ISSN No. 0976-5697

Volume 8, No. 5, May – June 2017

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Comparative Study of Multi Objective Task Scheduling using Metaheuristics in Multi-Cluster Environment

Er. Vishal Mehra
Department of Computer Engineering and Technology
Guru Nanak Dev University Amritsar, India

Er. Amit Chhabra
Department of Computer Engineering and Technology
Guru Nanak Dev University Amritsar, India

Abstract: This paper emphasises the study of various algorithms inspired by the nature in order to search for the optimal solutions when multiple criterions are to be considered. There are certain situations witnessed every time when a perfect solution is hard to be obtained the problems called NP Problems. In such situations no single solution or set of solutions as in case of multi-objective optimization can be considered as final result. Thus it is required to consider the solution/set of solutions those best fulfil our requirements. The author puts emphasis on study of Flower Pollination Algorithm for scheduling purpose only. The allocation of processors is based on non deterministic manner as per the availability of processors.

Keywords: Task Sequencing, multiple clusters, multi-objective optimization, scheduling, meta-heuristics, multi cluster allocation

I. INTRODUCTION

A **cluster** is a group of multiple server instances, spanning across more than one node, all running identical configuration. All instances in a cluster work together to provide high availability, reliability, and scalability.

The situations when compute intensive tasks/jobs are to be executed on multiple clusters scattered geographically can be used in order to maximize performance as well as reliability. For this Task scheduling is done to decide the task sequence for processor allocation. But Processor allocation is concerned with the assignment of the required number of processors to the tasks. Both are crucial to the overall performance of entire system[5].

Energy consumption is an important issue for development of computational systems in order to meet present and future needs. In the past few years, energy-efficient scheduling has received much attention for both single-processor and parallel-processor environments. As power consumption is convex function of processor speed. As higher speed allow faster execution, but at the same time result in higher energy consumption[3].

1.1Scheduling

Scheduling is the process of deciding the sequence in which processes will be executed and after that processes/tasks are allocated processors i.e. first job sequencing is performed and then processor allocation is done.

There are various criterions that can be stressed upon while scheduling the tasks as per requirements:

1. CPU Utilization
2. Throughput
3. Turnaround time
4. Wait Time
5. Load Average
6. Response Time

1.2Multi Cluster Allocation/CO Allocation

The jobs are scheduled by the scheduler. First the scheduler tries to find the cluster that can fulfill the resource requirements of the job. If it finds such cluster then it allocates the processors to the job.

If there is no such cluster in the environment that can fulfill the demand then scheduler CO Allocates the job by assigning processors from different clusters such that the required total number of processors are allocated to the job and thus it can be executed.

If the scheduler does not find adequate number of processors among all the clusters the process has to wait till the jobs which were allocated the processes earlier finish their execution and relinquish the processes.

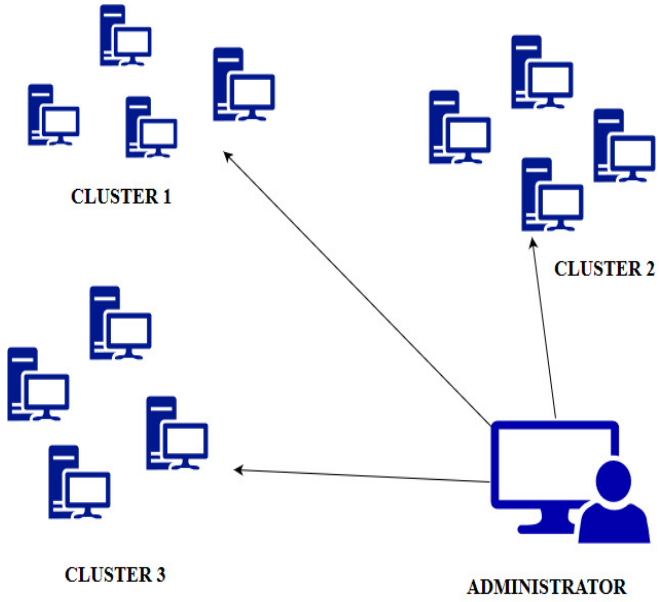
The allocation of jobs among various clusters sometimes raise dependency issues and high inter-process communication hence resulting in degraded performance

1.3 Optimization

In computer science optimization stands for deliberate modification in behavior and functioning of computer systems in order to produce the optimal results as per requirements[4].

When the optimization is performed there is no possibility of obtaining the optimal results always. There may be the situations where we may produce the most optimal or best suited solution. But there might be the cases when we are not able to optimal solution rather w obtain worst solutions.

In case of Multi-Objective optimization problems situations get even worsened. There are always the cases when we are able to obtain the best possible solution for one objective function but the same solution is not the optimal for another objective. For example, we can reduce computation time of processes by allocating them processors with highest computation speeds but contrary to that it also raises the consumption of electricity which is a big issue in current scenario and is a concern for future.



II.

Fig. 1. Multi Cluster Allocation

III. NATURE INSPIRED META-HEURISTICS

Most conventional or classis algorithms are deterministic. But in case of those problems which possess discontinuity and are cannot be solved with deterministic algorithms. For example, well known Newton-Raphson Algorithm is gradient based and works extremely well for smooth unimodal problems. However, if there is some discontinuity in objective function, it does not work well. In such cases non-gradient algorithm is preferred.

2.1Flower Pollination Algorithm

The flower pollination algorithm[1] (FPA) was developed by Xin-She Yang in 2012, inspired by the pollination process of flowering plants. The FPA can be easily extended for purpose of multi-objective optimization. The following four rules are used:

1. Biotic and cross-pollination can be considered processes of global pollination, and pollen-carrying pollinators move in a way that obeys Levy flights .

2. For local pollination, abiotic pollination and self-pollination are used .

3. Pollinators such as insects can develop flower constancy, which is equivalent to a reproduction probability that is proportional to the similarity of two flowers involved.

4. The interaction or switching of local pollination and global pollination can be controlled by a switch probability $p \in [0, 1]$, slightly biased toward local pollination.

2.2Lion Optimization Algorithm

Lion Optimization Algorithm[6] was developed by Maziar Yazdani in year2015. Lion Optimization Algorithm is based on special lifestyle of lions and their cooperation. In Lion Optimization Algorithm(LOA),an initial population is formed by a set of randomly generated solutions called Lions. Some of the lions in the initial population(%N) are selected as nomad lions and rest population (resident lions)is randomly partitioned into P subsets called prides. S percent of the pride's members are considered as female and rest are considered as male, while this rate(sex rate (%S)) in nomad lions is vice-versa. For each lion, the best obtained solution in passed iterations is called best visited position, and during the optimization process is updated progressively. In LOA, a pride territory is an area that consists of each member best visited position. In each pride, some females which are selected randomly go hunting. Hunters move towards the prey to encircle and catch it. The rest of the females move toward different positions of territory. Male lions in pride, roam in territory. Females in prides mate with one or some resident males. In each pride, young males are excluded from their maternal pride and become nomad when they reach maturity and, their power is less than resident males.

2.3Dolphin Swarm Algorithm

Dolphin Swarm Algorithm[7] was developed by Tian-Qi Wu in year 2015.Dolphins have many biological characteristics and living habits such as echolocation, information exchanges, cooperation, and division of labor.

The Dolphin performs the job in six phases:

1. Initialization Phase

2. Search Phase

3. Call Phase

4. Reception Phase

5. Predation Phase

6. Termination Phase

2.4Elephant Search Algorithm

Elephant search algorithm[8] was developed by Suash Deb in year 2015.ESA emerges from the hybridization of evolutionary mechanism and dual balancing of exploitation and exploration. The design of ESA is inspired by the behavioral characteristics of elephant herds; hence the name Elephant Search Algorithm which divides the search agents into two groups representing the dual search patterns. The male elephants are search agents that outreach to different dimensions of search space afar; the female elephants form groups of search agents doing local search at certain close proximities.

2.5Genetic Algorithm

A Genetic Algorithm[12] (GA) is an evolutionary algorithm which involves search and optimization. Genetic algorithms were first used by J.H. Holland. It mimics the process of natural evolution so as to create artificial processes for a “clever” algorithm in order to find the

solution of complex problems like job scheduling in computational grid.

IV. ALGORITHM

Flower Pollination Algorithm (or simply Flower Algorithm)

Objective min or max $f(x)$, $x = (x_1, x_2, ..., x_d)$
Initialize a population of n flowers/pollen gametes with random solutions
Find the best solution g_* in the initial population
Define a switch probability $p \in [0, 1]$
while (t <MaxGeneration)
  for $i = 1 : n$ (all n flowers in the population)
    if rand < p,
      Draw a (d-dimensional) step vector L from a Lévy distribution
      Global pollination via $x_i^{t+1} = x_i^t + \gamma L(g_* - x_i^t)$
    else
      Draw ϵ from a uniform distribution in [0,1]
      Do local pollination via $x_i^{t+1} = x_i^t + \epsilon(x_j^t - x_k^t)$
    end if
    Evaluate new solutions
    If new solutions are better, update them in the population
  end for
    Find the current best solution g_*
end while
Output the best solution found

Figure 2. Flower Pollination Algorithm

V. LEVY FLIGHTS

The Flower Pollination Algorithm is Levy Flight based search criterion. A **Levy flight**, named for French mathematician Paul Levy, is a random walk in which the step-lengths have a probability distribution that is heavy-tailed. When defined as a walk in a space of dimension greater than one, the steps made are in isotropic random directions. The term "Levy flight" was coined by Benoit Mandelbrot, who used this for one specific definition of the distribution of step sizes. He used the term **Cauchy flight** for the case where the distribution of step sizes is a Cauchy distribution, and **Rayleigh flight** for when the distribution is a normal distribution (which is not an example of a heavy-tailed probability distribution).

VI. PARETO FRONT

Pareto efficiency, or Pareto optimality[1], is a state of allocation of resources in which it is impossible to make any one individual better off without making at least one individual worse off.

It becomes much more difficult to solve the problems when multiple objectives are required to be fulfilled. In multi objective optimization problems we obtain a set of solutions , called Pareto Optimal set.

To map the Pareto front accurately is very time-consuming, and there is no guarantee that these solutions points will distribute uniformly on the front.

Figure 3. Levy Flight

VII. RELATED WORK

Over last few years Nature Inspired Algorithms have gathered huge interest of researchers across all the branches of computer science. As the problem size increases most of the algorithms tend to fail and are no longer able to provide optimal solution. So to fulfill the needs according to the problem set researchers find a great interest in Nature Inspired Algorithms. Beauty of Nature Inspired Algorithms lies in the Hybridization of these algorithms which further helps to simplify the research phase and even helps to derive better results.

- Xin-She Yang[1] proposed that the multiobjective optimization problems require multiobjective optimization techniques to solve, and it is often very challenging to obtain high quality Pareto Fronts accurately.

- Hosseini,et al.[2] proposed a new problem solving method called "intelligent water drops" or IWD method which is based on the processes that occurs in the natural river systems and the actions and reactions which take place between water drops in the river and the changes that happen in the environment that river is flowing.

- Xibo Jin[3] has proposed an optimal scheduling algorithm for the case when all of the tasks have uniform computational work. In the

reasearch it is presented that a polynomial-time algorithm that achieves a bounded approximation factor when the tasks have arbitrary-size work.

- Kuo-Chan Huang[5] has proposed most-fit policy based solution to reduce resource fragmentation occurrences and evaluates it with several existing processor allocation policies.

- Eloi Gabaldon et al.[9] suggested that after the free sources happen to be configured, completely new possibilities occur intended for cutting down power utilization by giving ideal matching of parallel programs to free computing nodes.

- Eloi Gabaldon et al. [10] proposed a genetic algorithm for organizing job-packages connected with parallel jobs for source federated environments. With regards to proposition seemed to be to determine the task schedule and also offer allocation to enhance the application operation and also process throughput.

- Rahmani. A. M et al. [12] presented the Genetic Algorithm in order to solve the scheduling problem of dependent jobs in which the population quantity and the number of generations are dependent upon the number of tasks. In this the number of iterations are stable for any task and it offers the advantage if the number of tasks are less, so long computational time was not used and if the number of tasks are large the probability of finding the appropriate solutions is provided by the further repetition of algorithm. Also, SA was used to decrease the time of calculations.

- Blanco et al. [11] presented a new scheduling method that allocated multiple jobs from the system queue simultaneously on a heterogeneous multi-cluster, by applying co-allocation when it is necessary. This method was composed by a job selection function and a linear programming model to find the best allocation for multiple jobs. The proposed scheduling strategy has shown to reduce the execution times of the parallel jobs and the overall response times.

Table I. Meta-Heuristic Algorithms

Sr. No.	Algorithm	Operators	Control Parameters
1.	Flower Pollination Algorithm	Pollinators	Scaling Factor Proximity Probability
2.	Lion Optimization Algorithm	Mating	Population Size Nomad Lions Working Lions
3.	Dolphin Algorithm	Sound	Population Size Acceleration Directions
4.	Elephant Search Algorithm	Male Search Agents Female Search Agents	Probability Factor Population Size Life and Death Possibility
5.	Genetic Algorithm	Crossover Mutation Selection Inversion	Population Size Max Generation Number Cross-Over Probability Mutation Probability

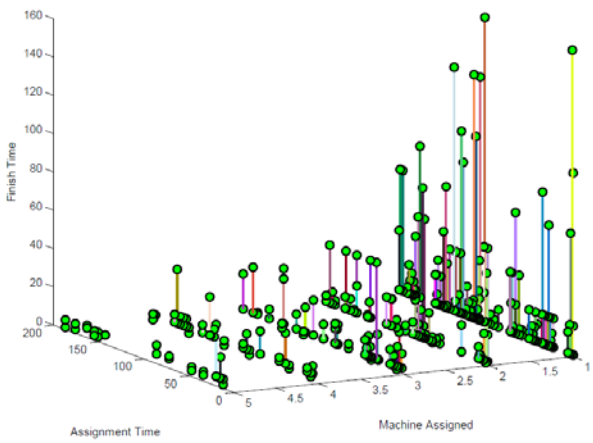


Figure 4. 3-D Representation of Schedule

VIII. CONCLUSION

Most of the research work done for task scheduling is in the case sequential processing systems. For scheduling jobs in Multi-Cluster Environment Flower Pollination Algorithm has been applied. Also, its convergence speed is slow. So, in future hybridization of flower pollination algorithm with some other meta-heuristic can be done to optimize the make span and flow time of parallel jobs.

IX. REFERENCES

[1] Yang, Xin-She, Mehmet Karamanoglu, and Xingshi He. "Flower pollination algorithm: a novel approach for multiobjective optimization." *Engineering Optimization* 46.9 (2014): 1222-1237.

[2] Hosseini, Hamed Shah. "Problem solving by intelligent water drops." *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*. IEEE, 2007.

[3] Xibo Jin, Fa Zhangb, Liya Fane, Ying Songc, Zhiyong Liu, "Fine Scheduling for energy minimization on restricted parallel processors, Journal of Parallel and Distributed Computing.

[4] Sapinderjeet Kaur, Amit Chhabra,"Survey of Various Meta-Heuristic Algorithms for Parallel Job Scheduling ".

[5] Kuo-Chan Huang, Kuan-Po Li, "Processor Allocation Policies for Reducing Resource Fragmentation in Multi-Cluster Grid and Cloud Environments " International Computer Symposium(ICS2010),2010.

[6] Maziar Yazdani, Fariborz Jolai, "Lion Optimization Algorithm: A Nature Inpired Meta Heuristic Algorithm," Journal. ofComputationalDesignandEngineering3(2016)24–36].

[7] Tian-Qi Wu, Min Yao,Jian-Hua Yang, "Dolphin Swarm Algorithm" Wu et al. / Front Inform Technol Electron Eng 2016 17(8):717-729.

[8] Suash Deb, Simon Fong, Zhonghuan Tian, "Elephant Search Algorithm for Optimization Problem" The Tenth International Conference on Digital Information Management (ICDIM 2015).

[9] Gabaldan, Eloi, et al. "Blacklist muti-objective genetic algorithm for energy saving in heterogeneous environments." The Journal of Supercomputing (2016).

[10] Gabaldon, Eloi, et al. "Particle Swarm Optimization Scheduling for Energy Saving in Cluster Computing Heterogeneous Environments." Future Internet of Things and Cloud Workshops (FiCloudW), IEEE International Conference on. IEEE, 2016.

[11] Blanco, Héctor, et al. "Multiple job co-allocation strategy for heterogeneous multi-cluster systems based on linear programming." *The Journal of Supercomputing* 58.3 (2011): 394-402.

[12] Rahmani, Amir Masoud, and Mojtaba Rezvani. "A novel genetic algorithm for static task scheduling in distributed systems." *International Journal of Computer Theory and Engineering* 1.1 (2009):