Nature Inspired Algorithms in Cloud Computing: A Survey

Gamal Abd El-Nasser A. Said

Department of Information Technology, Port Training Institute, Arab Academy for Science, Technology and Maritime Transport

Alexandria, Egypt

Email address:

jamalabdelnasser@hotmail.com

To cite this article:

Gamal Abd El-Nasser A. Said. Nature Inspired Algorithms in Cloud Computing: A Survey. *International Journal of Intelligent Information Systems*. Vol. 5, No. 5, 2016, pp. 60-64. doi: 10.11648/j.ijiis.20160505.11

Abstract: Cloud Computing consists of many resources, the problem of mapping tasks on unlimited computing resources in cloud computing is NP-hard optimization problem. In this paper, we provide a survey of popular nature inspired algorithms: Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), and Genetic Algorithm (GA) for solving NP-hard problems in cloud computing.

Keywords: Ant Colony Optimization (ACO), Cloud Computing, Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Metaheuristics

1. Introduction

Cloud computing is an emerging technology and it allows users to pay as you need and has the high performance. As it becoming popular, it is also becoming more challenging. In cloud computing resource allocation is one of the challenges. The goal of resource allocation is to allocate tasks to appropriate resources, resource allocation in cloud computing is NP-hard optimization problem. There are no algorithms which produce optimal solution within polynomial time to solve NP-hard problems.

Scheduling is the process of mapping the jobs, to the resources available. A job is usually split into multiple tasks. Hence scheduling can be redefined as mapping of tasks to a selected group of resources. Scheduling plays an important role in a cloud. Resource allocations is done with the help of scheduling which is used for supplying resources effectively based on request from users and availability of resources. Cloud computing is a paradigm that refers to buying computer resources which could be services, software, or infrastructure based on usage levels. Cloud service providers (CSP) are equipped with large data centers. Users request CSP for resources, and the CSP provides them from their pool of resources. This is profitable for both users and CSPs; users save the cost of design, setup and maintenance of data centers, while CSPs give access to resources to large number of users, and thereby make profit [1].

One of the challenges posed by cloud applications is Quality-of-Service management, which is the problem of allocating resources to the application to guarantee a service level along dimensions such as performance, availability and reliability. The main advantage of job scheduling algorithm is to achieve a high performance computing and the best system throughput. [2]. In Cloud Computing, Scheduling assumes a fundamental part in effectively dealing with the computer administrations; it is the movement of captivating choices in regards to the distribution of accessible limit and/or resources to jobs and/or clients on time. Million of clients offer cloud administrations by presenting their large number of processing tasks to the cloud computing environment. Scheduling of these huge numbers of tasks is a conflict to the cloud environment. The scheduling crisis in cloud makes it hard to work out, dominatingly on account of substantial composite jobs like workflows, so to solve these types of large problems many algorithms are proposed [3].

The task scheduling strategy in cloud environment plays an important role in making use of the cloud resources effectively. The total completion time, the average completion time and the load balancing are the most important factors to determine the performance of task scheduling in cloud computing. In this paper, we provide a survey of various nature-inspired algorithms in cloud computing based on three popular metaheuristics algorithms: Ant Colony Optimization (ACO), Particle Swarm

Optimization (PSO), and Genetic Algorithm (GA).

The rest of this paper is organized as follows: A brief overview about cloud computing is given in section 2. Section 3 provides overview of related work of different nature inspired algorithms for solving NP-hard problems in cloud computing. A brief overview of nature inspired algorithms is given in section 4. Conclusions and future work are given in section 5.

2. Cloud Computing: An Overview

Cloud computing, also known as 'on-demand computing', is a kind of Internet-based computing, where shared resources, data and information are provided to computers and other devices on-demand. Cloud computing is a kind of distributed computing with heterogeneous computing resources. It needs to use virtualization technology to change the heterogeneous resources into resources with the same or similar function to complete tasks with different types and large quantity [4]. Cloud computing is associate on demand service during which shared resources, data, package and alternative devices are provided per the clients demand at precise time. Cloud task scheduling is an NP-hard optimization problem, and many metaheuristic algorithms have been proposed to solve it [5].

Depending on the type of services offered, cloud services can be classified into three major categories: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) [6]. In the SaaS model, an application is hosted by a cloud vendor and delivered as a service to users, primarily via the Internet or a dedicated network. It eliminates the need to install and run the application locally, on a user's computer, and thereby also relieves the users from the burden of hardware and software maintenance and upgrades. Users are billed for the service(s) used, depending on their usage. Examples of SaaS include Webmail, and Google Apps.

In the PaaS model, the platform and tools for application development and middleware systems are hosted by a vendor and offered to application developers, allowing them simply to code and deploy without directly interacting with the underlying infrastructure. The platform provides most of the tools and facilities required for building and delivering applications and services such as Web service integration, database integration, security, and storage, Examples of PaaS include Google App Engine, and Microsoft Azure.

In an IaaS cloud, raw computer infrastructure, such as servers, CPU, storage, network equipment, and datacenter facilities, are delivered as a service on demand. Rather than purchasing these resources, clients get them as a fully outsourced service for the duration that they need them. Amazon Elastic Compute Cloud, GoGrid, and FlexiScale are some of the examples of IaaS clouds. An IaaS cloud exhibits the following characteristics: availability of a huge volume of computational resources such as servers, network equipment, memory, CPU, and disk space.

Based on where the cloud is deployed and by whom, who owns and manages it, and who its primary users are, clouds are classified into five categories: public cloud, private cloud, virtual private cloud, community cloud, and hybrid cloud [6]: The public cloud is the most common and widely known form of cloud, and is open for anyone – business, industry, government, and individuals. The cloud infrastructure is, however, owned and managed by the cloud service provider. Public cloud services are offered on a pay-per-usage model; A private cloud is deployed, provided, and controlled by an enterprise behind its firewall for its own use. Some enterprises deploy their own cloud computing environments for their own exclusive use. Thus, by having their own cloud, they gain operational efficiencies, effectively use their existing resources, if any, and have full control over the cloud, the applications, and data on the cloud; A virtual private cloud (VPC) is a segment of a public cloud, designated for a user with additional provisions and features for meeting that user's specific security and compliance requirements. An example of this type of cloud is Amazon's VPC; A community cloud is optimized and specially deployed for use by a particular industry sector or a group of users so that it meets specific requirements to address issues that are crucial to them; A hybrid cloud is a combination of two or more of the above cloud models. In this model, an enterprise makes use of both public and private clouds – deploying its less critical, low-risk services on a public cloud and business-critical core applications on its internal private cloud.

3. Related Work

Cloud computing enhances its performance and throughput by using an efficient task scheduling algorithm. Intensive explore has been carried on over the past decades in cloud to solve the problem of task scheduling in the cloud. The section throws light on the work of some renowned researchers who had been pillars and founders of the current research work.

In [7] paper proposed a cloud task scheduling policy based on Load Balancing Ant Colony Optimization (LBACO) algorithm. The main contribution of our work is to balance the entire system load while trying to minimizing the make span of a given tasks set. The new scheduling strategy was simulated using the CloudSim toolkit package. Experiments results showed the proposed LBACO algorithm outperformed FCFS (First Come First Serve). In [8] Authors presented a load-adaptive cloud resource scheduling model based on ant colony algorithm. By real-time monitoring virtual machine of performance parameters, once judging overload, it schedules fast cloud resources using ant colony algorithm to bear some load on the load-free node. So that it can meet changing load requirements. Analysis of results showed that the model improved the efficiency of the resource utilization.

In [9] authors focused on the Infrastructure as a Service model where custom Virtual Machines (VM) are launched in appropriate hosts available in a Cloud. This work described

and evaluated a Cloud scheduler based on Ant Colony Optimization (ACO). The main performance metrics studied are the number of serviced users by the Cloud and the total number of created VMs in online scheduling scenarios. Simulated experiments showed that the proposed scheduler succeeds in balancing the studied metrics compared to other schedulers. A particle swarm optimization (PSO) based heuristic to schedule applications to cloud resources that takes into account both computation cost and data transmission cost is presented in [10]. Experiments were carried out by varying computation and communication costs, results showed that PSO achieved good distribution of workload onto resources.

A new task scheduling model is proposed based on Particle Swarm Optimization algorithm in [11], the global search performance and convergence rate of the proposed algorithm were validated by the results. The model optimized the task execution time in view of both the task running time and the system resource utilization. In [12] Authors focused on two objectives, makespan and cost, to be optimized simultaneously using metaheuristic search techniques for scheduling independent tasks. A new variant of continuous Particle Swarm Optimization (PSO) algorithm, named Integer-PSO, is proposed to solve the bi-objective task scheduling problem in cloud which out performs the smallest position value (SPV) rule based PSO technique.

In [13] Authors proposed a Task-based System Load Balancing method using Particle Swarm Optimization (TBSLB-PSO) that achieves system load balancing by only transferring extra tasks from an overloaded VM instead of migrating the entire overloaded VM. The simulation results showed that the proposed TBSLB-PSO method significantly reduces the time taken for the load balancing process compared to traditional load balancing approaches. Authors in [14] proposed a task scheduling model based on the genetic algorithm. In the proposed model, the task scheduler calls the GA scheduling function every task scheduling cycle. This function creates a set of task schedules and evaluates the quality of each task schedule with user satisfaction and virtual machine availability. Experimental results showed effectiveness and efficiency of the genetic algorithm-based task scheduling model in comparison with other existing task scheduling models.

In [15] Authors established a scheduling model for cloud computing based on MO-GA algorithm to minimize energy consumption and maximize the profit of service provides under the constraint of deadlines. In this paper, authors proposed a job scheduling architecture under the environment of cloud computing, which contains several components to analyze the application, and allocate the suitable resources to the applications to improve the effectiveness and efficiency of the computing; then, the MO-GA based scheduling algorithm is proposed, several experiments are conducted to validate our scheduling models. An approach namely Effective Cloud Resource Allocation Using Improvised Genetic Approach is presented in [16] which directs to accomplish better virtual machine allocation

across cloud servers for maintaining vertical elasticity and minimizing response time. The proposed approach is focused on elasticity and Scheduling to improve resource allocation mechanism in cloud computing. Authors proposed innovative algorithm called Enhanced Genetic Algorithm (EGA) using Multipurpose Mutation Operator. The results showed that the EGA provides an optimal solution and proves better performance compared to the existing algorithms.

In [17] Authors focused is on a scheduling algorithm for hybrid cloud that tries to optimize both execution time and cost. Execution time and cost are conflicting objectives, i.e. when one is made better, the other becomes worse off. Multiobjective evolutionary algorithm is used to find the optimal schedule. The widely used scheduling implementations seen in hybrid cloud try to optimize either execution time or cost, but not both simultaneously. The proposed algorithm is compared with the more widely used scheduling optimization techniques and seen to have much better performance.

Authors in [18] proposed a comparison between metaheuristic algorithms, Genetic Algorithm, Tabu Search, and Simulated Annealing for solving Quadratic Assignment Problem. The computational results show that genetic algorithm has a better solution quality than the other metaheuristic algorithms for solving quadratic assignment problems.

4. Nature Inspired Algorithms

Nature inspired algorithm is a technique that is inspired by processes, observed from nature. Nature inspired optimization algorithms are mainly categorized into evolutionary algorithms and swarm intelligence based algorithms. Evolutionary algorithms are based on the evolutionary behavior of natural systems e.g. genetic algorithm (GA). Swarm intelligence based algorithms optimize the certain problem by mimicking the collective behavior of natural swarm's.e.g. particle swarm optimization (PSO), and ant colony algorithm (ACO).

4.1. Ant Colony Optimization

In 1992, Marco Dorigo proposed a new algorithm in his PhD thesis, Ant colony optimization. Ants are small creatures that can intelligently find the shortest path between their nest and food source. Each ant represents a potential solution to an objective function. Ants establish their communication through trails of pheromone. When an ant comes out of their nest to search for food source, they move randomly in any direction, leaving pheromone in their path. On reaching the food source, ants return back with food, leaving pheromone again on same path. Hence, the path with the highest amount of pheromone represents the shortest route from ant's nest to the food source. Pheromone concentration in each path represents the quality of solution (goodness of fitness value). The process is continued until stopping criteria is met. An ant K (k=1,2, …, m) represents a solution string. Pseudo code of ACO algorithm can be represented as follows:

```
Begin
    Initialize the pheromone trails and parameters;
        Generate population of m solutions (ants);
        For each individual ant k∈m: calculate fitness (k);
        For each ant determine its best position;
        Determine the best global ant;
        Update the pheromone trail;
    Check if termination=true;
End
```

4.2. Particle Swarm Optimization

PSO was developed by J. Kennedy and R. Ederhart in 1995. PSO mimics the flocking behavior of birds. The birds fly in a solution space and their flocking behavior determines the optimum solution. Particles tend to move towards its local best position (solution) (*pBest*) found by them so far. They also keep the track of global best (*gBest*) solution, the best (shortest) path found by any particle at particular instance. Each particle is allied with position and velocity and moves through a multi-dimensional search space. In each iteration, each particle adjusts its velocity based on its best position and the position of the best particle of the whole population. the position in '*n*' dimension space and the current position of particle with respect to *gBest* and *pBest*. Birds communicate with each other to find the most optimum (best) path to reach its food sources. Hence, they learn from the experience of their local best solutions and global best solutions. The algorithm continues till global optimum solution is achieved. Pseudo code of PSO algorithm can be represented as follows.

```
Begin
 For Each Particle
 {
 Initialize particle
 }
 Do until maximum iterations or minimum error criteria
 {
        For Each Particle
        {
        Calculate data fitness value
        If the fitness value is batter than pBest
 {
                Set pBest = current fitness value
        }
        If pBest is batter than gBest
        {
        Set gBest = pBest
        }
        }
        For Each Particle
        {
        Calculate Particle Velocity
        Use gBest and Velocity to update particle Data
        }
 }
 End
```

4.3. Genetic Algorithm

GA described by John Holland in the 1960s and developed by Holland and colleagues at the University of Michigan in the 1960s and 1970s. Genetic Algorithm is a nature inspired algorithm that aims to find solutions to NP-hard problems. The basic idea of Genetic Algorithms is to first generate an initial population randomly which consist of individual solution to the problem called Chromosomes, and then evolve this population after a number of iterations called Generations. During each generation, each chromosome is evaluated, using some measure of fitness. To create the next generation, new chromosomes, called offspring, are formed by either merging two chromosomes from current generation using a crossover operator or modifying a chromosome using a mutation operator. A new generation is formed by selection, according to the fitness values, some of the parents and offspring, and rejecting others so as to keep the population size constant. Fitter chromosomes have higher probabilities of being selected. After several generations, the algorithms converge to the best chromosome, which hopefully represents the optimum or suboptimal solution to the problem. Pseudo code of GA algorithm can be represented as follows.

```
Begin
    Generate randomly the initial population
    While termination condition not met do
        Evaluate the different chromosomes of the current
        population using the fitness functions
        Apply Genetic Operators
            Selection
            Crossover
            Mutation
    End While
End
```

5. Conclusions and Future Work

One of the important research issues which need to be focused in cloud computing for its efficient performance is scheduling. Scheduling in cloud computing belongs to a category of problems known as NP-hard problem due to large solution space and thus it takes a long time to find an optimal solution. This paper reviews the application of popular nature inspired algorithms, Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), and Genetic Algorithm (GA) for solving scheduling problems, reduction of makespan, execution cost, and resource utilization in cloud computing.

In future work, further investigations will be done for the application of nature inspired algorithms in the area of security to protect information associated with tasks and users in cloud computing.

References

[1] Leena, A., Ajeena, S., and Rajasree, S. (2013). "Inter-cloud scheduling technique using power of two choices" in Proc. IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-4.

[2] Lakshmi, R., and Srinivasu, N. (2016). "A dynamic approach to task scheduling in cloud computing using genetic algorithm". Journal of Theoretical and Applied Information Technology, Vol. 85, No. 2, pp. 124-135.

[3] Geetinder, K., and Sarabjit, K. (2016). "Improved Hyper-Heuristic Scheduling with Load-Balancing and RASA for Cloud Computing Systems ". International Journal of Grid and Distributed Computing, Vol. 9, No. 1, pp. 13-24.

[4] Zhenzhen, X., Xiujuan, X., and Xiaowei, Z. (2015). "Task Scheduling Based on Multi-objective Genetic Algorithm in Cloud Computing". Journal of Information & Computational Science, Vol. 12, No. 4, pp. 1429-1438.

[5] Parveen, K., and, Mandeep, K. (2015). "Load balancing in cloud using aco and genetic algorithm". International Journal of Scientific Research Engineering & Technology (IJSRET), Vol. 4, No. 7, pp. 724-730.

[6] Murugesan, S., and, Bojanova, I. (2016). "Cloud Computing: An Overview". Encyclopedia of Cloud Computing, IEEE, pp. 3-14.

[7] Kun, L., Gaochao, X., Guangyu, Z., Yushuang, D., and Dan, W., (2011). " Cloud Task Scheduling Based on Load Balancing Ant Colony Optimization", IEEE Sixth Annual Chinagrid Conference, pp.3-9.

[8] Xin, L., and Zilong, G. (2011). "A load adaptive cloud resource scheduling model based on ant colony algorithm", IEEE International Conference on Cloud Computing and Intelligence Systems, IEEE, pp.296-300.

[9] Pacinia, E, Mateosb, C, Garino, C. (2015). "Balancing throughput and response time in online scientific Clouds via Ant Colony Optimization", Advances in Engineering Software, Vol. 84, No. 1, pp. 31-47.

[10] Pandey, S., Wu, L., Guru, S., and Buyya, R. (2010). "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments", 24th IEEE International Conference on Advanced Information Networking and Applications, pp. 400–407.

[11] Zhanghui, L., and Xiaoli, W. (2012). "A PSO-Based Algorithm for Load Balancing in Virtual Machines of Cloud Computing Environment", Advances in Swarm Intelligence, Lecture Notes in Computer Science series, Springer, Vol. 7331; pp. 142-147.

[12] Beegom, S., and Rajasree, S. (2014). "A particle swarm optimization based pareto optimal task scheduling in cloud computing", Advances in Swarm Intelligence, Lecture Notes in Computer Science series, Springer,Vol.8795; pp.79-86.

[13] Ramezani, F., Jie, L., and Hussain, K. (2014) "Task-Based System Load Balancing in Cloud Computing Using Particle Swarm Optimization", International Journal of Parallel Programming, Vol. 42, No. 5, pp. 739–754.

[14] Jang, S., Kim, T, Kim, K., and Lee, S. (2012). "The study of genetic algorithm-based task scheduling for cloud computing". International Journal of Control and Automation, Vol. 5, No. 4, pp. 157–162.

[15] Jing, L., Xing-Guo, L., Xing-Ming, Z., Fan, Z., and Bai-Nan, L. (2013). "Job Scheduling Model for Cloud Computing Based on Multi-Objective Genetic Algorithm", International Journal of Computer Science Issues, Vol.10, No. 1, pp. 134-139.

[16] Durairaj, M., and Kannan, P. (2015). "Improvised Genetic Approach for an Effective Resource Allocation in Cloud Infrastructure", International Journal of Computer Science and Information Technologies, Vol.6, No.4, pp. 4037-4046.

[17] Leena, A., Ajeena, S., and Rajasree, S. (2016). "Genetic Algorithm Based Bi-Objective Task Scheduling in Hybrid Cloud Platform", International Journal of Computer Theory and Engineering, Vol. 8, No. 1, pp. 7-13.

[18] Gamal Abd El-Nasser, S., Abeer, M., and El-Horbaty, S. (2014). "A Comparative Study of Metaheuristic Algorithms for Solving Quadratic Assignment Problem," International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 5, No. 1, pp. 1-6.

[19] Mala, K., and Sarbjeet, S. (2015). "A review of metaheuristic scheduling techniques in cloud computing", Egyptian Informatics Journal, Vol. 16, No. 1, pp. 275-295.

Biography

Gamal Abd El-Nasser A. Said: Dr.Gamal Abd El-Nasser received his Ph.D. (2016) in computer science from Faculty of Computer and Information Sciences, Ain Shams University. M.Sc. (2012) in computer science from College of Computing & Information Technology, (AASTMT), Egypt and B.Sc. (1990) from Faculty of Electronic Engineering, Menofia University, Egypt. His work experience as a Researcher, Maritime Researches & Consultancies Centre, Egypt. Computer Teacher, College of Technology Kingdom Of Saudi Arabia and Lecturer, Port Training Institute, (AASTMT), Egypt. His research areas include optimization, simulation, artificial intelligence, and cloud computing.