



Symbiotic Organism Search optimization based task scheduling in cloud computing environment



Mohammed Abdullahi^{a,c}, Md Asri Ngadi^{a,*}, Shafi'i Muhammad Abdulhamid^b

^a Department of Computer Science, Faculty of Computing, Universiti Teknologi Malaysia, Malaysia

^b Department of Cyber Security Science, Federal University of Technology Minna, Nigeria

^c Department of Mathematics, Ahmadu Bello University Zaria, Nigeria

HIGHLIGHTS

- Discrete Symbiotic Organism Search algorithm for task scheduling is proposed.
- The proposed algorithm has better ability to exploit best solution regions than PSO.
- The proposed method has global ability in terms of exploring optimal solution points.
- The proposed algorithm performs significantly better than PSO for large search spaces.

ARTICLE INFO

Article history:

Received 16 November 2014

Received in revised form

7 August 2015

Accepted 11 August 2015

Available online 24 August 2015

Keywords:

Cloud computing

Task scheduling

Makespan

Symbiotic Organism Search

Ecosystem

ABSTRACT

Efficient task scheduling is one of the major steps for effectively harnessing the potential of cloud computing. In cloud computing, a number of tasks may need to be scheduled on different virtual machines in order to minimize makespan and increase system utilization. Task scheduling problem is NP-complete, hence finding an exact solution is intractable especially for large task sizes. This paper presents a Discrete Symbiotic Organism Search (DSOS) algorithm for optimal scheduling of tasks on cloud resources. Symbiotic Organism Search (SOS) is a newly developed metaheuristic optimization technique for solving numerical optimization problems. SOS mimics the symbiotic relationships (mutualism, commensalism, and parasitism) exhibited by organisms in an ecosystem. Simulation results revealed that DSOS outperforms Particle Swarm Optimization (PSO) which is one of the most popular heuristic optimization techniques used for task scheduling problems. DSOS converges faster when the search gets larger which makes it suitable for large-scale scheduling problems. Analysis of the proposed method conducted using *t*-test showed that DSOS performance is significantly better than that of PSO particularly for large search space.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

In cloud computing, resources such as processors, memory, storage, and applications are provisioned as services, thereby changing the way IT resources are designed and acquired [1]. The cloud computing paradigm had greatly reduced the financial cost of acquiring hardware and software for application deployment as well as maintenance cost. Because of the high scalability, users are not bothered with imprecise forecasting of service scale which

will amount to resource wastage if over provisioned, and revenue loss if under provisioned [2,3]. Cloud computing resources are shared among cloud clients through the concept of virtualization. Virtualization allows many remote running environments to be safely combined on physical servers for optimum utilization of physical resources and energy [2]. Virtual Machine (VM) is a vital component of software stacks in the cloud data center. Cloud data are located across servers which are interconnected through networked resources and accessed via virtual machines. Amazon Elastic Computing Cloud (Amazon EC2) [4] is an example of cloud platform that provides infrastructure services in the form of VMs. One of the cardinal objectives of cloud computing is maximization of revenue both on the part of the cloud provider and the user. Task scheduling has evolved as one of the focus in cloud computing [5] since inefficient task scheduling can lead

* Corresponding author.

E-mail addresses: abdullahilwafu@abu.edu.ng (M. Abdullahi), dr.asri@utm.my (M.A. Ngadi), shafi.abdulhamid@futminna.edu.ng (S.M. Abdulhamid).

to revenue loss, performance degradation, and breach of Service Level Agreement (SLA). Therefore, efficient scheduling algorithms are required to minimize both computation-based metrics such as response time, system utilization, makespan, system throughput and network-based metrics such as network communication cost, traffic volume, round trip time, data communication cost [6]. These metrics are central to monitoring cloud activities in order to address issues like load balancing, energy efficiency, SLA and Quality of Service (QoS) guarantee, fault tolerance [6].

There are encouraging research results for efficient task scheduling in the cloud, but task scheduling problems are still NP-complete [7]. Most of the task scheduling algorithms used in cloud computing are rule based [8–10] because they are easy to implement. Rule based algorithms perform poorly when it comes to complex task scheduling problems [5]. The most common metaheuristic techniques applied to task scheduling problems in grid and cloud computing are Genetic Algorithm (GA) [11–15], Particle Swarm Optimization (PSO) [16–26], and Ant Colony Optimization (ACO) [27–30]. PSO converges faster and obtains better solution than GA and ACO due to its exploratory capability for finding optimal solutions [16,31]. Owing to the better performance of PSO over ACO and GA, variants and hybrid versions of PSO have been used for benchmarking the proposed algorithm.

In this paper, DSOS is proposed and used to schedule a bag of tasks in a cloud environment, hence network communication cost and data transfer cost are not main optimization issues when dealing with independent tasks. In the experimental setup, four data set categories, normal, left-skewed, right-skewed, and uniform distributions, were used to test the suitability of the proposed method. We used different forms of distributions to gain insight into the performance trend of the proposed method. Web applications such as web services are usually run for a long time and their CPU demand is variable. Moreover, High Performance Computing (HPC) applications have short life span and place a high demand on CPU. Furthermore, chosen statistical models for task sizes represent different scenarios of concurrently scheduling HPC and web applications. Uniform distribution depicts a situation where HPC and web applications are of the same magnitude. Left skewed distribution represents a situation where HPC applications to be scheduled are more than web applications and right skewed distribution represents the other way round. Uniform distribution represents a scenario where a single application type is scheduled.

The proposed algorithm can be applied to obtain optimal solutions to well defined problems on discrete space including, production planning, scheduling, inventory control, optimization of network synthesis and design problems, etc. These problems arise in real-world situations like management, engineering, telecommunications, etc.

The main contributions of the paper are:

- Clearer presentation of SOS procedures.
- Design and implementation of discrete version of SOS algorithm for scheduling of tasks in a cloud computing environment.
- Evaluation of the proposed method using makespan, response and degree of imbalance among VMs as performance metrics.
- Statistical validation of the obtained results against that of PSO using significance test.

The organization of the remainder of the paper is as follows. Metaheuristic algorithms applied to task scheduling problems in the cloud and SOS are presented in Section 2. Section 3 describes the problem formulation. Design of proposed algorithm and its description are presented in Section 4. Results of simulation and its discussion are in Section 5. Section 6 presented a summary and conclusion of the paper.

2. Related work

2.1. Metaheuristic algorithms in cloud scheduling

Metaheuristic methods [11,13,25,27,29,32–37] have been applied to solve task assignment problems in order to reduce makespan and response time. The methods have proven to find an optimum mapping of tasks to resources which reduce the cost of computation, improve quality of service, and increase utilization of computing resources. ACO, PSO, GA, and their variants are the mostly commonly used nature inspired population based algorithms in the cloud. PSO outperforms GA and ACO in most situations [16,31] and has faster execution time. PSO is simple to implement as compared to GA and ACO respectively. Workflow scheduling problems have been widely studied using PSO [25, 26,38–40] with the aim of reducing communication cost and makespan. Scheduling of Independent tasks has also been studied in cloud using PSO [18,33,34,41,42] and it has proved to ensure minimal makespan. Improved and hybrid versions [5,20,25,40–42] of PSO were also proposed for scheduling of tasks in the cloud, and they obtained better solution than those of ACO and GA.

2.1.1. Symbiotic Organism Search algorithm

Symbiotic Organism Search (SOS) algorithm, a novel population-based metaheuristic algorithm, was presented in [43] for solving numerical optimization problems on a continuous real space. SOS mimics the symbiotic associations (mutualism, commensalism, and parasitism) among different species in an ecosystem. Mutualism simply means the relationship between different species where both individuals benefit from the association. Commensalism is the association of two different species where one benefits from the union and the other is not harmed while in parasitism relation one species benefits and other is harmed. Each member of the organism within an ecosystem is represented by a vector in the solution plane. Each organism in the search space is assigned a value which suggests the extent of adaptation to the sought objective. The Algorithm repeatedly uses a population of the possible solutions to converge to an optimal position where the global optimal solution lies. The algorithm used mutualism, commensalism, and parasitism mechanisms to update the positions of the solution vector in the search space.

SOS is a repetitive process for an optimization problem [44] given in Definition 2.1. The procedure keeps a population of organisms that depict the candidate solutions of the studied problem. The relevant information concerning the decision variables and a fitness value is encapsulated into the organism as an indicator of its performance. Essentially, the trajectories of the organisms are modified using the phases of the symbiotic association.

Definition 2.1. Given a function $f : D \rightarrow \Re$ find $X' \in D : \forall X \in D f(X') \leq$ or $\geq f(X)$. \leq (\geq) minimization (maximization), where f is an objective function to be optimized and D represents the search space while the elements of D are the feasible solutions. X is a vector of optimization variables $X = \{x_1, x_2, x_3, \dots, x_n\}$. An optimal solution is a feasible solution X' that optimizes f .

2.1.2. Procedures of Symbiotic Organism Search

The steps of the Symbiotic Organism Search algorithm are given below:

Step 1: Ecosystem initialization

Initial population of the ecosystem is generated and other control variables such as ecosystem size, maximum number of iterations are specified. The positions of the organisms in the solution space are represented by real numbers.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات