



HYBRID GREEN SCHEDULING ALGORITHM USING GENETIC ALGORITHM AND PARTICLE SWARM OPTIMIZATION ALGORITHM IN IAAS CLOUD

K. M. Uma Maheswari, Sujata Roy and S. Govindarajan

Computer Software Engineering Department, SRM University, Kattankulathur, India

E-Mail: umamaheswari.km@ktr.srmuniv.ac.in

ABSTRACT

Cloud computing is outsourcing of computing resources over the Internet where we can be connected to remote locations and can use the services over the Internet at another location to store our important information. The cloud service requirements provide access to advanced software applications. In cloud computing, the network of remote servers is used to process data. Workflow scheduling is one important issue in cloud. Scheduling of workflows is an NP complete problem. For NP complete problems, traditional scheduling algorithms do not provide optimal solution in polynomial time. In this paper a hybrid workflow scheduling algorithm is discussed for IaaS cloud environment. For simulation of the algorithm, WorkFlowSim simulator, an extension of CloudSim simulator, has been used. We have used varying types of workflows and it has been observed that the hybrid algorithm gives better result than the traditional PSO algorithm.

Keywords: cloud computing, workflowsim, cloudSim, PSO algorithm.

1. INTRODUCTION

In cloud environment, number of tasks and resources are very large and different resources have different computing capabilities. There are several cloud service models namely Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). Users who do not have sufficient resources, are provided with Infrastructure as a Service (IaaS) by public clouds. The IaaS cloud is considered to be flexible because users have to make payment for the resources as per their requirements. There are three types of IaaS cloud: Public IaaS cloud, Private IaaS cloud and Hybrid IaaS cloud.

Scientific workflows manage the problem of excessive complexity of scientific experiments and applications. Workflows are expressed as sequence of operations having some dependencies among them. Generally, the number of tasks in a workflow is several millions. Workflow applications are generally represented as a directed acyclic graph. Previously, workflow tasks were executed on High Performance Computing (HPC) infrastructure. But HPC infrastructure turned out to be time consuming and not cost effective. But, the cloud platform is more efficient in terms of cost-potency, extensibility and quantifiability. In virtualization platform, one physical machine contains several Virtual Machines (VMs) for sharing of resources. Resource sharing leads to better resource utilization and improves QoS. Workflow scheduling problem is said to be a notable instance of DAG scheduling problem. The DAG scheduling problem is an NP-Complete problem. For NP-Complete problems, traditional scheduling algorithms do not provide optimal solution in polynomial time. A meta heuristic is a favorable approach to identify a sufficiently good solution for an optimization problem. In the proposed work, two global search meta heuristic algorithms are used- Particle Swarm Optimization Algorithm and Genetic Algorithm.

2. THE BASIC PARTICLE SWARM OPTIMIZATION ALGORITHM

a) Background

Particle Swarm Optimization (PSO) is a nature inspired algorithm. The development of this algorithm is motivated from the social behavior of bird flocking or fish schooling. PSO algorithm was originally proposed by Doctor James Kennedy and Russell Eberhart in 1995. The basic idea of PSO is, when one of the members discovers a better path to the destination, others will follow the same path. This method helps to improve the search strategy to find the best path. Let us suppose a group of birds are searching for food in an area. The birds either move together or scattered before reaching the desired location. Among all the birds, there is one bird that can accurately locate the source of food, having better food resource knowledge. While searching for the food, the birds transmit the information. After that all the birds will gather at the food's location. The birds' movement can be compared to the growth of the solution swarm, and the food location can be compared to the best solution till the search ends.

b) Working principle

PSO algorithm starts with randomly initialized particle positions and velocities. In every iteration, each generation of those particles update their values to get the optimal value. Each particle updates two values- pbest and gbest. The best remembered particle solution calculated till that point of time, is called pbest, and gbest is the best remembered value of any particle in the swarm that is calculated till that point of time. Through each iteration, gbest converges to the goal. After finding the two best values, the particle updates its velocity and positions. The algorithm for PSO is as follows-

**I) Initialization:**

a) Randomly generate initial solutions(particles) of population.

b) A particle is assigned a random 'velocity'.

II) Repeat:

a) For each particle:

1) Evaluate fitness value

2) Compare the fitness value with the pBest

3) If the fitness value is better than pBest, update the present value to be new pBest

End

b) For each particle:

1) Search for the particle with the best fitness in the particle neighborhood

2) Compute velocity of the particle

3) Update position of the particle

End

While termination condition is not satisfied

Initial Proportion and velocity values are generated as random values. Fitness will be calculated for the initial proportion. Then initial proportions are updated in iterations.

Weight:

$w = w_{upperbound} - [(t/\max_iteration) * (w_{upperbound} - w_{lowerbound})]$;

where t is iteration number

Update Velocity:

$newVel = (w * velocityProportion) + (r1 * C1) * (pBestProportion - PreviousProportion) + (r2 * C2) * (gBestProportion - PreviousProportion)$;

Update Proportion:

$newProportion = PreviousProportion + newVel$;

Fitness:

$Fitness = f(\text{execution time, bandwidth})$

$= w1 * (1 - \text{execution time}) + w2 * \text{bandwidth}$

w1 and w2 are weight factors

Selection criteria for VMs are execution time and bandwidth. Three categories of VMs are there with the three different bandwidth and MIPS units.

pBest:

The pBest value for the swarm is updated in every iteration.

$pBest = fitness(swarmA-i1)$

If($fitness(swarmA-i2) > pBest$) {

$pBest = fitness(swarmA-i2)$

}

i1, i2 are the iterations 1, 2.

gBest:

The gBest value is updated from the fitness of all swarms from multiple iterations

$gBest = fitness(swarmA)$

If($fitness(swarmB) > gBest$) {

$gBest = fitness(swarmB)$

}

VM that provides less execution time and high bandwidth will provide high fitness for the task to be scheduled.

3. A COMPARISON BETWEEN GA AND PSO

If we observe some evolutionary algorithms, we can conclude that they share some common characteristics:

(1) Initially, system generates some random solutions of an initial population and searches for optima.

(2) Calculate a fitness value for every particle.

(3) Observe the fitness values and then reproduce the population accordingly.

(4) If termination condition is satisfied, then stop.

(5) Else return to 2.

After observing the above steps, we can conclude that PSO has some similar characteristics of GA.

(1) GA and PSO algorithms begin with a random population.

(2) Both use fitness value to analyze the population. But genetic operators like crossover and mutation are not applied by PSO algorithm.

(3) Both the algorithms update their current population to search for optimum value.

(4) Both the algorithms do not guarantee success.

4. PROPOSED HYBRID GENETIC AND PSO ALGORITHM

The Genetic Algorithm was introduced by John Holland in the mid 1970s. The Genetic Algorithm is based on genetics and evolution. The proposed hybrid PSO and Genetic Algorithm, the algorithm starts with PSO operations. The Genetic Algorithm operations are performed on the outcome of the PSO operations. The disadvantage of particle swarm optimization (PSO) algorithm is that it easily tends to succumb to local optima in high-dimensional space, also having a low convergence rate. The proposed hybrid algorithm overcomes this problem. The flowchart of the proposed algorithm is as follows-

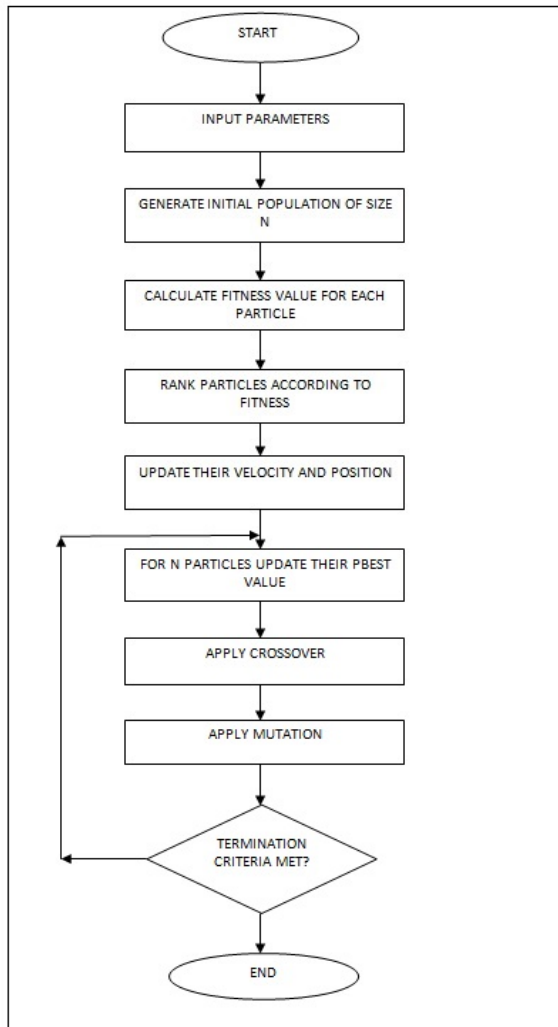


Figure-1. Flowchart of the proposed algorithm.

During scheduling if the PSO based VM is in busy state, in such case, the VM that is in idle state will be selected for scheduling instead of PSO based VM to reduce the scheduling delay. The operations are performed as follows-

VM for the task = PSO based VM

If (PSO based VM is busy) {

List of Idle VMs are considered as chromosomes for the cloudlet having the VM selected based on PSO

Number (6) of chromosomes are considered for GA

Each chromosome is composed of genes represents VM ID
Crossover is applied between two parent chromosomes randomly.

Mutation is applied followed by crossover

Fitness is computed for every chromosome after mutation

Execution time = task length/MIPS of VM

Fitness is indirectly proportional to execution time

High fitness chromosome is selected for scheduling

}

5. PERFORMANCE EVALUATION

WorkflowSim framework has been used in the simulation. Four different workflows have been used- Montage, Inspiral, CyberShake and Epigenomics. WorkflowSim is an extension of CloudSim having workflow level support for simulation. We have created a proper cloud environment with the simulator so that we can experiment with the algorithm there. The workflows are generated as DAX files. The graphs in the Figure-2 and Figure-3 shows the comparative analysis of traditional PSO and the proposed Hybrid GA and PSO algorithm.

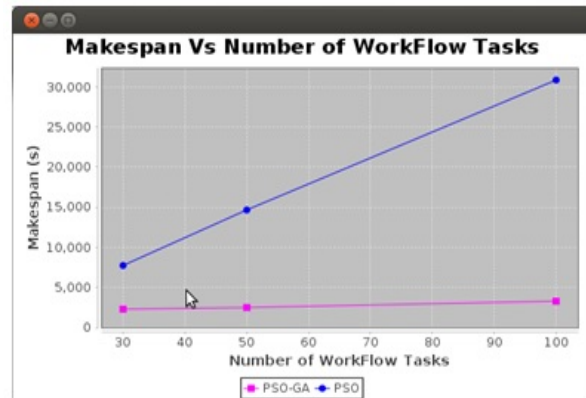


Figure-2. Performance analysis of makespan vs number of workflow tasks.

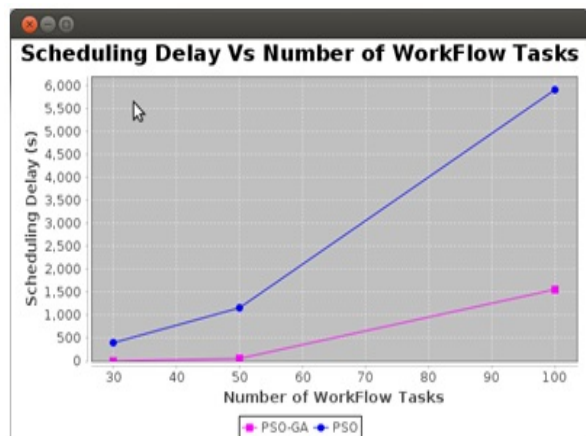


Figure-3. Performance analysis of scheduling delay vs number of workflow tasks.

In the following table, we have shown the comparative analysis of traditional PSO and hybrid PSO-GA algorithm. The makespan values for different types of workflows are compared and it is observed that the proposed algorithm performs better than the traditional algorithm.



Table-1. Comparison of makespan with traditional PSO and proposed HybridPSO-GA algorithm.

Workflow Type	Makespan with traditional PSO algorithm	Makespan with proposed hybrid PSO-GA algorithm
CyberShake_30	2150.7	972.3
CyberShake_50	5010.24	1117.32
Epigenomics_24	29444.33	29444.33
Epigenomics_46	180533.65	48569.38
Epigenomics_100	1683130.7	297418.06
Inspiral_30	20260.5	8567.01
Inspiral_50	46906	12442.84
Inspiral_100	80418.19	11763.67
Montage_25	800.69	381.69
Montage_50	2184.27	501.78
Montage_100	3737.82	622.16

6. CONCLUSIONS

In this paper, the hybrid GA-PSO workflow scheduling algorithm has been proposed for executing workflow applications on IaaS cloud. The comparative analysis of traditional PSO and the hybrid GA-PSO algorithm has been shown. The make span and scheduling delay for each workflow type is observed and we can conclude that the hybrid algorithm gives better result than the traditional PSO algorithm.

ACKNOWLEDGEMENTS

The author would like to thank her guide/mentor Mrs K. M. Uma Maheshwari, Assistant Professor, Faculty of Computer Science Engineering, SRM University for guiding and encouraging her throughout this study.

REFERENCES

- [1] <http://www.swarmintelligence.org/tutorials.php>
- [2] <http://mnemstudio.org/particle-swarm-introduction.htm>
- [3] <https://en.wikipedia.org/wiki/Metaheuristic>
- [4] <http://www.workflowsim.org/>
- [5] S. Jaya Nirmala, S. Mary SairaBhanu, Catfish-PSO based scheduling of scientific workflows in IaaS cloud, 14 June, 2016.
- [6] KeivanBorna, RaziehKhezri, A combination of genetic algorithm and particle swarm optimization method for solving traveling salesman problem,18 June,2015.
- [7] <https://doi.org/10.1080/23311835.2015.1048581>
- [8] MuraliManohar Sharma, AnjuBala, Survey Paper on Workflow Scheduling Algorithms Used in Cloud Computing, International Journal of Information & Computation Technology, ISSN 0974-2239 Volume 4, Number 10 (2014), pp. 997-1002
- [9] Dhanalakshmi. M, Dr. Anirban Basu, Energy Efficient Task Provisioning for Distributed Cloud Networks using Meta Heuristic Multidisciplinary Technique, 2016, ISSN 0973-4562 Volume 11, Number 9, pp 6195-6199.
- [10]Awad.A.I, El-Hefnawy.N.A and Abdel_kader.H.M, Dynamic Multi-objective task scheduling in Cloud Computing based on Modified particle swarm optimization, ACSIJ Advances in Computer Science: an International Journal, Vol. 4, Issue 5, No.17 , September 2015, ISSN: 2322-5157
- [11]Gang Zhao, Cost-Aware Scheduling Algorithm Based on PSO in Cloud Computing Environment, International Journal of Grid and Distributed Computing Vol.7, No.1 (2014), pp.33-42
- [12]TalwinderKaur, SeemaPahwa, An Upgraded Algorithm of Resource Scheduling using PSO and SA in Cloud Computing, International Journal of Computer Applications (0975-8887), Volume 74- No. 8, July 2013
- [13]Shaobin Zhan, HongyingHuo, Improved PSO-based Task Scheduling Algorithm in Cloud Computing, Journal of Information & Computational Science 9: 13 (2012) 3821-3829
- [14]LizhengGuo, Guojin Shao, Shuguang Zhao, Multi-objective Task Assignment in Cloud Computing by Particle Swarm Optimization, In Proceedings of 8th International Conference on Wireless Communications, Networking and Mobile Computing, 2012, 1-4.
- [15]Guo, Zhao, Shen, Jiang, Task Scheduling Optimization in Cloud Computing Based on Heuristic Algorithm, Journal Of Networks, Vol. 7, No. 3, March 2012, 547-553.
- [16]Nuttapong Netjinda, Booncharoen Sirinaovakul, Tiranee Achalakul, Cost Optimization in Cloud



- Provisioning using Particle Swarm Optimization. In Proceedings of 9th IEEE International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2012, 1-4.
- [17] Dian Palupi Rini, SitiMariyam Shamsuddin, SitiSophiyati Yuhani, Particle Swarm Optimization: Technique, System and Challenges, International Journal of Computer Applications (0975-8887), Volume 14- No. 1, January 2011
- [18] Pandey, Wu, Guru, Buyya, PSO based Heuristic for Scheduling Workflow Applications in Cloud Computing Environments. In Proceedings of 24th IEEE International Conference on Advanced Information Networking and Applications, 2010, 400-407.
- [19] Liu, Luo, Zhang, Zhang, Li, Job Scheduling Model for Cloud Computing Based on Multi-Objective Genetic Algorithm, IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 3, January 2013, 134-139.
- [20] Jang S.H., Kim T.Y., Kim J.K., Lee J.S., The Study of Genetic Algorithm-based Task Scheduling for Cloud Computing, International Journal of Control and Automation Vol. 5, No. 4, December 2012, 157-162.
- [21] Chang-Tian and Jiong, Energy-aware Genetic Algorithm for Task Scheduling in Cloud Computing, In Proceedings of Seventh IEEE ChinaGrid Annual Conference, 2012, 43-48.
- [22] Zhu K., Song H., Liu L., Gao J., Hybrid Genetic Algorithm for Cloud Computing Applications, In Proceedings of IEEE Asia-Pacific Services Computing Conference, 2011, 182-187.
- [23] Guo-ning G., Ting-Iei H. and Shuai, Genetic Simulated Annealing Algorithm for Task Scheduling based on Cloud Computing Environment, In Proceedings of IEEE International Conference on Intelligent Computing and Integrated Systems, 2010, 60-63.
- [24] Zhao C., Zhang S., Liu Q., Xie J., Hu J., Independent Tasks Scheduling Based on Genetic Algorithm in Cloud Computing, In Proceedings of 5th IEEE International Conference on Wireless Communications, 2009, Networking and Mobile Computing, 1-4.
- [25] R. S. Mohana, "A Position Balanced Parallel Particle Swarm Optimization Method for Resource Allocation in Cloud", Indian Journal of Science and Technology, 2015 Feb, 8(S3), Doi no: 10.17485/ijst/2015/v8i1/60501
- [26] Perla Ravi Theja, S. K. KhadarBabu, "An Evolutionary Computing based Energy Efficient VM Consolidation Scheme for Optimal Resource Utilization and QoS Assurance", Indian Journal of Science and Technology, 2015 Oct, 8(26), Doi no: 10.17485/ijst/2015/v8i26/77179