ISSN: 0976-3104
SPECIAL ISSUE: Emerging Technologies in Networking and Security (ETNS)
Sivakumar *et al.*

THE IIOAB JOURNAL

ARTICLE OPEN ACCESS

CLOUD WORKFLOW SCHEDULING ALGORITHMS USING CUCKOO SEARCH (CS) WITH NOVEL FITNESS FUNCTION

Sivakumar [1*], Chitra[2], Madhusudhanan[3]

[1]Dept.of Computer Science and Engineering, Mookambigai College of Engineering, Kalamavur, TN, INDIA
[2,3]Dept. of Computer Science and Engineering, Er. PerumalManimekalai College of Engineering, Hosur, TN, INDIA

ABSTRACT

Cloud Computing is an emerging technique. Certain parallel applications show a reduction in utilization of CPU resources when there is a rise in parallelism if jobs are not schedules correctly then it decreases the computer performance. Task scheduling is a valuable tool which influences performance of cloud service providers to a great extent. Conventional approach that is used in optimizations is deterministic, fast, and gives perfect answers but frequently bog down in local optimum. In this paper, Cuckoo Search (CS) is proposed for the optimization. Experiments conducted and the results revealed that the proposed method outperformed other methods.

Published on: 10th– August-2016

KEY WORDS

Cloud Computing, Scheduling, Cuckoo Search (CS).

***Corresponding author:** **Email:** sivakumarm.clo@gmail.com

INTRODUCTION

Cloud computing offers resources to users as per their demands. The users demand for available services as per their desired Quality of Service, and pay for them on a pay per use basis. A challenging problem in Cloud computing is workflow scheduling. The processing units in cloud environment are known as virtual machines. It is to make sure that tasks are not loaded heavily on a single VM and other VMs do not remain idle and/or under loaded [1].

There is great need for the cloud services and their scheduling. Scheduling will be followed by the task or job scheduling within the resources. There might be more instances of a single resource which can be run at the same instant. It is necessary to check availability and reliability while also load ought to be equalized amongst the resources of the same kind. For the above variables there needs to be a function or procedure which could check them and allotment should be done in an optimum way.

The best way is combine together the computability of network strategies with scheduling algorithms [2]. Typically when tasks are scheduled they are done as per user's requirements as well as requests but when looking into all the features the computation requires to be done. Application scalability is the primary aim for cloud services to attain. In cloud scalability of resources permits real time provisioning of resources. Cloud has complicated execution environment however it has to offer QoS to its users. Virtual resources are utilized effectively for the entirely customizable configuration environment for application.

Workflows enable arranging apps in a directed acyclic graph form, wherein every node denotes the constituent job and edges denote inter job dependencies of apps. One workflow comprises a set of jobs that may interact with other jobs in the workflow. Hence, workflow scheduling is important in managing workflow implementation.

For taking cloud computing, scientific workflow gains more utilizations. But, we face several new challenges, wherein data as well as task scheduling are one. How to effectively schedule all tasks in an application is the most crucial problem. In order to reduce the executing time, schedule the computing intensive tasks to the high performance computer [3]. As the task scheduling is an NP-complete problem, certain heuristic algorithms have been utilized to resolve it.

THE IIOAB JOURNAL ISSN: 0976-3104

For solving the NP complete as well as NP hard problems heuristic methods can be used. The several algorithms that may be utilized for the scheduling are evolutionary protocols which are based on the biological evolution of species. The evolutionary protocols are Particle Swarm Optimization as well as Genetic algorithm [4].

Optimization means finding the best solution for a given problem. The field of optimization algorithms studies algorithms derived from the observations and these algorithms are a source of inspiration for designing novel protocols for solving optimization as well as distributed control issues. Conventional methods need a lot of computational efforts which tend to fail when problem size rises [5].

Usage of bio-inspired algorithm is a motivational method for employing computationally effective alternative systems for deterministic method. Swarm intelligence denotes the group intelligence of social insects because of their efficacy in resolving complicated issues like discovering shortest route from nests to food sources or the organization of their nest. Despite the fact that the insects are not sophisticated as an individual unit, as a swarm through interactions with one another and the environment make them collectively intelligent.

In recent years, this activity of swarms is abstracted as numerical optimization techniques [6]. This collective intelligence arises out of a process of self-organizing of units that evolve automatically as per a set of rules denoting movement patterns as well as interactions with environment and other agents so that intelligent activity rises out of simplistic individual activity.

One of such nature inspired algorithms is Cuckoo Search (CS) algorithm. It owes its inspiration to the obligate brood parasitism of certain cuckoo species which lay their eggs in the nests of host birds [7]. CS is a recent meta-heuristic algorithm and is utilized in solving complex optimization problems. The optimum solutions acquired by CS are far better than those obtained by other swarm intelligence algorithms.

In the current paper we use CS algorithm to optimize the scheduling process. Remaining sections are formed as: Section 2 discusses the related work in literature. Section 3 describes the method used in the proposed work. Section 4 reveals the experiment results and Section 5 concludes the proposed work.

RELATED WORKS

Nandhakumar and Ranjithprabhu [8] compared and analyzed the performance of different heuristic workflow scheduling algorithms with various QoS parameters and scheduling factors.

Saxena and Saxena [9] proposed a new scheduling algorithm for workflow, which are having dependencies among tasks, taking into consideration important parameters of transfer time and bandwidth along with basic requirements of optimizing the execution time and cost. The simulation was experimented using cloudSim toolkit. The propsoed algorithm provided better results over other existing algorithm like PSO and CSO and is more closely related to real world scenario.

Wang et al., [10] proposed a competitive, dynamic and multiple DAG scheduling algorithm which takes link communication processor into consideration (CCRH). The algorithm used communication competition model to describe the communication between the processors. Simulation results showed that under the premise of ensuring reliable scheduling, the algorithm not only can improve the fairness of the multiple DAGs scheduling, but also effectively short the average multiple DAGs scheduling time, and make the robustness of algorithm is better.

Verma and Kaushal [11] proposed Bi-Criteria Priority based Particle Swarm Optimization (BPSO) for scheduling workflow tasks across the present cloud resources which minimized the execution costs as well as the execution time under the specified deadline as well as budget restrictions. The suggested algorithm was tested using simulations with four separate real world workflow applications and they are compared with Budget Constrained Heterogeneous Earliest Finish Time (BHEFT) as well as standard PSO. The simulation outcomes showed that the proposed scheduling algorithm considerably decreased the execution cost of schedule when compared with BHEFT as well as PSO with Deadline, Budget Constraints as well as using same pricing model.

Thanh et al., [12] proposed a metaheuristic algorithm called PSOi which based on the Particle Swarm Optimization method. The experiments that were arranged by utilizing simulation tool CloudSim reveal that PSOi

THE IIOAB JOURNAL
ISSN: 0976-3104

is better than the generic Random as well as RoundRobin, furthermore the deviation between the solutions found by PSOi as well as the optimal solution is negligible.

Tang et al., [13] proposed a DVFS-enabled efficient energy workflow task scheduling algorithm: DEWTS. Through merging the relatively inefficient processors by reclaiming the slack time, DEWTS may exploit the useful slack time once more employing DVFS method after severs combined. On the basis of the amount of arbitrarily generated DAGs workflows, experimental outcomes showed that DEWTS is capable of reducing the total power utilization with various parallel applications and balancing the scheduling performance.

Bittencourt et al., [14] analyzed the problems in the scheduling the workflows in the hybrid clouds and surveyed few scheduling algorithms used for the cloud systems. To impact of the communication channels on the allocation of the jobs was compared and evaluated for various scheduling algorithms.

Lu and Gu [15] proposed a load-adaptive cloud resource scheduling model based on ant colony algorithm. By analyzed an example result, the model could meet the goals and requirements of self-adaptive cloud resources scheduling and improved the efficiency of the resource utilization.

A scheduling algorithm based on Genetic Algorithm (GA) was proposed by Wei and Tian [16]. Scheduling scheme is encoded in integer sequence, as well as a fitness function on the basis of influence degree is formulated. The genetic operations are selection, crossover, mutations as well as elitist selection. An optimal method was proposed according to the practical application. Finally resources scheduling problem in a cloud design platform proved the validity of the scheduling algorithm and the effectiveness of the optimization method.

Ge and Wei [17] presented a scheduling algorithm to make a scheduling decision by comparing and judging about the entire group of tasks in the job queue. For the optimization of the parameters of the scheduler, a GA was designed. Simulations were conducted and the results were proved that the proposed scheduling policy balanced the load among the nodes in better than the First in First Out (FIFO) and delay scheduling.

Raghavan et al., [18] utilized a metaheuristic method known as bat algorithm. It is specifically formulated for optimizing hard problems. Bat algorithm with the help of binary bat algorithm was utilized for scheduling workflow in a cloud. Particularly, the mapping of tasks and resources is performed through this method. The optimum resources were selected such that the overall cost of the workflow is minimal.

METHODOLOGY

Cuckoo Optimization Algorithm (COA) is improved to select the resources and schedule the tasks to minimize the overall completion time of tasks optimally by combining random local search and basic CS. Mapping of the real world problem to the meta-heuristic algorithm is very important and explained in detail. For jobs to be executed in cloud resources it compute the time taken by to compute and so on. A time to execute matrix is created as shown in **[Table -1]**.

Table: 1. Mapping of resource to jobs

	CR1	CR2	CR3
j1	3.54	0.27	2.1
j2	5.28	1.12	9.07
j3	0.81	0.83	3.05
j4	8.12	5.33	5.22
j5	1.06	9.19	2.76
j6	0.29	0.18	1.79
j7	2.27	7.93	9.98
j8	4.92	8.9	3.14
j9	3.47	2.3	2.18
j10	4.39	1.58	8.04

THE IIOAB JOURNAL
ISSN: 0976-3104

j11	0.3	0.77	8.85
j12	8.35	4.71	3.03
j13	5.98	8.69	3.57
j14	8.35	4.77	8.35
j15	4.77	9.55	1.5

Generating initial cuckoo habitat: To resolve optimizing issues, it is required for the values of the issue variables be grouped as arrays. In GA as well as PSO terms, the array is known as 'Chromosome' and 'Particle Position' respectively. But here, in COA is known as 'habitat'. To begin the optimization algorithm, candidate habitat matrices are created. Few, arbitrarily produced quantity of eggs are assumed for every original cuckoo habitat. In the real world, cuckoos lay around five to twenty eggs [19]. These numbers are utilized as upper and lower limits of eff designated to every cuckoo at various iterations. Another habit of cuckoos is that they lay an egg at the farthest distance from their own habitat. This is known as 'egg laying radius' (ELR). Every cuckoo possesses an ELR that is suitable for the overall quantity of eggs, number of cuckoo eggs and also differing limits of var_{hi} and var_{low}. So ELR is given by:

$$ELR = \alpha \times \frac{Number\ of\ current\ cuckoo's\ eggs}{Total\ number\ of\ eggs} \times (Var_{hi} - Var_{low})$$

Which α is an integer, supposed to handle the maximum value of ELR.

Immigration of cuckoos: Once young cuckoos grow old, they fly to live in their own region and when the season for laying eggs rolls around, they shift to fresh habitats with most similar host eggs and more food for the young birds. Then the cuckoo groups are created in several regions, the society with greatest fitness value is chosen as target point and all cuckoos move toward it. When mature cuckoos that live in those environments identify cuckoos belonging to other groups is a tough task. Most benefit is defined by target group and subsequently the group's most optimal habitat is the new destination habitat for moving cuckoos. When moving to the target point, cuckoos do not fly directly straight to destination. They cross partial distance and deviate.
Pseudo code of Cuckoo Optimization Algorithm

1. *Initialize cuckoo habitats with random points*
2. *Define ELR for each cuckoo*
3. *Let cuckoo to lay eggs inside their corresponding ELR*
4. *Kill those eggs that are identified by host birds*
5. *Eggs hatch and chicks grow*
6. *Evaluate the habitat of each newly grown cuckoo*
7. *Limit cuckoos maximum number in environment and kill those that live in worst habitats*
8. *Cuckoos find best group and select goal habitat*
9. *Let new cuckoo population move toward goal habitat*
10. *If stop condition is satisfied end, if not go to* 2

CUCKOO SEARCH (CS)

CS optimization algorithm is one of evolutionary algorithms and it was introduced by Yang and Deb in the year in 2009 [20]. The lifestyle and behavior of a bird called the Cuckoo was inspired by the developers of this algorithm. The brooding nature of this bird is different from the other birds. Cuckoo bird does not use its nest for laying the eggs and use other bird's nest for laying eggs. If the host bird finds that the eggs are not belongs to other bird, it will throw away or leave the nest. The grown cuckoo bird becomes a mature bird, and then it continues the mother's life instinctively [21].

THE IIOAB JOURNAL
ISSN: 0976-3104

Cuckoo Behavior: Certain cuckoo species have evolved such that female parasitic cuckoos are typically specialized in mimicry in colour as well as pattern of the eggs of a few particular host species. This decreases the probability of eggs being discarded and increases their reproductively. Cuckoos often select nests wherein host birds have just laid their own eggs. Generally, cuckoo eggs hatch a little earlier than the host eggs. When first cuckoo chick hatches, the instinctive action will evicting the host eggs by blindly pushing the eggs out of the nest, thereby increasing the cuckoo chick's share of food given by the host bird. Cuckoo characteristics could be described, as a model for good behavior other animals have extensive use in computing Intelligence Systems.

Levy Flights: The activity of animals to scour for food is quasi-random in practice. In recent research, it has been proven that flight activity of several creatures demonstrate generic features of Levy flights. Typically, foraging routes of creatures is technically an arbitrary walk as the next move is on the basis of current position and transition probability to the next locale [22]. Selecting the direction relies on a certain probability that may be abstracted mathematically. Several researches have employed these activities in optimizations, optimal searches and initial outcomes reveal its promise.

Levy flight is the most popular technique used and handles [23]:

- The generation of how each step should be
- The random direction of flight which is given by equation:

$$L = \frac{u}{|v|^{1/\beta}}$$

Where β is the scaling value with a range of [1, 2]. $u$and v are generated from normal distribution and shown in equation:

$$u \sim N(0,\sigma_u^2),\ \mathrm{v} \sim N(0,\sigma_v^2)$$

Where σu and σv are calculated using equation:

$$\sigma_u = \left\{ \frac{\Gamma(1+\beta)\sin(\pi\beta/2)}{\Gamma[(1+\beta)/2]\beta 2^{(\beta-1)/2}} \right\}^{1/\beta}, \ \sigma_v = 1$$

Where Γ is the gamma function.

Cuckoo Rules and Parameters: To simplify the principles of the new CS algorithm, three exemplary rules can be used.

- Every cuckoo lays a single egg at a time, and deposits it in an arbitrarily chosen nest,
- The best nests with excellent quality of eggs (solutions) will be carried over to the next generation.
- The quantity of available host nests is static, and hosts can discover alien eggs with a probability of pa ∈ [0, 1]. In such a case, host birds can either discard the egg or abandon nest and build an entirely new nest in a fresh location.

In this overall, minimizing the overall task completion time is taken as the fitness function. The parameters updation and usage are given by:

Step 1: Assign the nests randomly. This indicates selection of random solution

Step 2: Select one random nest and replace it by a best solution. Best solution is found by a levy flight operation. When generating new solutions, $x^{i\,(t+1)}$for the *i* th Cuckoo, a Lévy flight is performed using the equation:

$$x_i^{(t+1)} = x_i^{(t)} + \alpha.S$$

Here *α* > 0, the parameter *α* is used as the step size parameter. It must be selected based on the problem scale. Most of the times it is set to unity in the CS [24] and reduced in the improved CS algorithm. The solutions or nests in the current positions are then used for finding best solution so far as the origin of the Lévy flight. The step size also increases the efficiency and performance of the CS algorithm. The parameter value *S* represents the length of random walk with the Lévy flights.

Step 3: In the next step, the fraction value *pa* is used to discover the worst nest, so that they can be and replaced by the best ones. This parameter *pa* is considered as the probability of a solution's that will be discovered. Therefore, a probability matrix is created by the equation:

$$P_{ij} = \begin{cases} 1 & if\ \ rand < pa \\ 0 & if\ \ rand \geq pa \end{cases}$$

Where *pa* represents the discovering probability and maximum number of analyses is the stopping criterion. *rand*is an arbitrary number within the range [0, 1] and *Pij*is the probability of discovering *j*th variable of *i*th nest. It assigns either 0 or 1 to each variable on the nest. Using random walk, the point wise multiplication of random step sizes with probability matrix from their current positions according to quality.

Pseudo code for The CS algorithm

THE IIOAB JOURNAL ISSN: 0976-3104

$Input:$

$function\ to\ optimize,\ fit$

$Population\ of\ \ n\ host\ nests\ x_i = i(1,2,......,n)$

$Output:$

$best\ solutions\ (nests\ with\ quality\ solutions),$

$Initialize:$

$While\ (t < MaxGeneration)$

$Get\ a\ cuckoo\ randomly\ by\ Levy\ flights,$

$Evaluate:$

fit

$Randomly\ choose\ nest\ among\ n\ available\ nests$

$If\,(fit_i < fit_j)$

$\mathrm{Re}\ place\ j\ by\ the\ new\ solutions;$

$End\ \ if$

$bandona\ fraction\ (pa)\ of\ \ worse\ nests\ and\ build$

$\mathrm{Re}\ peat$

$new\ nest$

$New\ locations\ via\ Levy\ flights;$

$Keep\ the\ best\ solutions;$

$Rank\ the\ solutions\ and\ find\ the\ current\ best;$

$end\ \ while$

$Post\ process\ results\ and\ visualization;$

END

Fitness Value: Fitness or quality value reveals how fit a solution is, i.e. how well it can adapt to its environment. For maximization problems, the fitness of solutions are proportional to values of objective functions. For simplicity, suppose every egg in a nest denotes a solution, and cuckoo egg denotes a new solution. The aim is to utilize the new and potentially improved solutions (cuckoos) to substitute a not-so-good solution in the nests. Here, it use the simplest approach where each nest has only a single egg [22].

EXPERIMENTAL SETUP

Thirty tasks are assigned to Cloud with 5 resources and 10 resources. CloudSim simulator is used for conducting the experiments. The resources are located at two data centers. Each resource has 1 CPU with 512 Mb RAM. Each task is of size 1, 2, 3 or 4 units. The simulations were conducted using random local search and CS, the overall task completion time or makespan is used for comparing the performance.

RESULTS

The overall task completion time is shown in **[Table -2]**.

Table: 2. Overall completion Time

Technique used	10 resource	5 resource
Random Local Search	7.96 second	15.36 second

COMPUTER SCIENCE

THE IIOAB JOURNAL ISSN: 0976-3104

Cuckoo Search	7.42 second	15.18 second

From the numerical results, it is noted that the overall task completion time of the proposed CS optimization drastically reduces the overall completion time.

CONCLUSION

Task scheduling problem concerns about the dynamic distribution of the tasks over the Cloud resources to achieve the best results. In the current paper, a task scheduling algorithm has been suggested to the independent task over Cloud Computing. The suggested algorithm is the CS algorithm. CS algorithm is based on the obligate brood parasitic behavior of some cuckoo species in combination with the Levy flight behavior of some birds and fruit flies. In the proposed CS algorithm, all the nests are ranked then a random local search is initiated with the average value of the top three nests (solutions). The best solution obtained by CS and the best solution obtained by Random Local Search are sent to the next iteration. For the simulation 30 tasks are taken with number of resources as 5 and 10.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

ACKNOWLEDGEMENT

None

FINANCIAL DISCLOSURE

The authors report no financial interests or potential conflicts of interest.

REFERENCES

[1] Lovesum SJ, Krishnamoorthy K, Prince P. [2014] An optimized qos based cost effective resource scheduling in cloud. *Journal of Theoretical and Applied Information Technology*, 66(1).

[2] Manimaran V, Prabhu S. [2010] A Survey on Resource Scheduling and Allocation Policy in a Cloud Environment.

[3] Guo L, Zhao S, Shen S, Jiang C. [2012] Task scheduling optimization in cloud computing based on heuristic algorithm. *Journal of Networks*, 7(3): 547-553.

[4] Jacob L, Jeyakrishanan V, Sengottuvelan P. [2014] Resource Scheduling in Cloud using Bacterial Foraging Optimization Algorithm. International *Journal of Computer Applications*, 92(1).

[5] Singh G, Kaur A. [2015] Bio Inspired Algorithms: An Efficient Approach for Resource Scheduling in Cloud Computing. *International Journal of Computer Applications*, 116(10).

[6] Milani FS, Navin AH. [2015] Multi-Objective Task Scheduling in the Cloud Computing based on the Patrice Swarm Optimization.

[7] Soneji H, Sanghvi, RC.[2012, October] Towards the improvement of cuckoo search algorithm. In Information and Communication Technologies (WICT), *IEEE* 878-883.

[8] Nandhakumar C, Ranjithprabhu K. [2015] Heuristic and meta-heuristic workflow scheduling algorithms in multi-cloud environments—A survey. In Advanced Computing and Communication Systems, 2015 International Conference on (pp. 1-5). *IEEE*.

[9] Saxena S, Saxena D.[2015] EWSA: An enriched workflow scheduling algorithm in cloud computing. In Computing, Communication and Security (ICCCS), 2015 International Conference on (pp. 1-5). *IEEE*.

[10] Wang Y, Jia C, Xu Y. (2014] Multiple DAGs dynamic workflow scheduling based on the primary backup algorithm in cloud computing system. In Broadband and Wireless Computing, Communication and Applications (BWCCA), 2014 Ninth International Conference on (pp. 177-182). *IEEE*.

[11] Verma A, Kaushal S. [2014] Bi-criteria priority based particle swarm optimization workflow scheduling algorithm for cloud. In Engineering and Computational Sciences (RAECS), 2014 Recent Advances in (pp. 1-6). *IEEE*.

[12] Thanh TP, LN, Doan CN.[2015] A novel workflow scheduling algorithm in cloud environment. In Information and Computer Science (NICS), 2015 2nd National Foundation for Science and Technology Development Conference on (pp. 125-129). *IEEE*.

[13] Tang Z, Cheng Z, Li K, Li K. [2014]An Efficient Energy Scheduling Algorithm for Workflow Tasks in Hybrids and DVFS-enabled Cloud Environment. In Parallel Architectures, Algorithms and Programming (PAAP), 2014 Sixth International Symposium on (pp. 255-261). *IEEE*.

[14] Bittencourt LF, Madeira ER, Da Fonseca, NL. [2012] Scheduling in hybrid clouds. Communications Magazine, *IEEE*, 50(9):42-47.

THE IIOAB JOURNAL ISSN: 0976-3104

[15] Lu X., Gu Z. [2011] A load-adapative cloud resource scheduling model based on ant colony algorithm. In Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference on (pp. 296-300). IEEE.
[16] Wei Y, Tian L. [2012] Research on cloud design resources scheduling based on genetic algorithm. In Systems and Informatics (ICSAI), 2012 *International Conference on* (pp. 2651-2656). *IEEE*.
[17] Ge Y, Wei G. [2010] GA-based task scheduler for the cloud computing systems. In Web Information Systems and Mining (WISM), 2010 International Conference on , 2: 181-186). *IEEE*.
[18] Raghavan S, Marimuthu C, Sarwesh P, Chandrasekaran K. [2015] Bat algorithm for scheduling workflow applications in cloud. In Electronic Design, Computer Networks and Automated Verification (EDCAV), 2015 International Conference on (pp. 139-144). *IEEE*.
[19] Rabiee M, Sajedi H. [2013] Job scheduling in grid computing with cuckoo optimization algorithm. *International Journal of Computer Applications*, 62(16): 38-44.
[20] Yang XS, Deb S. [2009] Engineering optimization by cuckoo search. International *Journal of Mathematical Modelling and Numerical Optimization*, 1: 330–343.
[21] Al-maamari A, Omara FA. [2015] Task Scheduling using Hybrid Algorithm in Cloud Computing Environments. *Journal of Computer Engineering (IOSR-JCE*), 17: 96-106
[22] Navimipour NJ, Milani FS. [2015] Task scheduling in the cloud computing based on the cuckoo search algorithm. *International Journal of Modeling and Optimization*, 5(1): 44.
[23] Baskan, O. [2013] Determining optimal link capacity expansions in road networks using Cuckoo Search algorithm with Lévy Flights. Journal of Applied Mathematics, 2013.
[24] Yang XS, Deb S. [2009] Cuckoo search via Lévy flights. In Nature and Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on (pp. 210-214). *IEEE*.