DOI: http://dx.doi.org/10.26483/ijarcs.v9i1.5470

ISSN No. 0976-5697

Volume 9, No. 1, January-February 2018

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

HYBRID BAT AND CUCKOO BASED META-HEURISTIC ALGORITHM FOR PARALLEL JOB SCHEDULING IN MULTI-CLUSTER ENVIRONMENT

Amandeep Kaur
Computer Engineering and
Technology Department
Lovely Professional University
Phagwara, India

Abstract: Scheduling in multi cluster environment is NP Hard but yet it is a fundamental to satisfying and improving the performance of parallel processing system. In this paper the parallel job scheduling problem is being addressed in multicluster computing environment using hybrid BAT+Cuckoo Algorithm. Scheduling policy is considered in which BAT algorithm is used to allocate the available resources and Cuckoo algorithm is used to look for constraint resources. Constraint resources may be located within distinct clusters.A mating pool in which resources from distinct clusters are maintained. In case resources are limited, Cuckoo algorithm look for available resources from each cluster and combine them within the mating pool. Parameters considered for evaluation includes Makespan and Flowtime which are substantially reduce using the proposed literature.

Keywords: BAT, Cuckoo, Scheduling, Mating pool, Cluster

1. INTRODUCTION

Various researchers have shown an intense interest in scheduling techniques. There are many algorithms for scheduling the jobs and then assigning those jobs to the resources or processors by using some criteria.[1] Some algorithms are bi-objective and others are multi objectives. In early stages only one objective function is optimized whereas latter two or more than two objective functions are optimized concurrently. The basic aim of algorithm is optimization. In any problem optimization means finding an alternative method or is a process of modifying a system to make some features work more accurately under some given constraints, as feasible by maximizing required parameters and minimizing the parameters.

Job scheduling is the way to allocate jobs to be executed by a single processor and even we can execute these tasks according to our choice. When tasks are scheduled on a multiprocessor system it is known to be NP-Complete problem. But the main issue behind this process is to minimize the execution time of a task. To achieve optimal result under different situations many heuristics have been developed. In our purposed paper we are using hybridization of bat and Cuckoo heuristic algorithms for taking better result and we also compare it with genetic algorithm. After comparison of make span and flow time we find that the make span of our hybridised algorithm is better than the genetic algorithm. [6]

Scheduling is a process which is used to schedule the tasks. These tasks can follow single cluster approach or multi cluster approach to achieve optimal result. In single cluster approach every job will be scheduled to execute on single cluster with similar types of serversBut in case of Multi cluster approach multiheuristic algorithms are used like ACO, BAT, Cuckoo and Firefly are used to execute the jobs.[8]

Proposed literature concentrate on Parallel jobs for resource allocation. In other words jobs are assigned to clusters with available resources.Problems starts to appear in case as job requirements increases beyond the scope of current cluster. In that case jobs must wait. Starvation problem starts to appear. Genetic algorithm convergence is at stake while evaluating such jobs. In order to resolve the problem, proposed literature uses hybrid approach of BAT and Cuckoo algorithm. Worth of study is proved considering Existing work in next section. [14]

Background Analysis

Scheduling is the most important issue which focuses on better performance and optimization is the basic parameter to achieve better performance. In scheduling every job get resources to achieve the objective or multiple objectives. In parallel computing scheduling is considered to solve NP hard problems because ittakes more time to find the optimal and best solution. Metaheuristic algorithm provides optimal solution in short span of time for certain kind of problems. These techniques can resolve large problems with more efficiency and effectiveness. The main aim to process the jobs parallel is to achieve better performance in less time as well as cost.(Yang and Hossein Gandomi 2012)

Heuristic and metaheuristic approaches mimic their behaviour according to biological systems like in [3] swarm algorithm flocks of birds to solve problems such as effective foraging for food, prey evading, or colony re-location[3] Ant Colony Algorithm with the behaviour as ants for finding food optimal path is achieved for certain problems, [18] BAT algorithm is inspired by the echolocation behaviour of bats to achieve global optimization, [10] Cuckoo algorithm inspired by the obligate brood parasitism of some cuckoo species by laying their eggs in the nests of other host birds[14] Firefly algorithm was inspired by the behaviour of firefly to find optimal distance.

Meta heuristic algorithms provides mechanism to allocate jobs that exceed resources availability using pre-emption of resources on completion of existing jobs known as effective finish time.

1.1 Genetic Algorithm

Genetic algorithm is one of the most commonly used Meta heuristic algorithms used to process complex jobs. Jobs processing under this algorithm have multiple objectives associated with it. Genetic algorithm has multiple phases associated with it. All of these phases are as explained below

Initialization

Jobs and resources at first place required to be initialized. Jobs are known as chromosomes. These chromosomes are initially selected randomly for allocation. Generations, objective function, crossover and mutation probabilities are initialized in this phase. [4]

Selection

In genetic algorithm, Jobs are referred to as Chromosomes. Resources are required by chromosomes. Chromosomes are selected initially for resource allocation. This process is known as selection. In genetic algorithm, this selection is performed using variety of mechanism known as selection criteria's. Chromosomes are selected for later breeding using this phase. Steps in selection are listed as under

1 Fitness function is evaluated associated with each job. Fitness values are normalized by adding the fitness values and then dividing it with total number of jobs.
2 Population is sorted according to descending value of fitness values
3 Population with highest fitness value is selected for mating.

Crossover or Mating

Selected Chromosomes are then go through this phase in order to determine chromosomes which are to be mutated. Uniform crossover is preferred in proposed thesis.

Mutation

Chromosomes selected for mating in crossover are mutated to generate new chromosomes and then evaluated again.

After performing all the phases, fitness values are analysed again. This process continues until desired level of fitness values is achieved or generations terminates.

1.2 Ant Colony Optimization

This is another metaheuristic approach used to schedule resources among jobs. Ant colony optimization is an extension of Genetic algorithm used to handle more complex jobs as compared to Genetic approach. Resources may be located within same or distinct clusters and these jobs are handled using local and global solutions. Results obtained are compared against the previous solution to obtain optimal solution. Iterations terminates as optimal solution is reached or iterations expires against the specified value [2]. Ant Colony optimization follows the following steps

1 All the Ants must visit all the resources at least once
2 Distant resource is less likely to be visited due to visibility problem.
3 In case resource is fond by Ant, pheromone is laid down.
4 After previous iteration, Pheromone evaporated.
5 Process continues until optimal solution gbest and lbest is found out or generations terminates.

Convergence problem exists while using Ant colony Optimization.

1.3 BAT Algorithm

A nature inspired BAT algorithm was developed by Xin-She Yang. To detect their chase and obstacles microbats have special quality i.e. "sonar". In the dark nights they can also detect the cracks for their nest. Microbats hear the sound pulse and wait for reverberate of sound that comes back after striking the chase or obstacles.[5] Microbats usually use pulse rate i.e. 10-20 times per second and can also detect tiny obstacles like human hair. Time delay between emissions and reverberate of sound is used for navigation and can also detect distance, type of chase, its moving speed.[13] By variations of Doppler affect introduced by the wing flutter rates of the target insects they are able to differentiate targets. To easily detect the chase and its navigation they combine their all senses. Micro bats uses [0.7,17] mm range of wavelength or [20-500] kHz inbound frequencies. Algorithm is developed using three rules which are given below:

1. To sense the distance, hunt, food and obstacles echolocation is used.
2.Bats usually used to fly randomly with some velocity at position Si with fixed frequency [feq max, feq min],varying wavelength □, pulse rate r[0,1] and loudness A.
3.Loudness varies from Amax to Amin. A[0,1].

Micro bats use varying loudness and frequency while their velocity, frequency and position remains fixed. Frequency can be varying corresponding to pulse rate and pulse emitted.

Movement of virtual bats:

For experiment, rules are updated with the frequency feq, position xi and velocity veli in d-dimensional search space. The corresponding updated solutions for SI^t and velocity veli tat time step t are represented as:

$feq_i=feq_{min}+(feq_{max}-feq_{min})U(0,1)$

$vel_i^t= vel_i^{t-1}+(x_i^t-best)feq_i$

$S_i^t= S_i^{t-1}+ vel_i^t$

Distribution preferred is random indicated through U[0,1]. Best value availed from U[0,1] is used to find the location of resources used within the job allocation criteria. The following equation is used to calculate current best solution:

$s_{new}=s_{old}+EA^t$

where E [-1,1] is a random number between -1 to 1, and At is used for average loudness of all the bats at time step t. Pseudo code is as under

1 Obtain Job Sequence(J1,J2,J3,-------,Jn)
2 Associate Objective Function with Each Job

$$F(y) = \{y_1, y_2, -----, y_n\}$$

3 Define range of BAT in terms frequency(F) and Loudness(A)
4 Repeat following Steps until i<Max_Iterations
Generate Solutions by adjusting Existing frequency
Update location of BATS in case solution found
If(Optimal_Solution==True)
Select this solution and replace it with existing solution
End of if

Randomize the BATS flying and relocate the resources.
In case optimal solution is located increase pulse rate and decrease loudness(A).
Rank the BATS and sort the solution.
End of loop
5 Output Solution and job Ordering(J)

1.4 Cuckoo Search Algorithm

This algorithm is also nature inspired optimization metaheuristic. It was introduced by Young and Deb in 2009. It gives guarantee for solving many hard real world optimization problems. The main attraction of the [11] cuckoos is reproduction strategy. It is a parasitism in which a cuckoo lays its eggs in the nest of host species. To increase the hatching probability of their own eggs some cuckoo species lay their eggs in other bird's nests and may remove the host bird's eggs.[15,12]As many host birds do not like intruders and conflict with them. However, the host bird will throw their eggs out or may simply abandon its nest and build a new nest at some other place. This algorithm is a population based global search stochastic[9.] Each egg retains possible solution in cuckoo search algorithm. Simplification of natural systems is required to successfully implement them by computer algorithms because natural systems are complex in nature. Three assumptions to simplify the cuckoo search algorithm is as follows:

1. At a time each cuckoo lays only one egg and can leave this egg in a randomly choosen nest.
2.To maintain the elitist property, the best nest with the highest quality of eggs will make it to the next generations.[13]
3.The number of available host nests to lay eggs is fixed in nature. It's egg can be identified by host bird with a probability pa□[0, 1]. The host bird can either throw it away or the host may abandon its own nest and builds a new nest at some other location if they analyse it. [17]

1 Initialize Jobs(J) for Execution along with number of resources available(R), y=0
2 Input number of generations G.
3 Define Objective function F(x).
4 Define initial population in terms of nest
5 Solution is in terms of eggs which must be assigned with rank examined through fitness values
6 Repeat while stopping criteria is matched or y<G
Y=y+1
Produce new solution using levy flight of cuckoo
Calculate Fitness value
Select the nest randomly
If (F_i<F_{i+1})
Replace current solution with the new solution z
End of if
Acesses fitness and sort the solution with maximum fitness value.
End of loop
7 Produce result in terms of Job Ordering

Existing literature suggest problems of convergence in case job requirements exceeded the available resources. It is possible that convergence never occurs and algorithm terminates without optimal solution. In order to overcome the problem hybridization of BAT and Cuckoo Algorithm is proposed. Next section describes the proposed system.

2. PROPOSED WORK

Proposed system hybridised BAT with Cuckoo to provide optimal solution to the complex job scheduling problem. BAT algorithm in this system is used to allocate the jobs to available resources and Cuckoo algorithm look for resources in distinct clusters in case resources are not available on contiguous basis. Mating pool is maintained to determine available resources from clusters. Mating pool is defined through following equation

$$Mating_{Pool} = \sum_{i=1}^{5} Cluster_i\_Machines$$

Equation 1: Mating Pool Configuration

Total of 5 clusters are considered for evaluation in current meta heuristic environment. Jobs are allocated to resources that consume resources. Number of resources required by jobs is then subtracted from Mating Pool. In case mating pool contains less resources then job requirements then Jobs must wait and early finish time of allotted jobs are examined. Mating pool in terms of job requirements is defined as under

$$Mating_{Pool} = Mating_{Pool} - \sum_{i}^{5} Jobs_{Requirments}$$

Equation 2: Job Allocation to Mating Pool

Parameter considered for evaluation are described as under

MAKESPAN

(Abdulal et al. 2012)Makespan is the total time consumed in executing all the jobs present within the system. In other words Makespan is the finish time of last job executed.

$$Makespan = \max(finished)$$

Equation 3: Makespan evaluation formula

FLOWTIME

[7] It is the time taken by individual job to be executed on machine. The Flowtime of a job consists of execution time and the waiting time.

$$Flowtime_i = \sum_{i=1}^{n} job\ (executiontime_i + waitingtime_i)$$

Equation 4: Individual Finish time of jobs indicating Flowtime

TIME CONSUMPTION OR EXECUTION TIME

Time consumption is total execution time required to execute the jobs presented to the system.

$$Time_{consumption} = Finish_i - Start_i$$

Equation 5: Time Consumption equation

WAITING TIME

This parameter defines the Jobs waiting time before allocation of resources to that job. Waiting time is given in terms of following equation

$$Waiting_{time} = Start_{time} - Arrival_{time}$$

Equation 6: Waiting time calculation

Job execution and ordering depends upon the fitness function calculations. Fitness function involves constant parameter α. Value of α is fixed to 0.5. Fitness function is evaluated using following equation.

$$F(x) = \alpha * \text{makespan} + (1 - \alpha) * \text{flowtime}$$

Equation 7: Fitness Function Evaluation

BAT algorithm is a meta heuristic algorithm used for global optimization. BAT algorithm look for resources from mating pool. In case resources are located Bat frequency is reduced and located resources parameter r is increased. As resources from mating pool decreases, Jobs may face starvation. To overcome the problem Cuckoo search is applied. Cuckoo search algorithm pre-empt the resources from finished jobs. Size of mating pool thus increase again. To prove worth of the study, proposed system is compared against existing Genetic algorithm. Terminology is described as under

Table 1: Analogy associated with considered approaches

GA	BAT	BAT+CUCKOO
1.Chromosomes=Number of Jobs 2.Mutation=Selection of Jobs with minimum Makespan and Flowtime for mating 3.Crossover=specify selection of parents for mating 4.Selection=Involves selection of jobs for allocation	1.BATs=Jobs 2.Prey=Resources 3.BAT_Frequency=Path toward resources 4.Gbest=Global Best solution consuming least time 5.Pulse emission Rate=Local Best solution which can change over time	1.BATs=Jobs 2.Pray=Resources 3.Gbest=Global Best solution consuming least time 4.Pulse emission rate=Local Best solution which can change over time 5.Rate=Speed with which resources are located 6.Distance=distance between jobs and resources

Parameters used within distinct algorithms have distinct characteristics. Algorithms considered for evaluation can be used in simulated environment because of their correlation in terms of parameters. In all the algorithms prime objective is to minimize Makespan, Flowtime and Waiting time. In order to accomplish this goal parameter tuning is done. In other words resources are varied along with mutation and crossover probabilities and results are obtained.

The proposed system utilized hybrid approach of BAT and Cuckoo algorithm using methodology given as under

3. METHODOLOGY

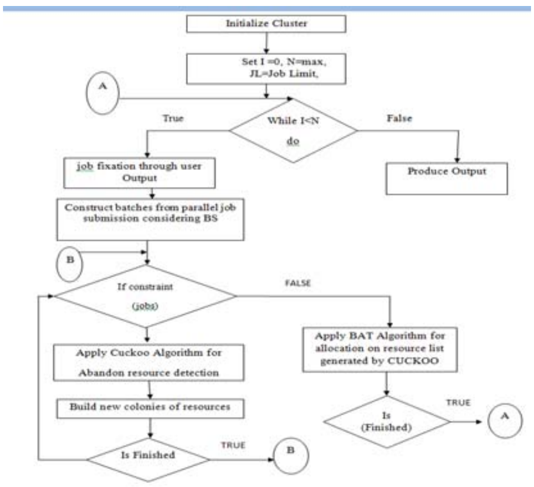

Figure : Proposed Methodology

Proposed methodology combine BAT and Cuckoo algorithm with the objective to minimize Makespan, Flowtime and Waiting time. Next section describes performance analysis of hybrid algorithm with genetic approach.

4. RESULTS

Results obtained through various strategies including GA, BAT and BAT+Cuckoo is analysed in this section. Parameter tuning is performed in terms of mutation probability, crossover probability, α , number of jobs and job requirements. Result obtained indicates proposed approach show better performance in terms of Makespan. Result show variation on increasing and decreasing value of α. Number of jobs and jobs requirements also impact Makespan of overall schedule. Values of alpha subsequently impact fitness value and Makespan results. In other words Makespan increases as α increase. Results in the form of plots is given as under

4.1 Results in terms of Makespan

The total time that finish the job from beginning to the end, is known as makespan. In other words the shortest possible time in whichjob scheduling is completed, is known as make span.

Make Span of Schedule is given below:

As number of processors increased for fixed set of jobs then the value of Makespan is decreasing.

1. MP=0.03 CR=0.4 number of jobs=100

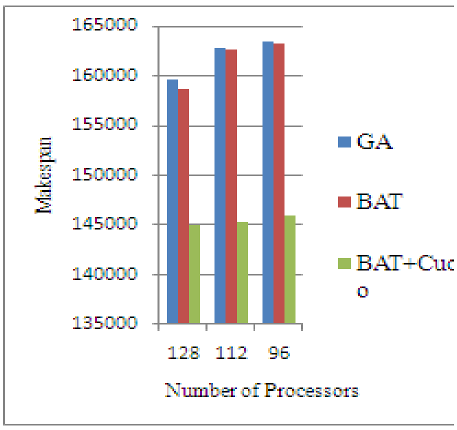

Figure 4.1 Makespan for 100 jobs

2. MP=0.03 CR=0.4 number of jobs=300

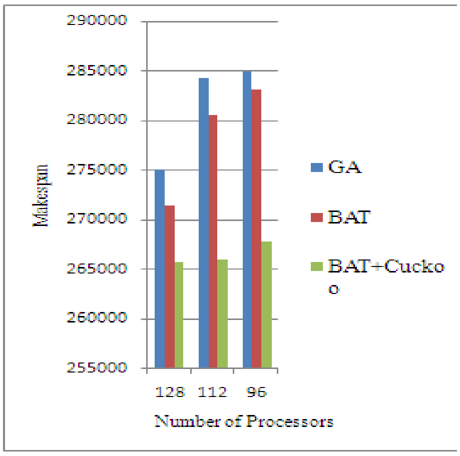

Figure 4.2 Makespan for 300 jobs

3. MP=0.03 CR=0.4 number of jobs=500

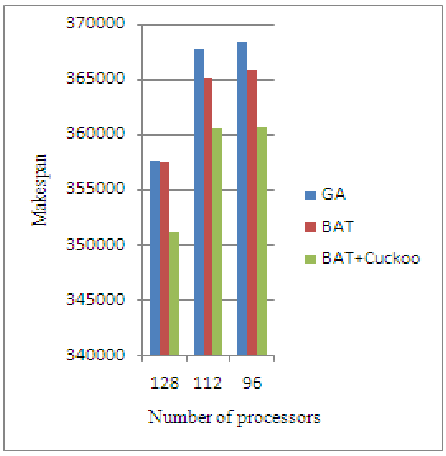

Figure 4.3 Makespan for 500 jobs

4. MP=0.03 CR=0.4 number of jobs=700

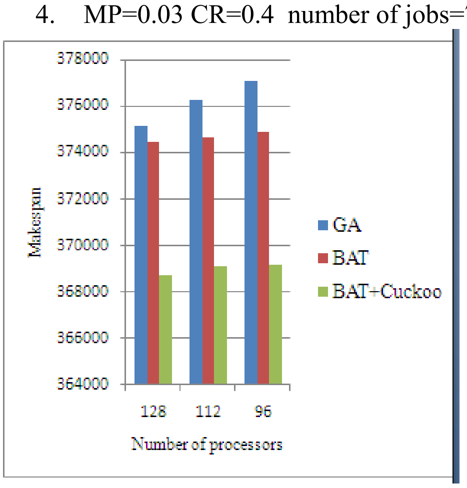

Figure 4.4 Makespan for 700 jobs

4.2 Results in terms of Flowtime

The time taken to complete a job is known as flow time.As the number of processors increased for particular set of jobs then theFlowtime is reduced.

5. MP=0.03 CR=0.4 number of jobs=100
6.

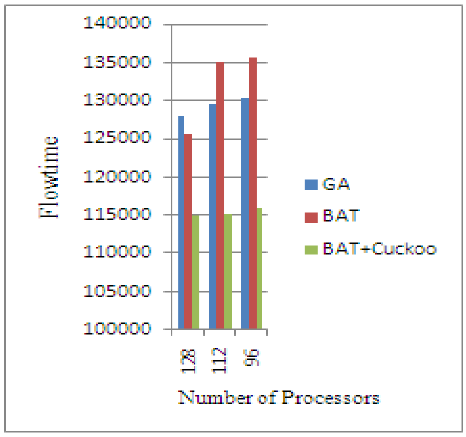

Figure 4.5 Flowtime for 100 jobs

7. MP=0.03 CR=0.4 number of jobs=300

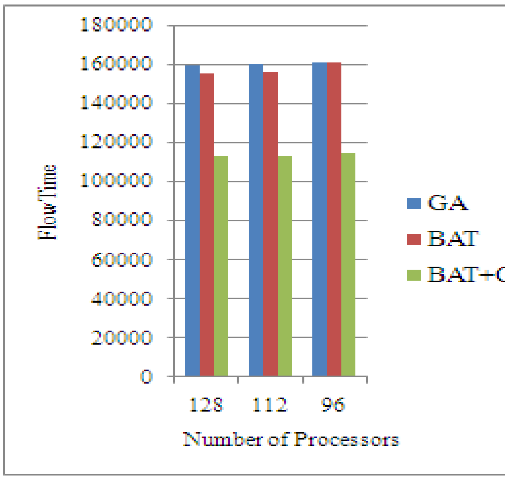

Figure 4.6 Flowtime for 300 jobs

8. MP=0.03 CR=0.4 NUMBER OF JOBS=500

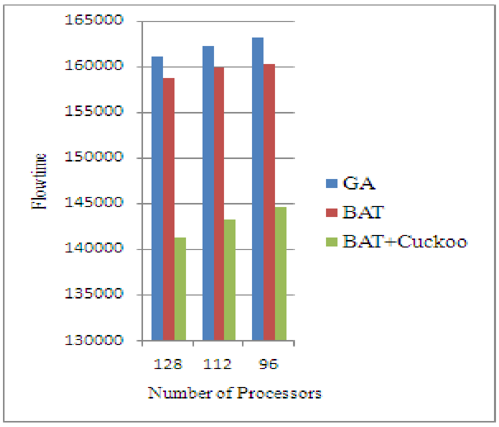

Figure 4.7 Flowtime for 500 jobs

9. MP=0.03 CR=0.4 number of jobs=700

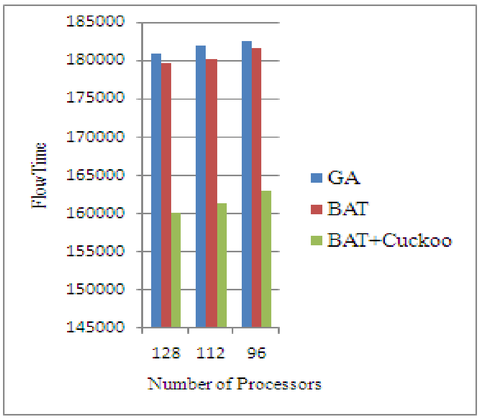

Figure 4.8 Flowtime for 700 jobs

4.3 Results in terms of Mean WT

The time interval for which a job is waiting for processor, is known as waiting time. As the number of processors increased for same set of jobs, then Mean WT reduced

10. MP=0.03 CR=0.4 number of jobs=100

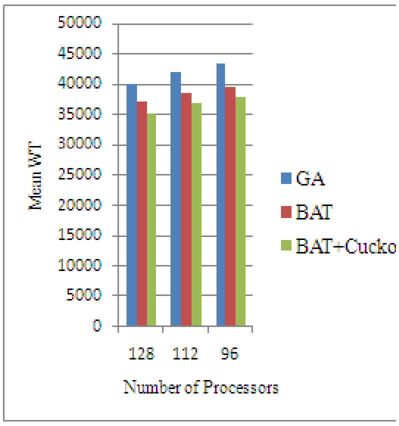

Figure 4.9 Mean WT for 100 jobs

11. MP=0.03 CR=0.4 number of jobs=300

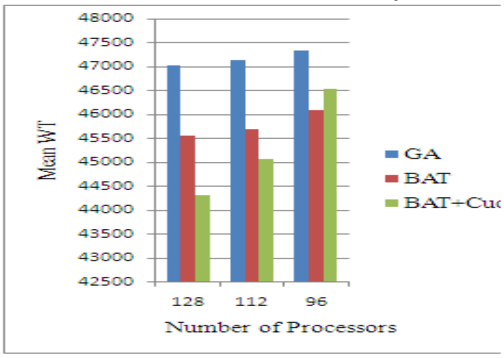

Fig: 4.10 Mean WT for 300 jobs

12. MP=0.03 CR=0.4 number of jobs=500

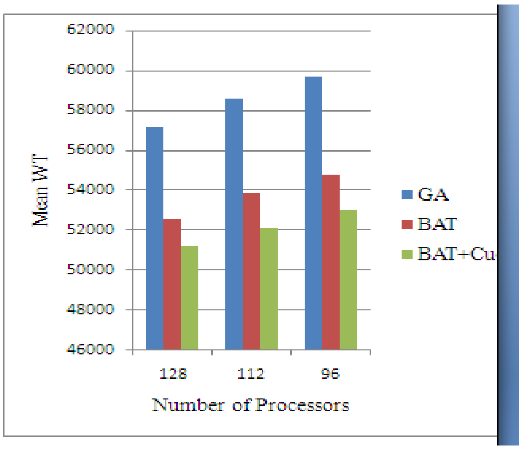

Fig: 4.11 Mean WT for 500 jobs

13. MP=0.03 CR=0.4 number of jobs=700

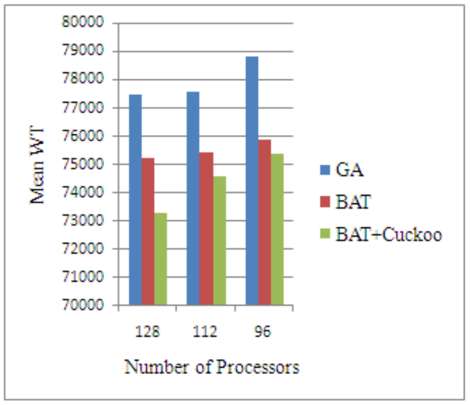

Fig: 4.12 Mean WT for 700 jobs

4.4 Results in terms of Normalization function

As the number of processors increased for particular set of jobs then the Normalization function is reduced.

14. MP=0.03 CR=0.4 number of jobs=100

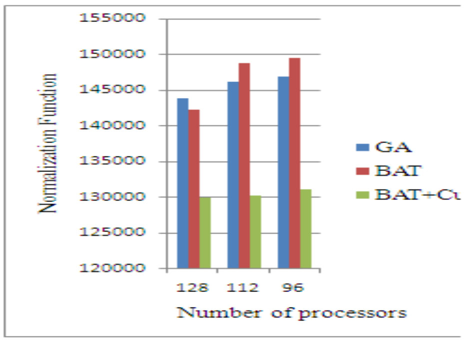

Figure 4.13 Normalization function for 100 jobs

15. MP=0.03 CR=0.4 NUMBER OF JOBS=300

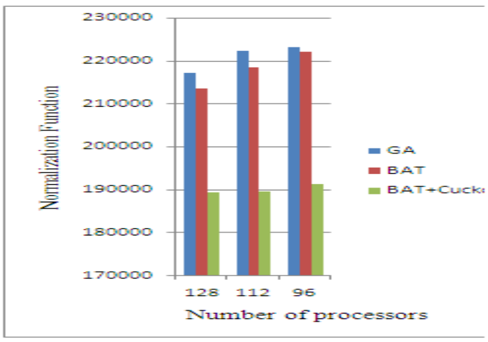

Figure 4.14 Normalization function for 300 jobs

16. MP=0.03 CR=0.4 NUMBER OF JOBS=500

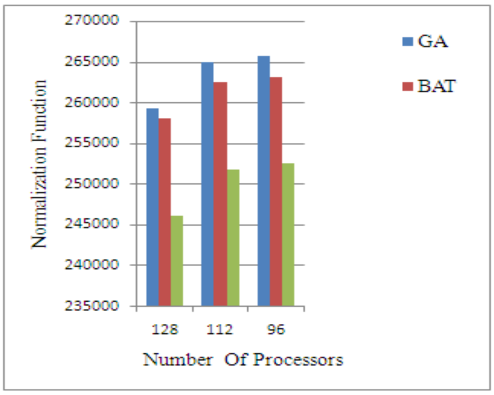

Figure 4.15 Normalization function for 500 jobs

17. MP=0.03 CR=0.4 number of jobs=700

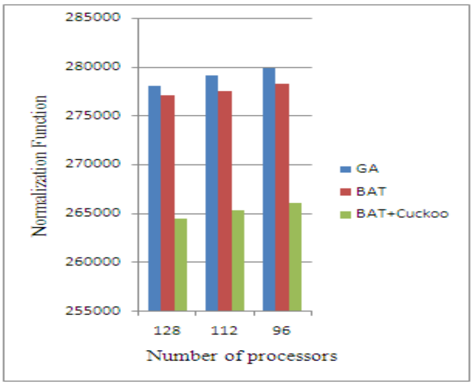

Figure 4.16 Normalization function for 700 jobs

5. CONCLUSION AND FUTURE SCOPE

The proposed literature uses hybrid approach of BAT and Cuckoo algorithm. The BAT algorithm is used in order to allocate the resources. The Cuckoo algorithm is an optimal searched algorithm. The BAT algorithm act first and locate the resources .In case resources are not available then Cuckoo algorithm will locate the resources. If the resources are not available then process must wait result in higher make span and flow time. The flow time and make span can be reduced further by merging the BAT algorithm with Cuckoo algorithm. Makespan is improved by 23% and Flowtime is improved by 36%. In future, the authors may hybrid some other algorithm to overcome the drawback of BAT algorithm. In future work, we aim to reduce the complexity of BAT andconsiderlargenumbersofjobstogether.We will also address the study of new optimization criteria, such as utilization, energy consumption, and some other performance metric criteria etc.

REFERENCES

[1] Abdulal, Wael, Ahmad Jabas, S. Ramachandram, and Omar Al Jadaan. 2012. “Task Scheduling in Grid Environment Using Simulated Annealing and Genetic Algorithm.”

[2] Adebisi, Adigun Abimbola, Adegun Adekanmi Adeyinka, and Asani Emmanuel Oluwatobi. 2015. “Performance Evaluation of Ant Colony Optimization and Genetic Algorithm for Facial Feature Selection.” 2(1):19–24.

[3] Al-Madi, Nailah, Ibrahim Aljarah, and Simone A. Ludwig. 2015. “Parallel Glowworm Swarm Optimization Clustering Algorithm Based on MapReduce.” IEEE SSCI 2014 - 2014 IEEE Symposium Series on Computational Intelligence - SIS 2014: 2014 IEEE Symposium on Swarm Intelligence, Proceedings 189–96.

[4] Allahverdi, A.; Gupta, J.N.D.; Aldowaisan, T. 1999. “A Survey of Scheduing Research Involving Setup Considerations.” OMEGA - International Journal of Management Science 27:219–39.

[5] Gigras, Yogita, Kusum Gupta, and Kavita Choudhary. 2015. “A Comparison between Bat Algorithm and Cuckoo Search for Path Planning.” 4459–66.

[6] Karatza, Helen D. and Ralph C. Hilzer. 2003. “Parallel Job Scheduling in Homogeneous Distributed Systems.” SIMULATION: Transactions of the Society for Modeling and Simulation 79(5):287–98. Retrieved (http://www.ingentaselect.com/rpsv/cgi-bin/cgi?ini=xref&body=linker&reqdoi=10.1177/0037549030 37148).

[7] Kołodziej, Joanna and Fatos Xhafa. 2010. “A Game-Theoretic and Hybrid Genetic Meta-Heuristics Model for Security-Assured Scheduling of Independent Jobs in Computational Grids.” CISIS 2010 - The 4th International Conference on Complex, Intelligent and Software Intensive Systems 93–100.

[8] Leinberger, William, George Karypis, and Vipin Kumar. 1999. “Job Scheduling in the Presence of Multiple Resource Requirements.” Performance Computing 1–14.

[9] Optimisation, Numerical. 2013. “A Comprehensive Review of Cuckoo Search□: Variants and Hybrids Iztok Fister Jr .*, Dušan Fister and Iztok Fister.” 4(4).

[10] Rabiee, Maryam. 2013. “Job Scheduling in Grid Computing with Cuckoo Optimization Algorithm.” International Journal of Computer Applications 62(16):38–43.

[11] Rajabioun, Ramin. 2011. “Cuckoo Optimization Algorithm.” Applied Soft Computing Journal 11(8):5508–18.

[12] Solanki, Sachin V. n.d. “Genetic Algorithm Approach for Implementation of Job Scheduling Problem.” 3867–72.

[13] Vijaya, Venkata, Geeta Pentapalli, and Ravi Kiran Varma P. 2016. “Cuckoo Search Optimization and Its Applications□: A Review.” 5(11):556–62.

[14] Xie, Jian, Yongquan Zhou, and Huan Chen. 2013. “A Novel Bat Algorithm Based on Differential Operator and Levy Flights Trajectory.” Computational Intelligence and Neuroscience 2013.

[15] Yang, Xin-she. 2011. “Bat Algorithm for Multi-Objective Optimisation.” 1–12.

[16] Yang, Xin-she. 2013. “Bat Algorithm□: Literature Review and Applications.” 1–10.

[17] Yang, Xin-she and Suash Deb. 2009. “Engineering Optimisation by Cuckoo Search.” 1–17.

[18] Yang, Xin-She and Amir Hossein Gandomi. 2012. “Bat Algorithm: A Novel Approach for Global Engineering Optimization.” Engineering Computations 29(5):464–83.