



Queensland University of Technology
Brisbane Australia

This may be the author's version of a work that was submitted/accepted for publication in the following source:

Abdessamia, Foudil, Zhang, Wei Zhe, & [Tian, Yu-Chu](#)
(2020)

Energy-efficiency virtual machine placement based on binary gravitational search algorithm.

Cluster Computing, 23(3), pp. 1577-1588.

This file was downloaded from: <https://eprints.qut.edu.au/200239/>

© **Springer Science+Business Media, LLC, part of Springer Nature**
2019

This work is covered by copyright. Unless the document is being made available under a Creative Commons Licence, you must assume that re-use is limited to personal use and that permission from the copyright owner must be obtained for all other uses. If the document is available under a Creative Commons License (or other specified license) then refer to the Licence for details of permitted re-use. It is a condition of access that users recognise and abide by the legal requirements associated with these rights. If you believe that this work infringes copyright please provide details by email to qut.copyright@qut.edu.au

License: Creative Commons: Attribution-Noncommercial 4.0

Notice: *Please note that this document may not be the Version of Record (i.e. published version) of the work. Author manuscript versions (as Submitted for peer review or as Accepted for publication after peer review) can be identified by an absence of publisher branding and/or typeset appearance. If there is any doubt, please refer to the published source.*

<https://doi.org/10.1007/s10586-019-03021-0>

Energy-efficiency Virtual Machine Placement based on Binary Gravitational Search Algorithm

Foudil Abdessamia · Wei Zhe Zhang a,b* · Yu-Chu Tian

Received: date / Accepted: date

Abstract cloud computing is a remarkable growing paradigm for hosting and offering services through the Internet. It attracted the most notorious business companies and resulted to an exponential increase of its users from simple end users to companies that deploy more and more of their system over the cloud. The amount of resources to provide the demand became tremendous. therefore, a great need energy supply. The world as we know is highly concerned about the environment and the energy-efficiency in all aspect of life and the domain of IT is one them. To deal with energy wastage in data centers, researches use Virtual machine placement as a main key to assure cloud consolidation and reduce power wastage. Several approaches were proposed for Virtual machine placement. This paper proposes a solution based on Binary gravitational search algorithm (BGSA) for the virtual machine placement in the heterogeneous data center. In this work, we compared the BGSA method to fit with virtual machines in data centers with particle swarm optimization, First-fit, Best-fit, and worst-fit. results showed significant difference of energy save comparing to other strategies. The results obtained gave the advantage to our

approach and its better response with the increase of number of virtual machines.

Keywords cloud computing · Green computing · Energy-efficiency · Optimization

1 Introduction

The world of technology and internet gained a huge success in the recent years, it allowed the emergence of the cloud computing as new model for scalable computing resources and on demand services model or also called “pay as you go” model. These services are provided to users, and access to them through the internet in three levels, Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) [1]. In one hand, computing resources are way cheaper than before and in the other hand they are more powerful and with high performance. In this case, the end users or consumers are provided with computing resources such as CPU run time and storage over the internet. The providers manage the allocation of these resources to consumers following a pricing model. The information technology IT was impacted by this change, the big names in the industry of IT rapidly invested this business. They compete, offer and provide more powerful platforms, less expensive and high quality of service. Leaders in the IT industry such as Amazon, Microsoft, Google and Alibaba offer now several business model for the cloud computing to attract the potential consumers. It is also noted that companies tend to choose and build their IT system over the cloud and 51% of them claim that they have built new polices for could usage (Figure 1) [2]. The cloud computing allows this companies as a consumer to lower the IT barrier and open more opportunities to innovation, as it can be seen in start-ups. In

Foudil Abdessamia
School of Computer Science and Technology,
Harbin Institute of Technology, Harbin, China
E-mail: foudil.abdessamia@hit.edu.cn

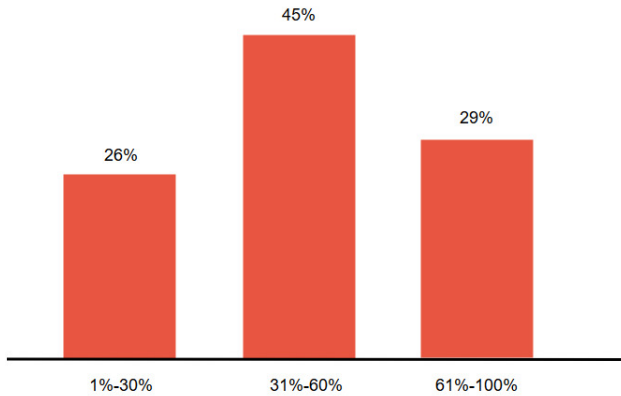
Wei Zhe Zhang *corresponding author
School of Computer Science and Technology,
Harbin Institute of Technology, Harbin, China

Peng Cheng laboratory. Shenzhen, China
E-mail: wzzhang@hit.edu.cn

Yu-Chu Tian
School of Electrical Engineering and Computer Science,
QUT, Brisbane QLD, Australia
E-mail: y.tian@qut.edu.au

other words, with cloud computing, providers can offer services that meet the needs of consumers without any excess requirement and adequate pricing. In the same time, the resources are scalable and easy to expand.

Fig. 1: Percentage of IT systems that are cloud-based. [2]



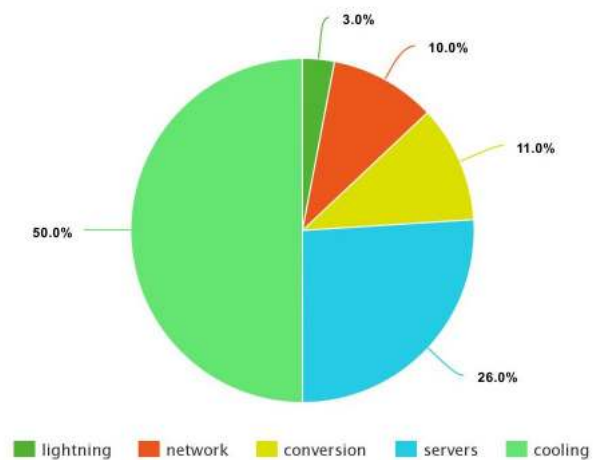
Moreover, new governmental politics focus and invest in the cloud and build data centers for their own usage, known as e-government.

The main side effect related to the success of cloud computing is the growth of energy consumption. It highlights the energy cost of powering and cooling the equipment. This has an impact in economy and environment. In fact, the world is facing big issues of global warming, the active players in economy, associations and government are taking responsibilities and proposing new policies to reduce the climate change and take earth friendly initiatives. In 2016, the Paris agreement brought together countries, states and governments to deal with CO₂ gas emission with a goal to keep the increase in global average temperature to well below 2 degrees C. In parallel, conferences are also organized to increase the awareness about the global warming and gas emission within society, industry and politics. They propose solutions and warn the potential danger from natural disasters and health problems.

The information technologies have been involved with the environment problem these recent years. The main issues are related to the energy consumption. In fact, with the fast growth and development of IT, it became a large fraction of business' energy costs and CO₂ emission. It is estimated that the IT industry account for 2% of global CO₂ emissions, a percentage equivalent to the amount generated by the aviation industry [3]. To deal with these problems, green information tech-

nology was emerged and opened new research topic. It refers to the practices and initiatives to use technology in favor of environment cause, such as reducing waste in components manufacturing, eco-friendly products and power efficiency. In 2013, U.S. data centers consumed 91 billion kWh of electricity -2.4% of total electricity consumption- at a cost of \$13 billion (U.S. Energy Information Administration (EIA) (2016) Monthly Energy Review May 2016). Data centers are one of the main part of the IT and cloud computing business' success. The huge usage of energy and hardware put the cloud computing and its big data centers in the front of the green IT and green computing issues to deal with. Info tech statistics (Figure 2) show that a half of the consumed energy in a data center is due to the cooling infrastructure (50%), the second part of power consumption goes to servers and storage devices (26%), the remained part is shared between network equipment, lightning, and others [4]. In a such model, the increase use of the hardware equipment releases more heat and directly affect the cooling system energy consumption.

Fig. 2: Typical Data Center Energy Consumption.



The data center is consisted of a large number of running servers equipped with processors, storage and network components. Measurement of power consumption in the idle state of a standard server showed that it could use up to 66% of the peak power, this can be explained by the running operation system (OS) and hardware power need [5]. The under-use of the servers will conduct to energy waste, therefore, reducing the number of physical machines (PM) in data centers will result to lower the energy consumption and consequently the energy cost in the business context. Furthermore, the cooling cost will decrease as a result of a

low energy consumption and enhanced heat dissipation.

Faced to these constraints, developers of the cloud computing services are required to provide solutions to insure an energy consumption reduction, in the meanwhile, deliver products with a standard service-level agreement (SLA) to end users. Therefore, cloud computing relies on the technology of virtualization. The main goal of virtualization is to abstract hardware and share the physical resources to multiple virtual machines (VM), this solution gives more isolation of users within data centers and energy saving [6]. A new main research topic concerning the VM allocation is brought to light. developers try to find the best strategies on how these VM are placed in servers and ensure the convenient delivery.

In this study, we proposed an energy-efficient virtual machine placement to face these issues in the area of cloud computing. We propose an optimization method based on Gravitational search algorithm. We adapted GSA in order to find an optimum placement for the virtual machine in the data center. This method shows better results than known methods such as best-fit, worst-fit and has better result than particle swarm optimization which is also based on population search for solution.

Paper organization: The remainder of this paper is organized as follows: Section 2 we survey related works about virtual machine placement in Section 3; we present a description about the virtual machine placement and the Gravitational Search algorithm in Section 4 we describe the problem of the energy-efficiency in virtualized environment; Section 5 showcase the BGSA for the VMP. we describe the algorithm in 6. We interpret the simulation and in Section 7 and we conclude in 8.

2 Related work

The area of cloud computing brings many issues, these problems attracted the interest of researchers and proposed solutions. The area of cloud computing brings many issues, these problems attracted the interest of researchers and proposed solutions. The works were interested to the problem of energy saving in cloud computing in many angles in the aim to reduce the power consumption. As a result, the concept of green cloud computing came into existence. The main Energy Efficiency Measures in the Data Center are [7]:

- **Infrastructural changes:** it affects the equipment related to the data centers such as cooling system and quality of the equipment. Providers of and owners of data centers make choices to provide equip-

ment with better energy usage and offer green building.

- **Data center location choice:** the location of the data center is very important and also has major effect on the architecture of the installation and its power use, very common now to choose cold area as location for data centers and use natural weather to cool the center. in other cases, the data centers uses energy issued from eco-friendly equipment such as wind turbines, the data center would be near places with high speed wind.
- **Hardware oriented optimizations:** as the name suggest, it's related to techniques and procedures on hardware to reduce the energy consumption. Deploy new servers with better energy use and more energy proportional. Optimization on parallel and multicore architecture. And finally the network and storage equipment optimization.
- **Software oriented optimizations:** it's the trend of developing energy-efficient algorithms for effectively provisioning resources to the tasks. Proposing strategies to better allocate the processes or virtual machine in cloud.

In this work we are more interested in the Software oriented optimization and there are multiple techniques and algorithms used to minimize the energy consumption in data centers.

Dynamic voltage and frequency scaling

(DVFS): is the adjustment of power and speed settings on a computing device's processors, controller chips and peripheral devices to optimize resource needed for tasks and maximize power saving when those resources are not needed. The technique is to decreased the clock frequency of a processor and reduce the supply voltage [8]. Several works were interested to use this technique in the field of cloud computing and data centers. Polices and methods were proposed using DVFS to reduce the energy consumption within data centers. The main idea is to find allocate resources to tasks or virtual machines in the data centers. In [9] proposes a DVFS policy that reduces power consumption while preventing performance degradation, and a DVFS-aware consolidation policy that optimizes consumption, using CloudSim toolkit and real Cloud traces and power model, results of the proposed work shows a saving up to 42 %. The same approach is used in [10] where they used DVFS to allocate virtual machines and make a compromise balance in between energy consumption and the set up time of different modes of hosts or VMs. They proposed a DVFS- energy model and algorithm for this purpose, result of the experiment compared to a random situation showed a positive reduce of energy consumption.

Virtualization:the abstraction given in Virtualization allows more possibilities to use the physical resources by the OS and a lot of scenarios emerges on how to use these resources in efficient way. as a result, many papers and work fall in this area of research.

Virtual Machine Allocation- and Scheduling-Based Techniques:

Describing the problem as Bin Packing, two well-known types of the heuristic methods was proposed at the early stage, the First Fit Decreasing (FFD) and Best Fit (BF). The two methods are simple to Implement and have good scalability [11]. However, it showed some weaknesses to deal with multiple dimensions. In [12] VM placement problem is abstracted as a combination of bin packing problem and quadratic assignment problem, and design a greedy algorithm by combining minimum cut with the best-fit.

In [13] considers the VM placement problem can be seen as a bin-packing problem and proposes a derivation of the best-fit heuristic to solve the problem. Servers have a carbon footprint in addition of energy consumption. The broker will place VMs reduces the CO2 emission and power consumption

Genetic algorithms (GA):

Other works explored the idea of using Genetic algorithms (GA), inspired form natural evolution processes, it proposes an optimization solution after a numerous generation in [14] considering both of the physical machine and the communication network energy consumption the data center, they have proposed a genetic algorithm for a new virtual machine placement problem. In addition to GA [15] applied different meta-heuristics and used for this problem which is Non-dominated Sorting Genetic Algorithm(I and II).

Their objectives were maximizing load balance and minimizing resource wastage in data center.

Bio inspired algorithms :

such as Ant Colony optimization (ACO) was present several works for virtual machine placement. In [16] they proposed the Ant Colony Optimization (ACO) for multidimensional bin packing problem. Comparing to greedy algorithm (FFD) their solution provided more gain in energy consumption. in [17] also based their proposition on (ACO) named ACO-VMP. the ant colony system (ACS)-based approach is developed to achieve the VMP goal. Coupled with order exchange and migration (OEM) local search techniques, it effectively minimizes the number of active servers used for the assignment of VMs .

optimization heuristic methods :

in [18] proposed novel adaptive heuristics that are based on an analysis of historical data on the resource usage for energy and performance efficient dynamic con-

solidation of VMs. They have evaluated their proposed algorithms through extensive simulations on a large-scale experiment setup using workload.

In [19] [20] [21], the work is based on Particle swarm optimization to find the best virtual machine placement strategy for energy efficiency of data centers in a cloud environment. Comparing to Best-Fit, First-Fit and Worst-Fit, this approach gives the better strategy for the cloud to reduce the energy consumption.

virtual machine consolidation :

the virtualization concept enables the possibility to reduce the number of active physical machines by leveraging live virtual machine migration. [22] presents a distributed approach to an energy-efficient dynamic virtual machine consolidation mechanism. Their approach virtual select machines to migrate, and time to do it. Then, the placement of the virtual machines selected for migration is achieved based on a generalization of the Knapsack known The results of experiment done in CloudSim using the real workload data, the results of the under-load detection methods proposed outperformed the other compared methods. In [23], achieved power saving through power efficient consolidation of virtual machines on a smaller number of servers by proposing algorithms for the power-aware allocation and migration of virtual machines.

diverse works :

In [12] the aim was to reduce and minimize the energy consumption Researchers addressed the problem of virtual machine placement as an NP-complete problem and proposed a virtual machine placement algorithm EAGLE that can balance the utilization of multi-dimensional resources and reduce the number of running virtual machines result of their simulation show that it reduces 15 % more energy than the first fit algorithm. Simulation results show that the approach can save as much as 15 % energy compared to the first fit algorithm.

in [24] they proposed an approach named EnaCloud, using Virtual Machine to encapsulate application and enables application live placement dynamically with consideration of energy efficiency in a cloud platform. supports applications scheduling and live migration to minimize the number of running machines the application placement is abstracted as a bin packing problem, and an energy-aware heuristic algorithm is proposed to get an appropriate solution

in [25] proposed different approach using framework that finds the best possible placement of virtual machines based on constraints expressed through service level agreements. The framework's flexibility is achieved by decoupling the expressed constraints from the algorithms using the Constraint Programming (CP). the

experimental and simulation results demonstrate the effectiveness and results shown that the presented approach is capable of saving both a significant amount of energy and CO2 emissions in a real world scenario on average 18%.

Work In [26] proposes a Simulated Annealing based algorithm, for VM placement problem to optimize the power consumption. Experimental results show that this SA algorithm can generate better results, saving up to 25 % more energy than First Fit Decreasing in an acceptable time frame. However, the authors mentioned that the proposed algorithm cannot handle migration well.

in [27] they consider the hypothesis that the VM could exist in q number of copies, and place them in servers the VM placement algorithm is based on the dynamic programming and local search methods, the dynamic programming method determines the number of copies for each VM and tries to minimize the energy cost by turning off the underutilized servers Simulation results show that the proposed algorithm reduces the total energy consumption by up to 20% with respect to previous work.

In [28] the paper proposes a joint energy-aware and application aware VM placement strategy based on the theory of multi objective optimization. In addition to server-side constraints their approach analyzes the impacts on VM placements of the dependencies among data center infrastructure and reduces network communication.

More recent work [usman2019] proposes an Energy-oriented Flower Pollination Algorithm for VM allocation based on Dynamic Switching Probability. By exploring the local search, the framework finds a near optimal solution. The performance is simulated on the framework CloudSim compared to Genetic Algorithm for Power-Aware (GAPA), ant colony system, and First Fit Decreasing. It showed significant saving power.

In [29] the work sees the issue of the electricity cost, in other word reduce energy consumption. They first study the impact of load placement policies on cooling and maximum data center and propose dynamic load distribution policies that consider all electricity-related costs it also concentrates on data availability to reduce migration and SLA violation. They compared their result to Round-Robin (RR). Worst Fit (WF). Static Cost-Aware Ordering (SCA).

3 Background

IT infrastructure and cloud computing relies essentially on virtualization, it allows optimization of the use of

physical resources, portability, ease of deployment, high availability of virtual machines. A virtual infrastructure has integrated high availability and removes many issues related to the hardware. In virtual machine architecture, several virtual machines (VMs) can share the same physical machine (PM). The Virtual machine monitor (VMM) or the Broker in other cases provides resource allocation to VM. Above the hardware layer, the VMM provides a consistent view of underlying hardware to the “OS guests” and offer the virtual machines a high level of abstractions to the application running on the VM [30], [11]. The VMM can also provide and share a pool of hardware resources between VMs and ensure isolation of these VMs with a same single PM. Virtual machine placement is the procedure that maps virtual machines to physical machines. The main aim is to chose the best host for the virtual machine. The criteria of the placement may vary depending on the final goal. In the case of energy efficiency, the aim of the placement is to minimize the energy and save the power. It can also maximize the usage of the pool of resources available. Virtual machine placement algorithms try to assure the goals above [31]. The VM may need to re-allocate to different server to fulfill some condition of QoS or to access to different data. The Virtual machine migration is a procedure which takes an entire running virtual machine from a physical machine and moves it to another. The migration must be transparent to the operation system and the application [32], [33].

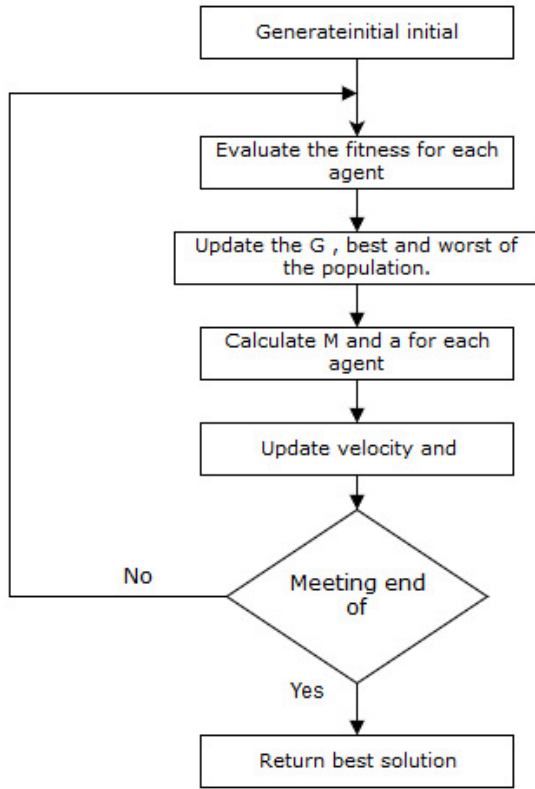
BINARY Gravitational Search Algorithm

For this work, we consider the problem as NP hard problem, the first objective is to place the virtual machines in servers with energy aware strategy, in this work we used a meta heuristic named Gravitational Search Algorithm (GSA).

GSA is based on the law of gravity [34]. In the proposed algorithm, agents are considered as objects and their performance is measured by their masses. As known in the word of physics, these objects attract each other by the gravity force. As a result of force, a global movement of all objects towards the objects with heavier masses. Thus, the object with the heavy mass mean it has a better solution than other objects. To guaranty the exploitation step of the algorithm, heavy masses move more slowly than lighter ones.

In GSA, each mass (agent) has four specifications: position, inertial mass, active gravitational mass, and passive gravitational mass. The position of the mass corresponds to a solution of the problem, and its gravitational and inertial masses are determined using a fitness function. To navigate through the search space of

Fig. 3: Gravitational search algorithm



solution, algorithm adjusts the gravitational and inertia masses. In every iteration, the masses are attracted by the heaviest mass. This mass will present an optimum solution in the search space. As in the Newtonian laws of gravitation and motion, masses in GSA obey two main laws [34]

Law of gravity: each particle attracts every other particle and the gravitational force between two particles is directly proportional to the product of their masses and inversely proportional to the distance between them R .

Law of motion: the current velocity of any mass is equal to the sum of the fraction of its previous velocity and the variation in the velocity. Variation in the velocity or acceleration of any mass is equal to the force acted on the system divided by mass of inertia. Now, consider a system with N agents (masses). We define the position of the i -th agent by:

$$X_i = (x_i^1, \dots, x_i^d, \dots, x_i^n) \text{ for } i = 1, 2, \dots, N \quad (1)$$

Where x_i^d presents the position of i -th agent in the d -th dimension. At a specific time 't', force acting on mass 'i' from mass 'j' is defined as following:

$$F_{ij}^d(t) = \frac{M_{pi}(t) * M_{aj}(t)}{R_{ij}(t) + \epsilon} (x_j^1(t) - x_i^1(t)) \quad (2)$$

Where M_{ai} is the active gravitational mass related to agent j , M_{pi} is the passive gravitational mass related to agent I , $G(t)$ gravitational constant at time t , ϵ is a small constant, and $R_{ij}(t)$ is the Euclidean distance between two agents i and j :

$$R_{ij}(t) = \|X_i(t) - X_j(t)\|^2 \quad (3)$$

in [8] proposed to consider the total force that acts on agent i in a dimension d to be a randomly weighted sum of d -th components of the forces exerted from other agents to give a stochastic characteristic to the algorithm (eq14)

$$F_i^d(t) = \sum_{j=1; j \neq i}^N rand_j F_{ij}^d(t) \quad (4)$$

Where $rand_j$ is a random number in the interval $[0,1]$. however, by the law of motion, the acceleration of the agent I at time t , and in direction d -th, is given as follows:

$$a_i^d(t) = \frac{F_i^d(t)}{M_{ii}(t)} \quad (5)$$

Where M_{ii} is the inertial mass of i -th agent.

to calculate the next velocity of an agent, the algorithm takes a fraction of its current velocity added to its acceleration. The position and velocity could be calculated as follows:

$$v_i^d(t+1) = rand_i * v_i^d(t) + a_i^d(t) \quad (6)$$

$$x_i^d(t+1) = x_i^d(t) + v_i^d(t+1) \quad (7)$$

Where $rand_i$ is a uniform random variable in the interval $[0,1]$. BGSA use this random number to give a randomized characteristic to the search. The gravitational constant, G , is initialized at the beginning and will be reduced with time to control the search accuracy. In other words, G is a function of the initial value (G_0) and time (t):

$$G(t) = G(G_0, t) \quad (8)$$

in the gravitational search algorithm the heavier mass means a more efficient agent. therefore, they have higher attractions and move slowly in the search space. gravitational and inertia masses are simply calculated by the fitness evaluation. Assuming the equality of the gravitational and inertia mass, the values of masses are calculated using the map of fitness. the gravitational and inertial masses update by the following equations:

$$M_{pi} = M_{ai} = M_{ii} = M_i \text{ for } i = 1, 2, \dots, N \quad (9)$$

$$m_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \quad (10)$$

$$M_i = \frac{m_i(t)}{\sum_j^N m_i(t)} \quad (11)$$

Where $fit_i(t)$ represent the fitness value of the agent i at time t . $worst(t)$ and $best(t)$ are defined as follows (for a minimization problem):

$$best(t) = \min_{i \in \{1, \dots, N\}} fit_i(t) \quad (12)$$

$$worst(t) = \max_{i \in \{1, \dots, N\}} fit_i(t) \quad (13)$$

4 Problem statement

The formulation problem in this work is defined as follow:

Notation	Description
V_m	a set of virtual machines
P_m	a set of physical machine
v_i	virtual machine in V_m
v_i^{cpu}	the CPU requirement of v_i
v_i^{mem}	the memory requirement of v_i
p_j	a physical machine in P_m
p_j^{cpu}	the CPU capacity of p_j
p_j^{mem}	the memory capacity of p_j
p_j^{Wcpu}	the total CPU workload of p_j
p_j^{Wmem}	the total memory workload of p_j
V_{pj}	the set of virtual machines assigned to p_j

The utilization of a CPU in one server is

$$U_j = \frac{p_j^{Wcpu}}{p_j^{cpu}} \quad (14)$$

$$p_j^{Wcpu} = \sum v_i^{cpu} ; v_i \in V_{pj} \quad (15)$$

Power energy model

The energy consumption is related to the utilization of the resources of the server; this relationship was described as linear relation [35]. Hence in this study, we take that power consumption is linearly related to CPU utilization. We can calculate the use of VM in a single CPU in 15. The CPU use of a PM is calculated as in 14; it is the total of CPU use of all the virtual machines running on the j -th machine.

The power consumption of a physical machine is defined as in Eq.16

$$P_j = \begin{cases} (P_j^{busy} - P_j^{idl}) * U_j + P_j^{idl} ; p_j^{Wcpu} > 0 \\ P_j^{idl} ; otherwise \end{cases} \quad (16)$$

In which,

- P_j^{idle} : the power consumption (Watt) of the host in Idle (0% of utilization)
- P_j^{busy} : the power consumption (Watt) of the host in maximum (100% of utilization)
- P_j : current power consumption of the j -th host

The aim of this study is to reduce the power consumption in a data center then we aim is to minimize 17.

$$\sum_{j=1}^m P_j = \sum_{j=1}^m [(P_j^{busy} - P_j^{idl}) * U_j + P_j^{idl}] \quad (17)$$

5 GSA for virtual machine placement

For the problem of virtual machine placement, we used an improved gravitational search algorithm, a binary version of GSA; it is more suitable for the problem. For every agent A we define it position by the matrix A_k :

$$A_k = \begin{bmatrix} x_k^{11}, x_k^{12}, \dots, x_k^{1n} \\ x_k^{21}, x_k^{22}, \dots, x_k^{2n} \\ \vdots \\ x_k^{m1}, x_k^{12}, \dots, x_k^{mn} \end{bmatrix} \quad (18)$$

A_k is the matrix position of the k object, when the i -th virtual machine is assigned to the j -th server, the correspondent bit is equal to 1 such in $A_k = 1$, otherwise equal 0.

The updating procedure of the force, acceleration and velocity may be considered similar to the continuous algorithm. The main difference between continuous and binary GSA is that in the binary algorithm, the position updating means a switching between “0” and “1” values. The mass velocity will the responsible function

to switching the correspondent bit for the position. the goal is to update the position in a way that the current bit value is changed with a probability that is calculated according to the mass velocity. That means the BGSA updates the velocity and considers the new position to be either 1 or 0 with the given probability.

to understand BSGA here some basics about the algorithm behavior

- the absolute value of the velocity shows how much the current position should move. larger value means the position of the mass is far from the optimum position
- in the case of a small absolute value of the velocity means that the current position of the mass is close to the optimum position and there is a small distance reaching to the optimum position. Then, , the velocity becomes converges to zero.

To implement the BGSA algorithm, the following concepts should be taken into account:

- A large absolute value of velocity must provide a high probability of changing the position of the mass respect to its previous position (from 1 to 0 or vice versa).
- A small absolute value of the velocity must provide a small probability of changing the position. In other words, a zero value of the velocity represents that the mass position is good and must not be changed.

Based on the above-mentioned concepts, few probability function must be implemented such that for a small $|v_i^d|$, the probability of changing x_i^d must be near zero and for a large $|v_i^d|$, the probability of x_i^d movement must be high. We define function $S(v_i^d(t))$ to transfer x_i^d into a probability function. $S(v_i^d(t))$ should be bounded within interval $[0,1]$ and increases with increasing $S|v_i^d(t)|$. $S(v_i^d(t))$ is defined :

$$S(v_i^d(t)) = |\text{tranh}(v_i^d(t))| \quad (19)$$

Once $S(v_i^d)$ is calculated, the agents will move according to

$$\begin{cases} x_i^d(t+1) = \text{complement}(x_i^d(t)) ; \text{rand} < S(v_i^d(t+1)) \\ x_i^d(t+1) = (x_i^d(t)) ; \text{otherwise} \end{cases} \quad (20)$$

As it is said above, the position in the case of the virtual machine placement is a matrix, the two dimensions of the matrix are the the set of VM and Servers. Therefore, for our virtual machine placement we must define another matrix for velocity of an agent. The velocity matrix with same dimensions contains binary value.

6 algorithm

Based on the BGSA algorithm and the new definitions the virtual machine placement, the implementation would be presented as follow at first we need to allocate the necessary matrices for the Algorithm 1.

Step1 : Initialization

- We define the list of server resources as vector with size of m , m is the number of servers in the data center.
- We define the list of VMs as vector with size of m , n is the number of servers in the data center.
- the number of agent and iteration

Step2: Every agent position is set randomly. the position must be in the search space.

Step3: In every iteration we calculate the fitness of all agents of the initial positions. It represent the mass of every agent; update the fitness value if it is better than the old fitness.

Step 5: For every agent we calculate the Mass based on the new fitness the gravitational value is updated as the iteration increase, it the change of the value in function of time. The acceleration is calculated with the value of the new mass and gravitation.

Step 5: At this step we lunch the function to move the agents in the search space, which means that the new velocities will create the movement for every agent Algorithm 2.

Algorithm 1: BGSA for virtual machine placement

Input : List of VMs , List of Servers
Output: matrix for optimum position

```

1 initialization;
2 while (iteration  $j$  max iteration) do
3   Calculate fitness;
4   Update fitness;
5   Calculate M;
6   Calculate Gravitational constant;
7   Calculate acceleration;
8   move();
9 end
10 return position;
```

7 simulation and results

To implement the improved BSGA algorithm we used Matlab version R2015b running on Windows 7 64 bits,

Algorithm 2: function to calculate new position

Input : acceleration , velocity ,current position
Output: new position position

- 1 Calculate new velocity ;
- 2 Calculate S ;
- 3 Generate the movement to be done;
- 4 **for** (*every agent*) **do**
- 5 find the bit to change;
- 6 Update position;
- 7 correct the position;
- 8 **end**
- 9 return new position;

and a processor Intel i7 3.4 GHz accompanied with 16 Go of memory. All the functions needed in the algorithms were implemented in Matlab in addition to the compared algorithms. To have an accurate comparison, we use the same set of virtual machines for all algorithms and the same amount of physical resources. We simulate a cluster of heterogeneous servers that have different energy consumption with static resources like CPU and memory, the energy proportionality is different when we have different server equipment in the data center. In this simulation, we used to the type of servers as shown in table 1.

Table 1: servers setup

	kWh	MIPS	memory
server 1	140	15000	16Go
server 2	100	10000	8go

The number of servers in every simulation is set to support the worst allocation scenario. For the virtual machine, we used tree types; the low, medium and high need of resources Table.2.

Table 2: virtual machine setup

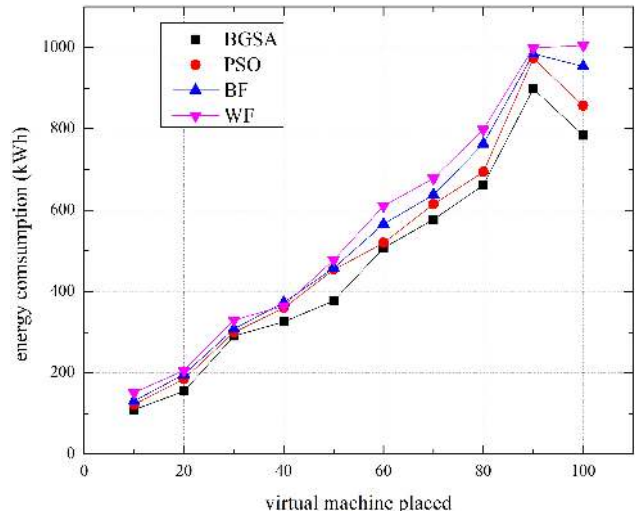
	MIPS	memory
VM type 1	1500	2Mo
VM type 2	1000	1Mo
VM type 3	500	512Ko

the result of simulation of the compared algorithms is shown in Figure 4 .all algorithm used the same setup of servers and placed the same set of VM based on table1 and 2.

For the first simulations, we tried to set place under 100 of VM and compare the energy consumption of the

over all servers, we notice that for the small amount of VM to be placed the algorithms give quite similar results due to the low complexity. it give us a positive sign that the BGSA is looking for the optimum solution and works as we wanted. for much higher number of VM , the divergence start and algorithms give different result; however our proposed solution still has a better energy with a slight difference. Under 50 VMs the placement problem is simple, the optimum solution is very quick to obtain that is why all the methods used in the simulation get very similar result. As the number of VMs increases our approach react well to the placement problem and gets better energy while the other strategies diverges from the optimum.

Fig. 4: Power consumption obtained using different virtual machine placement and based on setup table 1



the Figure 5 shows the number of servers active after every virtual machine placement for the different strategy placement. As expected, the number of active servers is low in the case of BGSA and that explains the lower energy consumption.

the Figure 6 shows the results of the same simulation with only VM type 2 in Table 2 , we wanted to explore the possibilities and the performance of the algorithms in different situations. GSA and PSO have better energy consumption than Best-fit and Worst-fit. however, when the number of VM is increase, the BGSA also gradually gives better energy consumption overall.

BGSA and PSO have similarities in their definition and how they work. the two algorithm have a number of agent and particles (for PSO) which represent a possible solution to a problem and for a number of iteration the global agent and particles converges to an optimal

Fig. 5: number of active server in every virtual machine placement

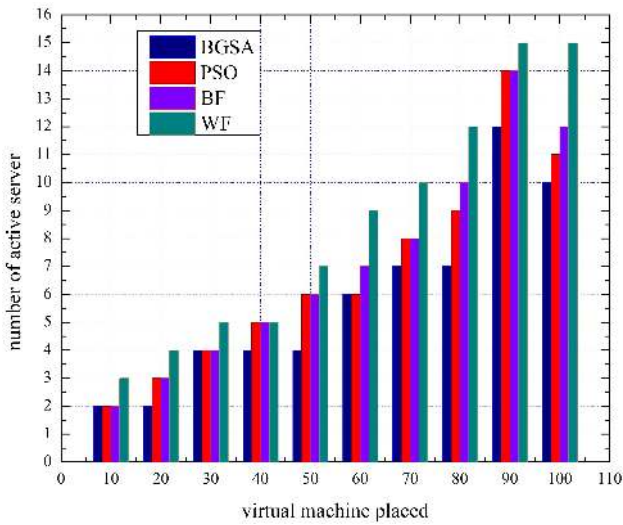
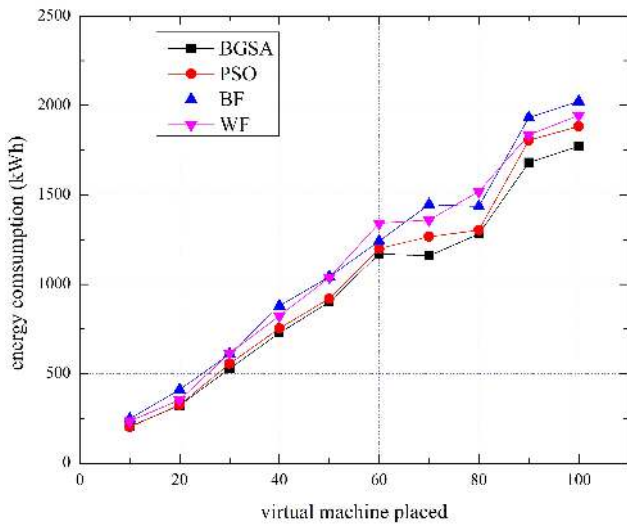


Fig. 6: Power consumption obtained using different virtual machine placement and based on setups table 2

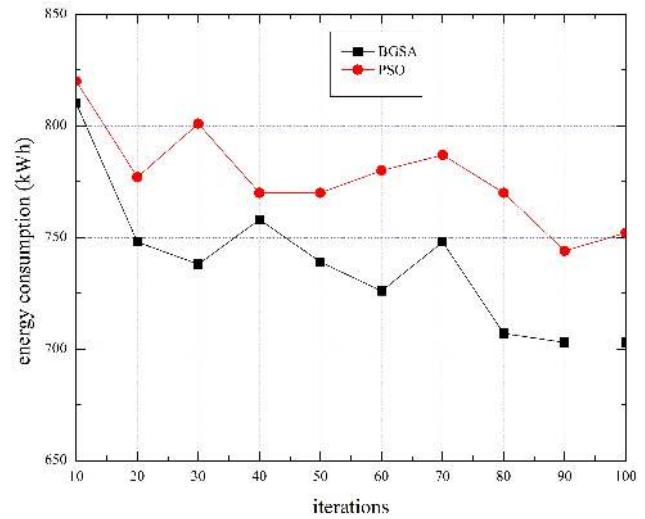


solution. we run different simulations for the two algorithms and changed the parameters of iteration and fix the number of agent or particles.

We made sure that for every simulation, the set of virtual machine to be placed are the same. therefore, we can notice from Figure 7 that in BGSA it reaches the optimum with less iterations than the PSO. BGSA clearly needs less iterations which means less time to get to the optimal solution.

For next simulation, we want to increase the number of VM and increase the diversity of the setups to cre-

Fig. 7: Power consumption for BGSA and PSO



ate more heterogeneous system. we present the second setups in table 3 and 4

Table 3: servers setup 2

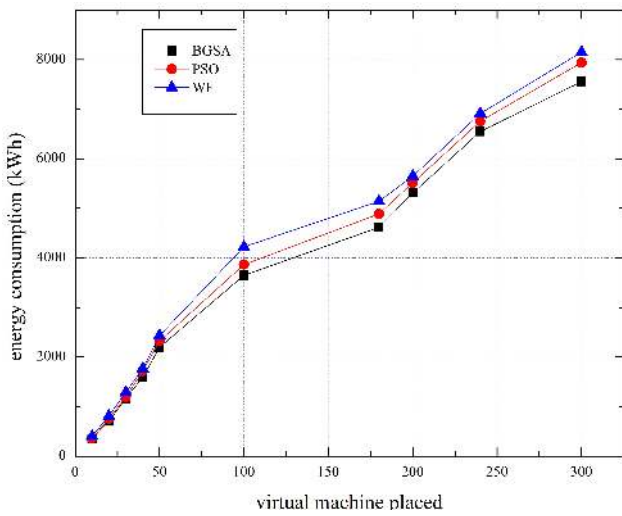
	kWh	MIPS	memory
server 1	140	15000	16Go
server 2	130	10000	8go
server 2	120	8000	4go
server 2	100	4000	2go

Table 4: virtual machine setup 2

	MIPS	memory
VM type 1	2500	2.5Mo
VM type 1	2000	2Mo
VM type 1	1500	1.5Mo
VM type 2	1000	1Mo
VM type 3	500	512Ko

In the second simulation we tried to place upto 300 VM with the setups in table 3 and 4, BGSA gave better results for the energy consumption. Under 50 virtual machine, the amount is not that important to make different results between the virtual machine placement. In this simulation, we made sure that the number of VM is higher than the precedent simulation and we noticed that our approach was still outperforming the other strategies and specially PSO which has similar way to find optimal solution based on population search (particle in the case of PSO and agents for BGSA).

Fig. 8: Power consumption obtained using different virtual machine placement and based on setups type 2



8 Conclusion

Energy-efficiency in the cloud computing became one of the main practice for green computing. to deal with this issue researchers and works were made to treat the problem of resource wastage, energy consumption and CO2 emission. virtual machine placement allow developers to propose solutions for better use of the resources. In this paper, we proposed a strategy for virtual machine placement based on Gravitational search algorithm to find the optimal energy consumption. The Binary version of the GSA appeared to be more suitable in our work for its compatibility with our definition of the problem. We compared the results with Particle swarm optimization, First fit, Best fit and worst fit strategies. In our approach proved its advantage and shows better energy consumption in different setups and situations proposed better use of the resource. This work can be extended in future with the possibility of combining other methods and use advantages for better Energy-efficiency.

9 acknowledgment

This work was supported in part by the National Key Research and Development Plan under Grant 2017YFB0801801, in part by the National Science Foundation of China (NSFC) under Grant 61672186 and Grant 61872110. Professor Zhang is the corresponding author.

References

1. Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: state-of-the-art and research challenges," *Journal of internet services and applications*, vol. 1, no. 1, pp. 7–18, 2010.
2. "2018 trends in cloud computing."
3. "Green it: The new industry shock wave."
4. M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 732–794, 2015.
5. A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future generation computer systems*, vol. 28, no. 5, pp. 755–768, 2012.
6. G. Chen, W. He, J. Liu, S. Nath, L. Rigas, L. Xiao, and F. Zhao, "Energy-aware server provisioning and load dispatching for connection-intensive internet services," in *NSDI*, vol. 8, pp. 337–350, 2008.
7. T. Kaur and I. Chana, "Energy efficiency techniques in cloud computing: A survey and taxonomy," *ACM computing surveys (CSUR)*, vol. 48, no. 2, p. 22, 2015.
8. E. Le Sueur and G. Heiser, "Dynamic voltage and frequency scaling: The laws of diminishing returns," in *Proceedings of the 2010 international conference on Power aware computing and systems*, pp. 1–8, 2010.
9. P. Arroba, J. M. Moya, J. L. Ayala, and R. Buyya, "Dynamic voltage and frequency scaling-aware dynamic consolidation of virtual machines for energy efficient cloud data centers," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 10, p. e4067, 2017.
10. S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi, "Cloud computing—the business perspective," *Decision support systems*, vol. 51, no. 1, pp. 176–189, 2011.
11. M. Rosenblum and T. Garfinkel, "Virtual machine monitors: Current technology and future trends," *Computer*, vol. 38, no. 5, pp. 39–47, 2005.
12. X. Li, Z. Qian, S. Lu, and J. Wu, "Energy efficient virtual machine placement algorithm with balanced and improved resource utilization in a data center," *Mathematical and Computer Modelling*, vol. 58, no. 5-6, pp. 1222–1235, 2013.
13. A. Khosravi, S. K. Garg, and R. Buyya, "Energy and carbon-efficient placement of virtual machines in distributed cloud data centers," in *European Conference on Parallel Processing*, pp. 317–328, Springer, 2013.
14. M. Tang and S. Pan, "A hybrid genetic algorithm for the energy-efficient virtual machine placement problem in data centers," *Neural Processing Letters*, vol. 41, no. 2, pp. 211–221, 2015.
15. A. C. Adamuthe, R. M. Pandharpatte, and G. T. Thampi, "Multiobjective virtual machine placement in cloud environment," in *2013 International Conference on Cloud & Ubiquitous Computing & Emerging Technologies*, pp. 8–13, IEEE, 2013.
16. E. Feller, L. Rilling, and C. Morin, "Energy-aware ant colony based workload placement in clouds," in *Proceedings of the 2011 IEEE/ACM 12th International Conference on Grid Computing*, pp. 26–33, IEEE Computer Society, 2011.
17. X.-F. Liu, Z.-H. Zhan, J. D. Deng, Y. Li, T. Gu, and J. Zhang, "An energy efficient ant colony system for virtual machine placement in cloud computing," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 113–128, 2016.

18. A. Beloglazov and R. Buyya, "Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers," *Concurrency and Computation: Practice and Experience*, vol. 24, no. 13, pp. 1397–1420, 2012.
19. S. Wang, Z. Liu, Z. Zheng, Q. Sun, and F. Yang, "Particle swarm optimization for energy-aware virtual machine placement optimization in virtualized data centers," in *2013 International Conference on Parallel and Distributed Systems*, pp. 102–109, IEEE, 2013.
20. D. Kumar and Z. Raza, "A pso based vm resource scheduling model for cloud computing," in *2015 IEEE International Conference on Computational Intelligence & Communication Technology*, pp. 213–219, IEEE, 2015.
21. F. Abdessamia, Y. Tai, W. Zhang, and M. Shafiq, "An improved particle swarm optimization for energy-efficiency virtual machine placement," *2017 International Conference on Cloud Computing Research and Innovation (ICCCRI)*, pp. 7–13, 2017.
22. A. Al-Dulaimy, W. Itani, R. Zantout, and A. Zekri, "Type-aware virtual machine management for energy efficient cloud data centers," *Sustainable Computing: Informatics and Systems*, vol. 19, pp. 185–203, 2018.
23. K. Ajmera and T. K. Tewari, "Greening the cloud through power-aware virtual machine allocation," in *2018 Eleventh International Conference on Contemporary Computing (IC3)*, pp. 1–6, IEEE, 2018.
24. B. Li, J. Li, J. Huai, T. Wo, Q. Li, and L. Zhong, "Enacloud: An energy-saving application live placement approach for cloud computing environments," in *2009 IEEE International Conference on Cloud Computing*, pp. 17–24, IEEE, 2009.
25. C. Dupont, T. Schulze, G. Giuliani, A. Somov, and F. Hermenier, "An energy aware framework for virtual machine placement in cloud federated data centres," in *2012 Third International Conference on Future Systems: Where Energy, Computing and Communication Meet (e-Energy)*, pp. 1–10, IEEE, 2012.
26. Y. Wu, M. Tang, and W. Fraser, "A simulated annealing algorithm for energy efficient virtual machine placement," in *2012 IEEE international conference on systems, man, and cybernetics (SMC)*, pp. 1245–1250, IEEE, 2012.
27. H. Goudarzi and M. Pedram, "Energy-efficient virtual machine replication and placement in a cloud computing system," in *2012 IEEE Fifth International Conference on Cloud Computing*, pp. 750–757, IEEE, 2012.
28. D. Huang, D. Yang, H. Zhang, and L. Wu, "Energy-aware virtual machine placement in data centers," in *2012 IEEE Global Communications Conference (GLOBECOM)*, pp. 3243–3249, IEEE, 2012.
29. K. Le, R. Bianchini, J. Zhang, Y. Jaluria, J. Meng, and T. D. Nguyen, "Reducing electricity cost through virtual machine placement in high performance computing clouds," in *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, p. 22, ACM, 2011.
30. Y. Li, W. Li, and C. Jiang, "A survey of virtual machine system: Current technology and future trends," in *2010 Third International Symposium on Electronic Commerce and Security*, pp. 332–336, IEEE, 2010.
31. B. Yuvaraj and K. Palanivel, "A survey of virtual machine placement algorithms in cloud computing environment," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 3, no. 1, 2015.
32. M. Nelson, B.-H. Lim, G. Hutchins, et al., "Fast transparent migration for virtual machines," in *USENIX Annual technical conference, general track*, pp. 391–394, 2005.
33. S. Akoush, R. Sohan, A. Rice, A. W. Moore, and A. Hopper, "Predicting the performance of virtual machine migration," in *2010 IEEE international symposium on modeling, analysis and simulation of computer and telecommunication systems*, pp. 37–46, IEEE, 2010.
34. E. Rashedi, H. Nezamabadi-Pour, and S. Saryazdi, "Bgsa: binary gravitational search algorithm," *Natural Computing*, vol. 9, no. 3, pp. 727–745, 2010.
35. X. Fan, W.-D. Weber, and L. A. Barroso, "Power provisioning for a warehouse-sized computer," in *ACM SIGARCH computer architecture news*, vol. 35, pp. 13–23, ACM, 2007.