A STUDY OF VARIOUS META HEURISTIC ALGORITHMS FOR SCHEDULING IN CLOUD

Aishwarya.V⁵, Subramanian E.R.S ⁶
¹M.Tech-CSE, School of Computing, SASTRA University, Thanjavur
²Associate Dean-CSE, ³,⁴,⁵,⁶B.Tech-CSE, School of Computing, SASTRA University, Thanjavur

Abstract: Scheduling plays a vital role in Cloud Computing. The way the tasks are assigned to the available resources determines the overall performance and utilization efficiency. Traditional scheduling algorithms even though easier to implement doesn’t work well for problems with higher complexity. Metaheuristic algorithms are based upon usage of operators such as transition, evaluation and determination for providing efficient scheduling strategies. The evaluation of Metaheuristic algorithms are done based on computation time, resource utilization, response rate and scheduling cost. In this paper a brief review of existing metaheuristic algorithms with their advantages and disadvantages.

Keywords: Cloud Computing, meta-heuristics, scheduling.

1. Introduction

Cloud computing systems provide a better way to carry out the submitted tasks in cloud. Submitted tasks are to be scheduled properly in the cloud. There are various scheduling algorithms used in cloud computing. Due to the immense development in cloud technologies, the scheduling of workflow in cloud is a challenging one. The problems that occur in scheduling can be, mapping of task-resource; various requirements; resource provisioning on the basis of demand; performance variation and fault tolerant scheduling of resource, effective transmission of data and storage.

2. Traditional scheduling

A main objective function of scheduling problem is to assign a number of tasks to a specified number of machines satisfying the requirement of enhancing certain factors. The problem in scheduling is mainly because of the problem of single machine scheduling. In the case of scheduling in two/more machines it is referred to as multi-processor scheduling problem. The primary roles involved are overall creation of workflow, latency, slowness, time of flow, and their alternatives which can be commonly used to study and estimate scheduling algorithms outcomes. The detailed explanation of further methods is given in literature [1].

The scheduling problems can be studied and compared based on the following parameters such as description of jobs (which includes time for processing, date of release, last date and load), number of machines and other factors. To express and distinguish the scheduling problems these constraints are considered. The notations α, β, γ[6], explain the problems in scheduling where α represents a single machine and parallel machine. The features of processing and constraints, that is sequence-dependent or sequence independent is referred using notation β, γ represents overall execution or the delayed jobs number. The scheduling was classified as five levels based on the length of time in [5] long range, middle range and short range planning.

3. Meta-heuristics scheduling algorithms

3.1 Framework

To search for the possible solutions on process, heuristics uses three operators. They are Transition, Evaluation, and Determination. In Metaheuristic algorithm [6], the solution is iterated until tmax which is the stop criteria or maximum number of iterations. The Transition operators create the possible solutions, Evaluation operators measure the fitness of solutions and the Determination operators determine the next search directions solution.

3.2 Scheduling algorithms

3.2.1 Metaheuristics based on single solution

a. Hill Climbing

A popularly used algorithm is Hill Climbing (HC), a greedy and iterative algorithm [6]. It illustrates a common framework and the design of metaheuristics consists of tasks and edge parameters in the algorithm specify two techniques: they are transition and evaluation of Hill Climbing. Determination mechanism of Hill Climbing is performed which compares the existing solution with new one. In the next iteration, the starting point of better search direction is assigned.
b. Simulated Annealing

The Simulated Annealing algorithm is generally designed to show that an iterative greedy algorithm can be a metaheuristics one by making a small modification as suggested by Kirkpatrick and Cern[4].

c. Tabusearch

Tabusearch was presented by Glover [2,3] to evade determining the same solutions repeatedly. The solutions that are recently visited are kept recorded into the tabu list to solve this problem. Similar to Hill Climbing and Simulated Annealing, the Tabusearch algorithm is initiated as a single outcome, it then tries to determine the current solution neighbor as the next one, and the assigned new solution should not be in the list of tabu. The tabu list will be added with the new solution, if it is accepted it will be in the list of tabu, if not the next new solution is replaced.

4. Scheduling on cloud

4.1 Problems on cloud related to scheduling

On cloud, scheduling is the main element. Even though many studies give a correct definition and explanation of cloud scheduling, there exists different expressions. A general method to express the cloud scheduling problem is directed acyclic graph where it consists of vertices and edges. To minimize the construction cost of cloud applications and complexity for designing the cloud environment, Cloudsim (a simulation tool)[7] is used to design and develop better algorithms for scheduling.

4.2 Metaheuristics scheduling algorithms

Meta-heuristics algorithms give an effective way to find the fine solution. The commonly used meta-heuristics algorithm for task scheduling is genetic algorithm (GA). Other heuristics algorithms such as simulated annealing (SA), particle swarm optimization (PSO) and ant colony optimization (ACO) are also studied and discussed later in this paper.

4.2.1 Genetic algorithm scheduling algorithms

The Genetic Algorithm constructs an initial solution which has a impact on process and speed. The initialisation can be done in two phases they are test and computing fitness phases. These phases are mentioned to build a fine initial set of Genetic algorithm solutions. Min-Min, Round robin are few examples of heuristics scheduling algorithm that are designed to build the initial candidate solutions first and then solutions to be passed after certain tests.

4.2.2 Ant colony optimization algorithm for scheduling

The general aim of Ant colony optimization is to design the solution by using the pheromone matrix. In that matrix, each node can be expressed as a computer node in a network. Because of this, Ant Colony Optimization is very suitable for imitating the real network status. In cloud computing, to resolve the problem of scheduling, expected time of execution matrix is used in many studies. The result of ACO is assigning the tasks to the machine based on the scheduling. ACO is used in the cloud computing to solve scheduling problem that occurs on the convergence process where various resources for computation can be considered at the same time, such as usage of CPU, storage, and bandwidth of the network.

4.2.3 Particle Swarm Optimization (PSO) scheduling algorithms

PSO uses the directed acyclic graph in workflow application to represent the tasks similar to the Genetic Algorithm. Because PSO is generally designed for continuous problems but various other techniques are designed to express the discrete problems solutions. Primary problem is that the transition operator has to be redesigned in order to satisfy the requirements based on the priority. Next, main problem is to encrypt the outcome of particle swarm optimization scheduling. To encrypt each particle a set of tasks and machine pairs is the general technique. The task and machine pair represents the assignment of task T to machine M. In PSO, each particle is encrypted by applying n by m matrix.

4.2.4 Hybrid metaheuristics algorithms for scheduling

A single algorithm is formed by combining two or more metaheuristics algorithms which is the basic design of hybrid metaheuristic algorithm to influence the advantages of all the basic algorithms in order to improve the quality of the time for computing. In hybrid metaheuristic scheduling algorithms, various combinations are used. They are combining single solution based algorithm with population based algorithm, combining metaheuristic algorithm with other heuristic algorithm and combing two population based algorithms.

4.3 Discussion

The metaheuristics algorithms are mainly concentrated on three functions. They are changing the operators, changing the value of fitness, and hybrid metaheuristics. Changing the operators concentrates on designing the operators of transition by appending the transition operators of transition in the scheduling algorithms. These factors have an impact on search technique of the real meta heuristic algorithm. The outcome can be, cost of data transferring, cost of
communication and computation, the gain, and the consumption of energy which are included to the fitness function. Hybrid metaheuristics is used by other scheduling algorithms to optimize the original algorithm performance.

5. Conclusion
Cloud computing has many functions and limitations of scheduling problems. This paper focuses on the study of various Metaheuristic algorithms. These scheduling algorithms are widely used on cloud. Though scheduling problems on cloud are not new, several problems arise in scheduling algorithm to satisfy the future requirements. In order to enhance the performance and satisfy the new constraints, a new scheduling algorithm considering the various level of services such as IaaS (Infrastructure as a service), PaaS (Platform as a service) or SaaS (Software as a service) has to be considered. By combining other Metaheuristic algorithms a new algorithm has to be developed to enhance the performance.

Acknowledgments
The authors wish to express their sincere thanks to the Department of Science & Technology, New Delhi, India (Project-ID:SR/FST/ETI-371/2014) and SASTRA University, Thanjavur, India for extending the infrastructural support to carry out this work.

References


