A Novel ACO algorithm with job selection technique for Scheduling in Grid

Shashank Girepunje,
Assistant Professor,
Department of Computer Science and Engineering,
SRIT II, Raipur, India.

Abstract : Grid computation is new emerging area for researchers. The computing capabilities of grid system is comes from heterogeneous resources in order to support and manage complicated computing problems. Job scheduling in computing grid is a complex issue to be solved. Current scientific applications need huge computing power and storage space. However, to utilize computational grid efficiently, require an efficient job scheduling algorithm for proper management of resources in grids. In this paper, we propose a efficient Ant Colony Optimization (ACO) algorithm for job scheduling in the Grid environment. Using ant colony optimization with job selection approach, we can improve the efficiency of grid scheduling policy. In our approach, ant colony optimization and TLBO (teacher learning based optimization) is considered. The job selection is performed using TLBO method and the resource allocation is done through ant colony optimization.

Index Terms: **- ACO, TLBO, Job scheduling.**

I. INTRODUCTION

In present decade the scientific problems are becoming more complicated and these complex problems need large processing capabilities and storage media. The other technologies such as parallel computing or distributed system are less efficient in the term of resource utilization and throughput. Grid computing is evolving technologies for solving enterprise problems. The word grid is employed for a platform in which different organizations or domains are connected and sharing their processing resources for providing set of services to the end entities. User can use this platform as a utility for accessing computational resources for specific applications. This new distributed platform assists organizations to solve complex problems that cannot be solved by single desktop. It can also be named as a kind of parallel computing that relies on devices, computers and data archives connected through network.

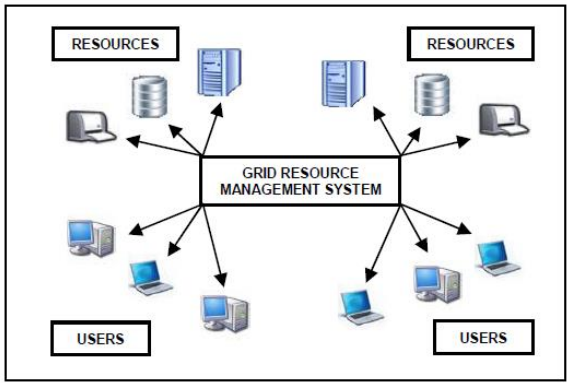

Figure 1: Grid Environment

In grid computing system does not follow centralized structure concept like server-client architecture. This infrastructure enables the integrated collaborative utilization of high end computers, scientific instruments and databases managed and owned by multiple organizations [11]. In Grid, resources are autonomous and heterogeneous. In Grid platform, resource management and job scheduling are very crucial task because a resource can leave or join the network at any time. An effective grid resource management with good job and resource scheduling algorithm is needed to manage the grid computing system (refer Fig. 1). The algorithm must consider the dynamically changes conditions in grid environment because the computational performance changes from time to time, networks connections may become unreliable, resources may join or leave the system at any time and resources may become unavailable without any notifications.

In this paper, we examine the grid scheduling technique ant colony optimization in detail. One type of ACO called multi criterion ant colony optimization and TLBO approach is also explained. Next section describes some of the related work on ACO algorithm and grid scheduling.

II. RELATED WORK

The objective of grid scheduling is to utilize the available resources in the distributed environment. Grid scheduler collects resource status information, selects appropriate resources, and determines the best schedule for the jobs to be executed for maximize performance. Because of the dynamic nature of grid environment, developing scheduling algorithms are always major challenges for researchers. The scheduling must be done in a way to maximize the resource utilization. In Grid system, scheduling decisions must be made in the shortest makespan, user competing for resource allocation, and time slots desired by users could be dynamic. The goal of any task scheduling is to achieve minimum makespan and suitable resource for each job so that system can satisfy the constraints that are imposed by the user. The Grid scheduling problem has been taken to be a NP-hard optimization problem.

Grid computing system submits jobs to available resources and also determines best resource for particular job. Best resources in term of processing speed, memory and availability status are more likely to be selected for the submitted jobs during the scheduling process [1]. Best resources are categorized as optimal resources. In a research by [5], Ant Colony Optimization (ACO) has been used as an effective algorithm in solving the scheduling problem in grid computing.

An ant colony optimization for dynamic job scheduling in grid environment was proposed by [10] which aimed to minimize the total job tardiness time. The process to update the pheromone value on each resource is based on local update and global update rules as in ACS. ACO algorithm found efficient as compared to First Come First Serve, Minimal Tardiness Earliest Release Date techniques and Minimal Tardiness Earliest Due Date.

In another [2] paper the author proposed an enhanced ant colony optimization algorithm for grid system. This algorithm is has taken main idea from ant colony algorithm for scheduling efficiently and it also considers the current load in available resources in grid environment. The objective of the algorithm is to distribute the balanced load and to obtain minimum make-span. This algorithm defined a framework for scheduling using the current instant value of resources. In our work, Job selection technique is done at initial state of algorithm and then the ant is applied to selected jobs.

Grid computing composes distributed computational resources to create a secure environment to accomplish a specific computational goal. Current scientific applications become more complicated and need huge computing and processing. [3] In this paper, the author has proposed a Balanced ACO algorithm for job scheduling in the Grid environment. There are two new schemes introduced regarding pheromone update. The main objective of our work is to balance the entire system load and minimize the makespan of a given set of jobs.

Dr. R. Venkata Rao, the author has proposed a teacher learning based optimization algorithm for solving complex problems in industrial environment. [9] Researchers have shown a lot of interest in algorithms considering natural behavior of the social species. TLBO is one of the recent population based algorithms which consider the teaching-learning process of the class room. TLBO does not require any algorithm-specific control parameters for problems. In this paper, elitism concept is introduced in the TLBO algorithm and its effect on the performance of the algorithm is shown.

III. GRID SCHEDULING

It is very important to understand difference between Grids computing environments, which is of type distributed computing, from other distributed computing environments. We first like to present brief introduction about types of distributed computing environments that are widely used for solving problems in distributed manner in the current era of network-computing.

Grid computing enables utilization of idle time of resources that are available at geographically diverse locations [9]. Grid computing can allow access of resources such as storage, sensors, application software/code, databases, and computing power. In Grid, resources are autonomous and heterogeneous. Present Grid computing, in most of the implementations, is on collaborative manner, in which resources expose varying availability. However, QoS oriented Grid computing is also possible. Grid computing has been used in drug discovery, GIS processing, sky image processing, industrial research, and scientific experiments. Usage of Grid infrastructure is done by moving traditional infrastructure into Grid applications. The independent tasks applications include Parameter Sweep and Task farming (embarrassingly parallel) problems [10], e.g. drug discovery. Dependent tasks applications include workflow applications, e.g., Montage.

In grid system, there are two sides; first are the set of heterogeneous resources and the group of grid users looking for the execution of job to be done. So the aim of scheduling is to manage both sides with the maximum efficiency. Grid scheduling is crucial to provide efficient and coordinated use of heterogeneous grid resources. Because of the dynamic nature of the grid, scheduling is significantly complex due to the difference in performance goals by various grid applications (users) and grid resources. Most grid systems use grid scheduler responsible for resource allocation and available resources, by selecting the most suitable resources that provide user's request, and allocating the task onto selected resources.

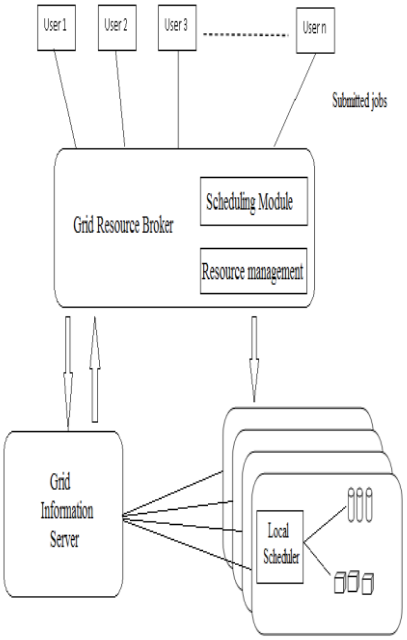

Figure 2: Grid scheduler

A lots of scheduling algorithms introduced for homogeneous and dedicated resources such as computer clusters.The scheduling algorithms (e.g. First Come First Serve, Shortest Job First, Backfill and so on) have a complete control over all computing nodes and the overall information. Thus, the scheduler can allocate set of jobs onto the cluster efficiently. However these algorithms cannot be suitable and work well for today's heterogeneous environments such as Grid. Grid must concern with large-scale heterogeneous resources across different domains. In such a dynamic distributed computing environment, resources availability varies dramatically. Task scheduling in Grid environments is significantly more complex than the conventional scheduling systems.

In grid environment, we need to consider the conditions like network status and resources status. If the network or resources are dynamic, jobs would be failed or the total computation time could be very large. Scheduling systems must take distant properties of both different Grid applications and various Grid resources into account. Therefore, scheduling becomes more challenging in this distributed environment. The dynamic behaviour of the grid systems makes scheduling policies even harder. Because the resources that has located at different location varies in terms of computation power and availability. The network that combine thses resoutce together having dynamic bandwidth.So the grid scheduler must concern behaviour of the system. So when any resource available for the processing, the scheduler must adopt it and when any resource become an available it must provide rescheduling for the jobs. These challenges create significant obstacles o\in the path effective scheduling system for Grid environments.

There have been some popular techniques that has been used for scheduling problems i.e. genetic algorithm, tabu search and ACO etc. in grid environment. A good scheduling policy must concern the changing status of the entire environment and the types of jobs. Therefore, a dynamic algorithm in job scheduling such as Ant Colony Optimization (ACO) is more suitable for grid environment. Aim of the scheduling algorithms is to minimize make-span, idle time of the available resources, turn-around time and meet the specified application deadlines.

IV. PROBLEM STATEMENT

Grid computing is an environment with dynamic behavior, inefficient scheduling affects the computing and utilization of computational resources. The crucial factor in grid scheduling is that, it cannot control all the jobs in grid system. The dynamic versatility of grid resources and the difference between the expected time and the actual time of execution in algorithm are the other challenges. The main goal of the grid scheduling is assigning the jobs to the available resources efficient way. The suitable match should be assigned from the list of resources available and the list of jobs looking for allocation. In order to solve the scheduling problem, many algorithms like ACO, Genetic algorithm and simulated annealing have been approached. The scheduling algorithm MCACO (multi criterion ant colony optimization) has been proposed. It is an extension to ACO algorithm .ACO algorithms have been applied to many real world application problems.

V. PROPOSED SOLUTION

In the past few years, researchers have proposed scheduling algorithms for parallel system and distributed system. However, the problem of grid scheduling is still more complex solutions. Therefore, this issue attracts the interests of the large number of researchers. Current systems of grid resource management was surveyed and analyzed based on classification of scheduler organization, system status, scheduling and rescheduling policies.

At the current scenario, job scheduling on grid computing is not only aims to find an optimal resource to improve the overall system performance but also to utilize the existing resources more efficiently. Recently, many researchers have been studied several works on job scheduling using heuristic algorithms. Heuristics algorithms find solutions among all possible ones, but they do not guarantee that the best will be found. Some of those are the popular heuristic algorithms, which have been developed, are min-min, the fast greedy, tabu search and an Ant System. The intention behind of the scheduling algorithms is to minimize make-span, high utilization of the available resources, turn-around time and meet the specified application constraints.

The scheduling optimization problem is very complex because the scheduling is a prominent component for organizations using grid system. An efficient scheduling must obtain the performance criteria for enterprises. Developed scheduling algorithms only consider an instant computing speed of the resource at the scheduling time, and assume this value remains same during the processing time. A more accurate scheduling policy must consider the dynamic changes during the processing time. A scheduling algorithm Ant Colony Optimization algorithm is defined that evaluates the quality of a schedule in computational grid. The algorithm shows an efficient technique in term of achieving high throughput.

Using ant colony optimization with job selection approach, we can improve the efficiency of grid scheduling policy. In our approach, ant colony optimization and TLBO (teacher learning based optimization) is considered. The job selection is performed using TLBO method and the resource allocation is done through ant colony optimization.

Proposed Grid scheduling Algorithm:

- Let n is the no. of jobs (j1, j2, and j3… jn).
- Let m is the no. of resources (r1, r2… rm).
- Compute the pheromone indicator value.
- For each resource obtain the information like bandwidth, computing capacity and current load from GIS.
- For each job obtain the job size and the time needed to complete the job.
- Create grid matrix for the job process and create ants.
- Generate the initial population of job and apply the TLBO method for selecting the job selection.
- Initially estimate the mean parameter mg for each job. The mean parameter of each job is given as:

$$Mg = [m1, m2, m3, \dots\dots\dots\dots md]$$

The job with minimum value of the objective function is considered as teacher.

$$Xnew = Xi + rand * (Xteacher - (Tf * mg))$$

Where Tf is teaching factor which is the probability value. The value of this factor is not giving as input in the algorithm.

- The selected no. of jobs is taken input for the ant colony algorithm. Create grid matrix for the process. Apply ants in dedicated fashion.

- Calculate the local pheromone and set the process priority order for the completion of job. The pheromone value of each job is computed by following formula.

$$Pij = \left[\left[\frac{JSi}{Bandwidth\ r} + \frac{Cj}{MIPS * (1 - Tj)}\right]\right]$$

Where Pij is the pheromone value of job with respect to resource. JSi is the value of size of the job, Cj is the computation time of job, Tj is current load and bandwidth is the available bandwidth for transferring job.

- In pheromone matrix, the highest value of pi is selected and resource i process the job for execution.
- The local updation is done for remaining jobs by calculating pi matrix again.
- After the execution of jobs the global updation is done.
- Again select population from ant and repeat the process until all jobs are processed.

The key idea of selection operator is to give preference to better individuals by allowing them to the pass through selection process and prohibited the entrance of worst fit individuals; here we are using TLBO approach to only select the job not to find the solution for scheduling policy.

VI. CONCLUSION:

This paper presents an approach for solving gird scheduling based on improved ant colony algorithm. An improved version of ACO algorithm based on selection procedure of jobs that is based on selection procedure of TLBO approach and pheromone matrix is calculated to transfer job. From our experimental results, the proposed system is more effective than the ACO algorithm in terms of convergence speed and the ability to finding better solutions.

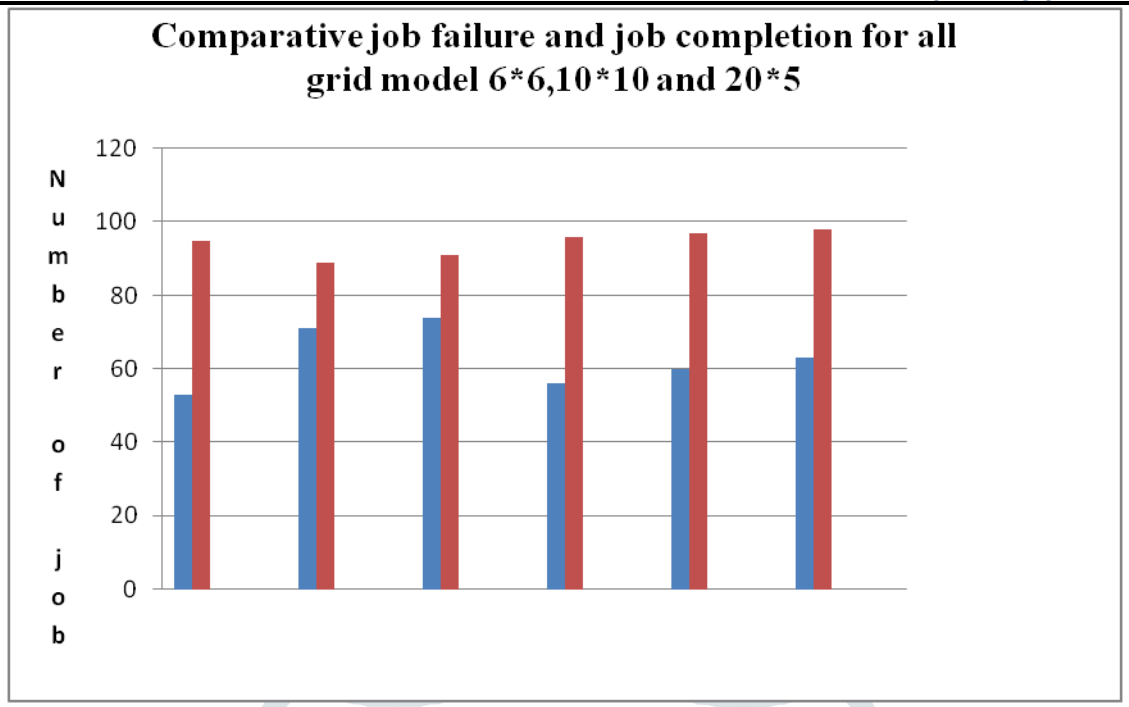

Figure 3: the comparative result analysis of all grid model

REFERENCES

[1] V. Naik, P. Garbacki, K. Kummamuru, and Y. Zhao, "On-line evolutionary resource matching for job scheduling in heterogeneous grid environments," *Proceedings of the 12th International Conference on Parallel and Distributed Systems(ICPADS'06), 2006*.

[2] Ruay Shiung Chang, Jih Sheng Chang and Po Sheng Lin, "Balanced Job Assignment Based on Ant Algorithm for Computing Grids", IEEE Asia-Pacific Services Computing Conference, 2007.

[3] S. Lorpunmanee, Sap, M.N., Abdullah, A.H., Chompoo-inwai, C.: "An Ant Colony Optimization for Dynamic Job Scheduling in Grid Environment." International Journal of Computer and Information Science and Engineering, 2007, pp 207–214.

[4] S. Fidanova and M. Durchova, "Ant algorithm for grid scheduling problem," *Lecture Notes in Computer Science*, vol. 3743, pp. 405-412, 2006.

[5] M. Dorigo, L.M. Gambardella, "Ant colony system: A cooperative learning approach to the travelling saleman problem", IEEE Transations on Evolutionary Computation, 1197, pp.53-66.

[6] Suresh Chandra Satapathy, Anima Naik and K Parvathi, "A teaching learning based optimization based on orthogonal design for solving global optimization problems", Springer open journal 2013.

[7] Venkata Rao and Vivek Patel, "An elitist teaching-learning-based optimization algorithm for solving complex constrained optimization problems", International Journal of Industrial Engineering Computations, 2012.

[8] S. Lorpunmanee, M. Sap, A. Abdullah, and C. Chompoo-inwai, "An ant colony optimization for dynamic job scheduling in grid environment," *International Journal of Computer and Information Science and Engineering*, vol. 1(4), pp. 207-214, 2007.

[9] D. W. Erwin and D. F. Snelling, "Unicore: A grid computing environment," in Euro-Par 2001 Parallel Processing. Springer, 2001, pp. 825–834.