ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

Improved Genetic and Memetic based Task Scheduling in Cloud Computing Environment

Ashwini Lokhande , Prof.Rahul Gaikwad

Research Scholar, Department of Computer Engg, Godavari College of Engineering, Jalgaon, India

Asst. Professor, Department of Computer Engg, Godavari College of Engineering, Jalgaon, India

ABSTRACT: In the current scenario, cloud computing has emerged widely used in companies and enterprises. In this paper, a task scheduling algorithm has been introduced which allocates and executes user's tasks. This proposed algorithm is based on Improved genetic algorithm (IGA) andimproved memetic algorithm(IMA). The main goal of this algorithm is to reduce completion time and execution cost of different tasks and maximize resource utilization. Cloudsim Toolkit is used to evaluate the performance of the proposed algorithm.

KEYWORDS: Cloud computing, Task Scheduling, Genetic algorithm, mematicalgorithm, Improved Genetic algorithm ,Improved memetic algorithm

I. INTRODUCTION

Cloud computing is an on demand service in whichshared resources, information, software and other devicesare provided according to the clients requirement atspecific time. It's a term which is generally used in case ofInternet. The whole Internet can be viewed as a cloud.Capital and operational costs can be cut using cloudcomputing. In case of Cloud computing services can beused from diverse and widespread resources, rather thanremote servers or local machines. There is no standarddefinition of Cloud computing. Generally it consists of abunch of distributed servers known as masters, providingdemanded services and resources to different clientsknown as clients in a network with scalability andreliability of datacenter. The distributed computers provideon-demand services. Services may be of softwareresources (e.g. Software as a Service, SaaS) or physicalresources (e.g. Platform as a Service, PaaS) orhardware/infrastructure (e.g. Hardware asa Service, HaaS or Infrastructure as a Service, IaaS).Amazon EC2 (Amazon Elastic Compute Cloud) is anexample of cloud computing services

II. RELATED WORK

In the Paper of "task scheduling in Cloud Computing" Aakansha Sharma and Sanjay Tyagi found that in task scheduling not only execution time, cost, response time, flow time, throughput, and average resource utilization, e but also improvement is required in some areas like make span, Time/space complexity & execution cost. They observed that Optimization based task scheduling may be further considered for achieving more effective task scheduling.

In this paper of "Review paper on various scheduling techniques in cloud environment "They taken , existing scheduling algorithm are considered and they all are compared by using different parameters as well as tools. Mostly they all are work on to minimize the execution time, faster response time and maximum utilization of resources. They found that Existing scheduling algorithms does not consider the load balancing, availability and reliability. Therefore, there is a need to implement such scheduling algorithm that can improve the reliability, availability and load balancing in cloud computing environment. In future, algorithm based on migration of task from one machine to another can also be introduced.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

In the paper of "Research of cloud computing task scheduling algorithm based on improved genetic algorithm "GE Junwei and YUAN Yongsheng presents a the genetic algorithm (MGA) which comprehensive consideration of the total task completion time, average task completion time and costconstraints, it can obtained not only total task completion time and average task completion time is short, but also therequired costs is low. The result of this task scheduling is ideal.

III. METHODOLOGY

Existing Algorithm

A. Genetic Algorithm

Genetic Algorithm (GA) is based on the biological concept of generating the population. GA is considered a rapidly growing area of Artificial IintelligenceBy Darwin's theory of evolution was inspired the Genetic Algorithms (GAs). According to Darwin's theory, term "Survival of the fittest" is used as the method of scheduling in which the tasks are assigned to resources according to the value of fitness function for each parameter of the task scheduling process The main principles of the GA are described as follows

1) Initial Population

The initial population is the set of all individuals that are used in the GA to find out the optimal solution. Every solution in the population is called as an individual. Every individual is represented as a chromosome for making it suitable for the genetic operations. From the initial population, the individuals are selected, and some operations are applied on them to form the next generation. The mating chromosomes are selected based on some specific criteria.

2) Fitness Function

The productivity of any individual depends on the fitness value. It is the measure of the superiority of an individual in the population. The fitness value shows the performance of an individual in the population. Therefore, the individuals survive or die out according to the fitness or function value. Hence, the fitness function is the motivating factor in the GA.

3) Selection

The selection mechanism is used to select an intermediate solution for the next generation based on the Darwin's law of survival. This operation is the guiding channel for the GA based on the performance. There are various selection strategies to select the best chromosomes such as roulette wheel, Boltzmann strategy, tournament selection, and selection based on rank.

4) Crossover

Crossover operation can be achieved by selecting two parent individuals and then creating a new individual tree by alternating and reforming the parts of those parents. Hybridization operation is a guiding process in the GA and it boosts the searching mechanism.

5) Mutation

After crossover, mutation takes place. It is the operator that introduces genetic diversity in the population. The mutation takes place whenever the population tends to become homogeneous due to repeated use of reproduction and crossover operators. It occurs during evolution according to a user-defined mutation probability, usually set to fairly low. Mutation alters one or more gene values in the chromosome from its initial state. This can produce the entirely new gene values being added to the gene pool. With this new gene values, the genetic algorithm may be able to produce a better solution than was previously

B.Memetic Algorithm

The genetic algorithm is not well suited for fine-tuning structures which are close to optimal solution. The memetic algorithms can be viewed as a marriage between a population-based global technique and a local search made by each of the individuals. They are a special kind of genetic algorithms with a local hill climbing. Like genetic algorithms, memetic Algorithms are a population-based approach. They have shown that they are orders of magnitude faster than traditional genetic Algorithms for some problem domains. In a memetic algorithm the population is initialized at random or using a heuristic. Then, each individual makes local search to improve its fitness. To form a new population for the next generation, higher quality individuals are selected. The selection phase is identical inform to that used in the classical genetic algorithm selection phase. Once two parents have been selected, their chromosomes are combined and the classical operators of crossover are applied to generate new individuals. The latter are enhanced using a local search technique. The role of local search in memetic algorithms is to locate the local optimum more efficiently then the genetic algorithm. Figure 3 explains the generic implementation of memetic algorithm.

1. Encode solution space
2. (a) set pop_size, max_gen, gen=0;
   (b) set cross_rate, mutate_rate;
3. initialize population
4. while(gen <gensize)
   Apply generic GA Apply local search

end while Apply final local search to best chromosome

Hill climbing local search algorithm

The hill climbing search algorithm is a local search and is shown in below.It is simply a loop that continuously moves in the direction of increasing quality value

While (termination condition ins not satisfied) do
  New solution ← neighbors(best solution);
If new solution is better then actual solution then
  Best solution ← actual solution
End if End while

Proposed Algorithm

In this work, a improved GA is proposed to solve task scheduling problem in Cloud computing environment to enhance the completion time for executing all tasks on the VMs, in the same time, minimize the total cost of usage the resource and maximize utilization of the resource. The main idea of this proposed algorithm (i.e., IGA) is that after each selection in the population, there is a solution that might satisfy good fitness function, but it is not selected to crossover process. By the proposed algorithm, this solution is not removed from the population, but it is chosen and added to the population when next iteration is started.

This step is considered as a good step as some of the iterations can generate the best solution.

1) Initialize Population

According to the proposed IGA algorithm, the population is randomly generated using encoded binary (0, 1).

2) The Fitness Function Representation

The main objective of task scheduling in the Cloud computing is to reduce completion time for execution all tasks on the available resources. Therefore, the completion time of task on as is defined using equation Where denotes maximum time for complete Task ion .*n*and *m* denote the number of tasks and virtual machines respectively

VM 1:- TS1, TS5, TS3

Gene 1

Gene 2

Gene 3

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

VM 2:- TS7, TS2, TS6
VM 3:- TS4, TS8, TS9

3) Selection Process
Tournament selection is computationally more efficient and more amenable to parallel implementation . Therefore, the developed IGA algorithm, Tournament Selection is used to overcome the limitation of the population size. Two individuals are chosen at random from the population. A random number r is then chosen between 0 and 1. If r < k (where k is a parameter, for example, 0.75), the fitter of the two individuals is selected to be a parent; otherwise the less fit individual is selected. The non-chosen individuals are then returned to the original population and could be selected again.

4) Crossover
In the proposed IGA algorithm, the new crossover has been used differently from the used crossover in the original GA. Therefore, two chromosomes which are selected to crossover process to generate two offspring will be considered as offspring also. So, the proposed crossover produces four children (see Figure 4). After that, the two best children are chosen from these.

5) Initialize Subpopulation
After each iteration, subpopulations (i.e., new populations after crossover) are added into old populations (i.e., parents). This step can enhance the diversity of population. This idea is introduced by the modified IGA algorithm.

6) Keep Best Solution
There is a solution that might satisfy good fitness function, but it is not selected during the crossover process. By the proposed IGA algorithm, this solution is not removed from the population, but it is chosen and added to the population when next iteration is started. This step is considered as good step as some of the iterations can generate the best solution.
Generally, according to the modified IGA algorithm, a set of modifications have been introduced. These modifications are as follows.
 The tournament is used instead of the roulette wheel in the selection process to select the best solution.
 The solutions not chosen in the selection process are considered and added to the new population. This might help in generating the best solution in the next generations.
 The new crossover is introduced by considering parents individuals as new child
After each iteration, subpopulations (i.e., new populations after crossover) are added into old populations (i.e., parents).

Improved Memetic Algorithm
The memetic algorithm can be used to optimize several conflicting objectives by combining them as a single objective. The memetic algorithm is an evolutionary algorithm that simulates the natural evolution and is successful in the past. For the workflow scheduling problem, the memetic algorithm generates a set of schedules selected by the user constraints. The schedules are refined using the genetic operators and the optimal solution is reached.

Crossover
The scheduling for jobs here is precedence constrained. The order should always be maintained in the dependencies between jobs. If *Jk*has to get the output of *Ji*, then *Ji* should precede *Jk*in all the possible orders generated. The crossover used here is Partially Matched Crossover(PMX) shown in Algorithm 1. Given two schedules ‘A’ and ‘B’ the PMX randomly picks two crossover points. This crossover point is used for the construction of next generation schedule. Here substring is a sub sequence of job in the given schedule referred with starting and ending position. The crossover is performed by considering the following facts.
i. Same job should not repeat.
ii. The dependencies among job are maintained.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

The procedure for Crossover operation is given in Algorithm 1:

Algorithm 1: Partially matched crossover
i. Procedure Partially_Matched (Schedule A, Schedule B)
ii. n is the number of jobs
iii. Select a random value P1 between 0 and n-1
iv. Select second random value P2 between 0 and n-1 where P1 ≠P2
v. If P1 > P2, swap P1 and P2
vi. From the Schedule A, copy he substring from position P1 to P2 as O1.
vii. From the Schedule B, copy he substring from position P1 to P2 as O2.
viii. For all allels(Jobs) in substring O1 and O2
ix. If the allels in substring (A,0,P1)and substring (A,Pi+1,n-1) does not contain entries from substring (A,P1,P2) and substring (B,P1,P2), then replace substring O1 by O2,
x. else repeat steps 3 to 9.
xi. For string O1 and O2
xii. Find the remaining allels in O1 and O2.
xiii. Place them from left to right.
xiv. Check if the dependencies among jobs are satisfied. If not repeat steps 3 to 12

Mutation
The mutation is a memetic operator that maintains alteration in one or more gene values shown in algorithm 2. It is an occasional random alteration of a value in a schedule with small probability.

Algorithm 2: Mutation
i. Procedure Mutation(Schedule A)
ii. n is the number of jobs
iii. Select a random value P1 between 0 and n-2
iv. Select a random value P1 between 1 and n-1 and P2 > P1
v. Swap the allels in positions P1 and P2 and generate new schedule .
vi. Check if dependency among jobs is maintained
vii. If not repeat step 3 to 6
viii. Check the fitness of new string ix. Repeat step 3 to 8 till a better fit solution is obtained or if the number of mutation done is 'n'

5.3 Initial population
In any scheduling algorithm, the solution space is massive. The generation of initial population plays a major role in the convergence to the solution. For a problem size of '*n*' jobs, the initial population of '*n*' possible schedules are generated. Out of the '*n*' schedules, first four schedule are based on
i. Topological ordering
ii. Earliest Finish Time algorithm
iii. Shortest Job First
iv. Minimal Cost Ordering algorithm
The remaining *n-4* schedules are generated on random ordering. Any schedule generated will be considered only if the dependency is maintained. This is depicted in algorithm 3.
The four algorithms included in generation of initial population helps in quicker convergence towards the optimal solution. For every schedule generated, the instance-type and job to instance are mapped accordingly.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

Algorithm 3: Initial population
i. Procedure initial population
ii. n is the number of jobs
iii. m is the number of instance types
iv. Generate the first schedule by topological ordering based on indegree of any vertex.
v. Generate the second schedule by heterogeneous earliest finish time algorithm.
vi. Generate the third schedule by Shortest job first algorithm
vii. Generate the fourth schedule based on the descending order of cost .
viii. For i=1 to n-4
ix. Generate a random schedule 'I' so that the dependency is maintained.
x. Find the fitness of each schedule. The best fit is the queen. The next N/2 -1 fit schedules are Drones. The remaining N/2 schedules are workers.
xi. End procedure

IV. RESULT AND OUTCOME

1. **Completion Time:-**The number of instructions the processor can execute per second:
TQ=(NP * MIPS)/1000 whereas, TQ - Time Quantum, NP - Number of Processors, MIPS - Million Instructions Per Second,

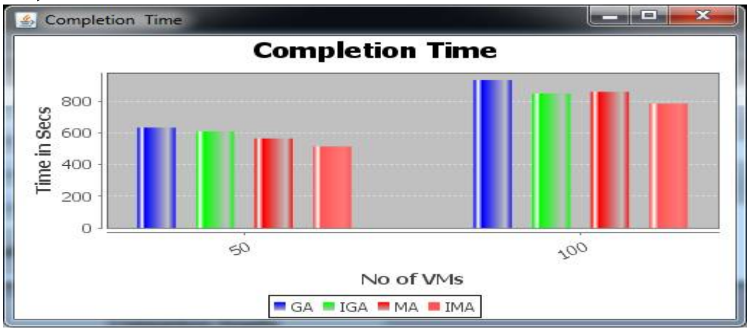

Fig 1. Completion time graph

Above fig shows the completion time graph of 50 and 100 Virtual machines for Genetic(GA),Improved genetic(IGA),Memetic(MA) and Improved Memetic algorithm(IMA) .Genetic algorithm required largest completion time and Improved Memetic algorithm required lowest completion time .

2. **Execution Cost:**
In addition, the total cost of execution of all tasks on the available VMs is calculated as

$$Total\ Cost = \frac{Task\ length\ *\ Cost\ per\ seconds}{VM\ mips} + Processing\ Cos$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 7, Issue 5, May 2019

Fig 2. Execution cost graph

Above fig shows the Execution cost graph of 50 and 100 Virtual machines for Genetic(GA),Improved genetic(IGA),Memetic(MA) and Improved Memetic algorithm(IMA) .Genetic algorithm required largest Execution Cost and Improved Memetic algorithm required lowest Execution cost .

V. CONCLUSION

Due to the shortcomings of past task scheduling problem, the introduction of improved genetic and improved memetic algorithm under the cloud environment has many good points. For example, it can meet the balance of cost and performance, and up to the requirements of the load balancing fairness. At the same time it can reduce the task execution time and increase the chance of success, and also improve the optimized solution for task scheduling in virtual machin.users' comprehensive QoS significantly.

REFERENCES

1. Yang Cui, Zhang Xiaoqing**"Workflow Tasks Scheduling Optimization Based on Genetic Algorithm in Clouds**"2018 the 3rd IEEE International Conference on Cloud Computing and Big Data Analysis
2. ShuangYINPengKELingTAO**"An improved genetic algorithm for task scheduling in cloud computing"**201813thIEEE Conferenceon IndustrialElectronics andApplications(ICIEA)
3. RUBA ABU KHURMA, HEBA AL HARAHSHEH, AHMAD SHARIEH (2018),**"Task Scheduling Algorithm In Cloud Computing Based On Modified Round Robin Algorithm**",Journal of Theoretical and Applied Information Technology 15th September 2018. Vol.96. No 17
4. Jinsong Zhang , Hua Zhang "**The application of improved genetic algorithm in resource scheduling in cloud computing**"Revista de la Facultad de Ingeniería U.C.V., Vol. 31, N°11, pp. 291-300, 2016