum time completion. It can be observed from the graph that the Makespan are improved fairly in many methods.

Table 1. Comparison of Different Scheduling Algorithms used in Cloud Services

S.No	Scheduling Algorithms	Author's Name	Performances	Nature of Task	Advantage	Disadvantage
1.	Round Robin [12]	Durairaj M Menaka A	• Completion time: Maximum • Execution time: Minimum • Reduces Makespan	Grouped Task	Load balance is done fairly.	Improve time scheduling of tasks.
2.	Round Robin[13]	Athokpam Bikramjit Singh, Sathyendra Bhat J, Rajesh Raju, Rio D'Souza	• Completion time: Average • Execution time: Minimum • Reduces Makespan	Grouped Task	The process dispatched in FIFO techniques.	The CPU is pre-empted and given the next task is waiting in a queue.
3.	Min-Min Algorithm [14]	Pardeepkumar, Amandeep Verma	• Completion time: Minimum • Reduces Makespan	Independent Task	Better Makespan compared to other algorithms.	Poor load balancing and QoS factors are Not considered.
4.	Min-Min Algorithm [2]	Tarun Goyal	• Completion time: Minimum • Execution time: Average	Independent Task	Speedy processing	Ineffective resource sharing
5.	Min-Min Algorithm [15]	Teena Mathew	• Completion time: Minimum • Reduce Makespan	Independent Task	Load balancing is considered.	There are not best resources in the optimization selection.
6.	Max-Min Algorithm [16]	Sonali Jain	• Completion time: Maximum • Execution time: Maximum • Better Makespan	Independent Task	Better makespan compared to other Algorithms.	Poor load balancing and QoS factors are Not considered
7.	Resource-Aware-Scheduling algorithm (RASA) [13][17]	Athokpam Bikramjit Singh, Pinal Salot	• Execution time: Minimum • Reduces Makespan	Grouped Task	Lower makespan with Good QoS in the RASA techniques	Poor Resource Utilization and don't consider for computing cost.
8.	Enhanced Max-Min Algorithm [18]	Jeni Patel, Jignesh Prajapati	• Completion time: Minimum • Execution time: Average • Reduce Makespan	Independent Task	Improves makespan and Load balancing when large numbers difference occurred in the Length of the longest task and other tasks or speed of Processors	Only limited parameter considered.
9.	Independent Task Scheduling based on Genetic Algorithm In cloud computing [19]	Shekhar Singh, Mala Kalra	• Execution time: Minimum • Resource utilization	Independent Task	Fitness function deadline is maximized.	Reducing the solution space in genetic algorithms
10.	Independent Task Scheduling in the Cloud Computing by Improved Genetic Algorithm [1]	Sourabh Budhiraja, Dheerendra Singh	• Completion time: Minimum • Execution time: Minimum • Resource utilization	Independent Task	The initial population is generated by max-min and min-min algorithms.	Execution cost is based on fitness criteria.
11.	An Efficient Approach to Genetic Algorithm for Task Scheduling in Cloud Computing [19]	Shekhar Singh, Mala Kalra	• Execution time: Minimum • Reduces Makespan	Independent Task	The makespan and execution cost is less the standard genetic algorithm (GA)	Improve search space.

The Figure 4 illustrates completion time, execution time, performance and energy consumption which are considered in a moderate rate of the analyzed algorithms. Some of the methods are based on scheduling of priority resource utilization, cost, efficiency etc. Some algorithms primarily focus on to reducing a Makespan and to have better resource usage inside the cloud frameworks alongside improvement.

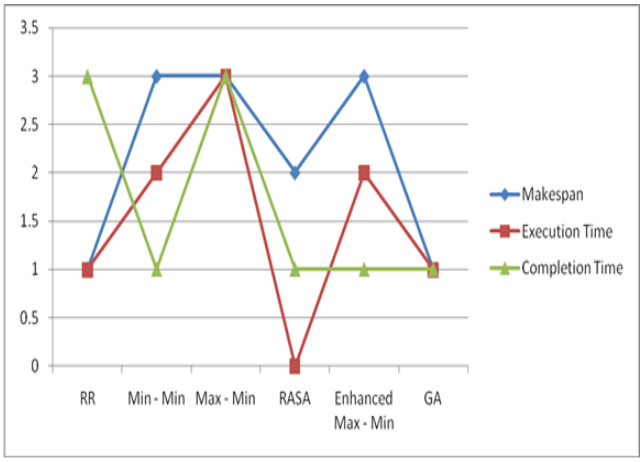

Figure 5. Makespan (Execution Time Vs Completion Time)

In Figure 5, the comparison of different scheduling algorithms based on Makespan, Execution time and Completion time are given. It is observed from the results obtained, the algorithms such as Enhanced Max-Min and GA performs relatively better than other algorithms.

V. PROPOSED METHODOLOGY

A genetic algorithm is an optimistic search algorithm based on the mechanics of natural selection and natural genetics [20]. The objective of the Genetic algorithm (GA) is to have maximum utilization of resources and to reduce the execution time. The comparative study between the proposed Scheduling Algorithm and Existing Scheduling Algorithms is carried out. The performances of scheduling algorithms are developed using Python programming and evaluated. The Python is an easy and effective programming tool which can be effectively used to program Genetic algorithms to test and implement. In this work, the Python is used to implement Genetic Algorithms.

VI. CONCLUSION AND FUTURE SCOPE

In this paper, different scheduling algorithms proposed by the different researchers during last five years are taken for study. The process of scheduling is one of the main tasks in Cloud Computing. In this work, the most popular algorithms which commonly used for scheduling in cloud computing are reviewed and effectively analysed. The performances of these algorithms are tabulated and compared. Every algorithm has its own advantages and disadvantages. The performance of algorithms is mainly depends on the size of the Tasks assigned. The metrics such as cost efficiency, completion time and effective resource utilization are used for performance analysis. It is observed from the performance analysis and comparative study, the Makespan of a scheduling algorithm improves only when the execution time of a task is reduced. This work concludes that the enhanced Max-Min algorithm and GA performs relatively better than others.

Minimization of execution times of scheduling algorithm is primary task of enhancement. The extension of this work is to focus on developing methodologies to combine two or more algorithms which reduce task execution time.

REFERENCES

[1] Sourabh Budhiraja, Dheerendra Sing, "*Task Scheduling using Genetic Algorithm in Cloud Computing Environment: A Review*", International Journal of Info. Tech. and Knowledge Management, Vol.8, No.1, pp.46-49, 2014.

[2] Tarun Goyal and Aakanksha Agrawal, *"Host Scheduling Algorithm Using Genetic Algorithm in Cloud Computing Environment",* International Journal of Research in Engineering & Technology, Vol.1, Issue.1, pp.7-12, 2013.

[3] Simply, Jagandeep Sidhu, "*Different Scheduling Algorithms in Different Cloud Environment",* International Journal of Advanced Research in Computer and Communication Engineering, Vol.3, Issue.9, pp.8003-8006, 2014.

[4] Hongyan Cui, Xiaofei Liu, Tao Yu, Honggang Zhang, Yajun Fang, and Zhongguo Xia, "*Cloud Service Scheduling Algorithm Research and Optimization*", Security and Communication Networks, Hindawi, Vol.2017, Article ID 2503153, pp.1-7, 2017.

[5] Durairaj, M and Kannan, P, "*Improvised Genetic Approach for an Effective Resource Allocation in Cloud Infrastructure",* International Journal of Computer Science and Information Technologies, Vol.6, Issue.4, pp.4037-4046, 2015.

[6] R. Durga, Lakshmi, N. Srinivasu *" A Review and Analysis of Task Scheduling Algorithms in Different Cloud Computing Environments "* International Journal of Computer Science and Mobile Computing, Vol.4, Issue.12, pp.235–241, 2015.

[7] S. Ravichandran, E.R. Naganathan, "*Dynamic Scheduling of Data Using Genetic Algorithm in Cloud Computing*", International Journal of Computing Algorithm, Vol.2, Issue.1, pp.127-133, 2013.

[8] M. Padmavathi, Shaik. Mahboob Basha, Srinivas Pothapragada, "*A Survey on Scheduling Algorithms in Cloud Computing*", *IOSR Journal of Computer Engineering, Vol.16, Issue.4, pp.27-32, 2014.*

[9] Safwat A. Hamad, Fatma A. Asmara *"Genetic-Based Task Scheduling Algorithm in Cloud Computing Environment "* International Journal of Advanced Computer Science and Applications, Vol.7, No.4, pp. 550-556, 2016.

[10] A. Kaleeswaran, V. Ramasamy, P. Vivekanandan, "*Dynamic Scheduling of Data using Genetic Algorithm in Cloud Computing*", International Journal of Advances in Engineering & Technology, Vol.5, Issue.2, pp.327-334, 2013.

[11] R. Durga Lakshmi, N. Srinivasu, "*A Dynamic Approach to Task Scheduling in Cloud Computing using Genetic Algorithm*", Journal of Theoretical and Applied Information Technology, Vol.85, No.2, pp.124-135, 2016.
[12] M. Durairaj, A. Menaka, "*Load Balancing in Cloud Computing*", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 5, Issue.8, pp. 862-870, 2015.
[13] Athokpam Bikramjit Singh, Sathyendra Bhat J, Rajesh Raju, Rio D'Souza, "*A Comparative 8dy of Various Scheduling Algorithms in Cloud Computing*", American Journal of Intelligent Systems, Vol.7, Issue.3, pp.68-72, 2017.
[14] Pardeep Kumar, Amandeep Verma, "*Scheduling Using Improved Genetic Algorithm in Cloud Computing for Independent Tasks*", International Conference on Advances in Computing, Communications, and Informatics, India, pp.137-142, 2012.
[15] Teena Mathew, K. Chandra Sekaran, John Jose, "*Study and Analysis of Various Task Scheduling Algorithms in the Cloud Computing Environment*", International Conference on Advances in Computing, Communications, and Informatics, India, pp.658-664, 2014.
[16] Sonali Jain, "*Task Scheduling in Cloud Computing using Genetic Algorithm*", International Journal of Computer Science, Engineering, and Information Technology Research, Vol.6, Issue.4, pp. 9-22, 2016.
[17] Pinal Salot, "*A Survey of Various Scheduling Algorithm in Cloud Computing Environment*", International Journal of Research in Engineering and Technology, Vol.2, Issue.2, pp. 131-135, 2013.
[18] Jeni Patel, Jignesh Prajapati, "*A Survey of Various Scheduling Algorithms and types of Resources Provisioning in Cloud Environment*", International Journal of Engineering and Computer Science, Vol.4 Issue.1, pp.10132-10134, 2015.
[19] Shekhar Singh, Mala Kalra, "*Task Scheduling Optimization of Independent Tasks in Cloud Computing using Enhanced Genetic Algorithm*", International Journal of Application or Innovation in Engineering & Management, Vol.3, Issue.7, pp.286-291, 2014.
[20] M. Durairaj, T. Sathyavathi, "*Applying Rough Set Theory for Medical Informatics Data Analysis*," International Journal of Scientific Research in Computer Sciences and Engineering, Vol. 1, Issue. 5, pp. 1–8, Oct. 2013.

Authors Profile

Dr. M. Durairaj completed his Ph.D. in Computer Science at Bharathidasan University on April 2011. Prior to that, he received my master degree (M.C.A.) in 1997 and bachelor degree (B.Sc. in Computer Science) in 1993 from Bhrathidasan University. His Ph.D. work was to study different possibilities and device a methodology for hybridizing two Machine-learning techniques for making an effective prediction system for processing clinical/medical data. At present, he is Assistant Professor in Computer Science at Bharathidasan University, prior to this he was Research Associate at National Research Centre on Rapeseed-Mustard (Indian Council of Agricultural Research), Rajasthan and National Institute of Animal Nutrition and Physiology (ICAR), Bangalore, India for 12 years. He has 20 years of teaching and research experience in his credit. He has more than 90 publications in his credit. His area of research includes Data Mining, Data Science, Soft Computing, Cloud Computing and Big Data Analytics.

Mr. C. Dhanavel is currently researh scholar pursuing M.Phil. in Computer Science from Bharathidasan University, Tiruchirappalli, India and done his B.Sc. in Computer Science and M.Sc. in Computer Science from Bharathidasan University, Tiruchirappalli, in the year of 2016. His Area of Interest is Cloud Computing, Scheduling, Mathematical Algorithms, etc.

