Hybrid biogeography algorithm for reducing power consumption in cloud computing

Conference Paper · September 2017

DOI: 10.1109/ICACCI.2017.8125827

Sunita Bansal
Manipal University Jaipur

Jyoti Grover
Malaviya National Institute of Technology Jaipur

Hybrid Biogeography Algorithm for Reducing Power Consumption in Cloud Computing

Sunita Singhal
School of Computing and Information Technology
Manipal University Jaipur
Jaipur 303007 Rajasthan India
Email: sunita.singhal@jaipur.manipal.edu

Jyoti Grover
School of Computing and Information Technology
Manipal University Jaipur
Jaipur 303007 Rajasthan India
Email: jyoti.grover@jaipur.manipal.edu

***Abstract*—This paper purposes a hybrid biogeography algorithm that is the combination of biogeography and particle swarm optimization algorithm. A hybrid algorithm performs mutation, based on velocity instead of random mutation. The simulation results show that our proposed scheduling algorithm reduces the energy footprint of cloud data centers and improves a convergence rate of biogeography algorithm.**

***Index Terms*—Particle swarm optimization, biogeography based optimization, clusters, K-means.**

I. Introduction

Today, cloud computing has become an important and integral part of our life. It exists in one form or another in our digital life. Be it as a mail service or a video streaming service. With this increased dependence on cloud computing, there has arisen a need for developing big data centers that can service our needs. All these data centers are heterogeneous and quite different from each other. The only aspect through which they are similar is the huge power consumption. The data centers require a huge amount of energy to run and any approach which can reduce the consumption even marginally will have a great impact on the power bill and ultimately the profit.

The power consumption of the data centers is a key factor that shows us how efficiently that center is running. Thus, finding a technique that reduces the power consumption greatly increases the resource efficiency. The main aim of our work is to find a suitable framework that would allow maximum possible power saving. Our work will focus on machine learning techniques to learn the users tasks and utilize this knowledge to efficiently deploy resources accordingly.

The rest of the paper is organized as follows. Section II presents the related work. Section III describes our proposed algorithm named Hybrid Biogeography Algorithm. The results are presented and discussed in Section IV. Finally, in Section V we conclude the work and suggest future research directions.

II. Related Work

There have been mostly three ways to reduce the power consumption. Most of the earlier work on power reduction of the data centers revolved around actual control of hardware resources such as processors by controlling its voltage and frequency to save power. Initial work done by [1] revolves around controlling CPU voltage to reduce power consumption. In their technique, they proposed a real-time service model to complete real-time tasks using the minimum power. In general it is difficult to obtain. Even it can lead to missed deadlines and increased response time minimum power.

Later technique by [2] defines SLA as the basis for saving energy. The VMs are consolidated for energy efficiency and the unused hosts are kept in sleep mode for speedy restarting of the physical infrastructure. This technique is mostly based on architectural styles that dont use any of the machine learning algorithms.

The framework proposed by [3] in their paper of virtual machine customization and task mapping architecture for efficient allocation of cloud data center resources. In this, they have devised a way of learning from the history of the task load. They first collect the history of the cloud load for a predetermined time and then try to learn the VM configuration from that history. To learn they used K-means clustering algorithm which clustered similar type of tasks in the same cluster. From this clustering, they formulated an optimized VM configuration for each cluster. Then future tasks are mapped to appropriate clusters and appropriate VM configuration. Drawback of this technique is that it uses K-means for clustering which gave locally optimized clusters and thus poorer clustering than the globally optimized clusters. The technique that we are proposing will give globally optimized clusters which will give better clusters.

III. Hybrid Biogeography Algorithm

The primary algorithm used for clustering in [3] was K-means which generates a locally optimized solution and thus may not converge to a globally optimized solution. So, to overcome this we propose the hybrid algorithm that globally optimize the clustering formed by the K-means to get more accurate clustering and in return save more energy.

The machine learning algorithms such as Genetic algorithm [4], Teacher Learner algorithm [5], Gravitational Search algorithm [6], Biogeography based algorithm [7], [9], [10] (BBO) are all based on the natural phenomenon of finding the best solution by observing the best solution that is measured by the cost function of the problem. We are proposing a new

978-1-5090-6367-3/17/$31.00 ©2017 IEEE

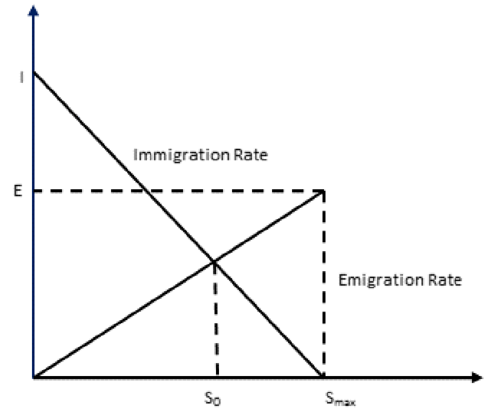

Fig. 1. Relationship between Species Count and Rates

BBO+PSO hybrid technique to lessen the time of convergence of BBO so that this technique can be used in real life cloud environment. The clusters formed from Hybrid BBO will be used for defining the VM configuration step in the framework proposed in [3] i.e. finding the average usage of each computing resources for a particular cluster and making that usage as the VM configuration for that cluster.

BBO algorithm mainly works on migration and mutation principals. Migration means habitat share features with each other. Mutation means sudden changes in habitat. Migration principal works on emigration and immigration rate. Emigration and immigration rate are affected by various factors such as the distance of an island to the nearest neighbor, a size of the island and Habitat Suitability Index (HSI). The relationship between immigration and emigration rate and species count is shown in figure 1. Where I refers the maximum immigration rate, E refers to the maximum emigration rate, S_0 is the equilibrium number of species and S_{max} is the maximum count.

Habitat is basically a solution and solution is a collection of Suitability Index Variables (SIV). The habitat which is high HSI have high emigration and low immigration rate. The habitat that has low HSI have a small number of species.

Mutation of the BBO is random but this phenomenon of pure random mutation does not really apply on natural mutations. Mutations mostly happen as a response to the external factors such as climate change or availability of food. This intelligent mutation is still random but it sees the external factors such as the individual that are surviving the most in a given condition and mutates itself to be in line with those individuals. Keeping this in mind we propose an intelligent mutation function derived from the Particle Swarm Optimization (PSO) algorithm. Mutation of BBO is computed using equation 1.

$$\begin{aligned} Habitat([i][SIV]) = W + Habitat[i][SIV] + \\ (C1 * R1 * (Pbest[SIV] - Habitat[SIV]) + \\ (C2 * R2 * (Gbest[SIV] - Habitat[SIV]))) \end{aligned} \tag{1}$$

Here $Gbest$ is the best value obtained by any habitat in island. $Pbest$ is best obtained by an habitat in an island. $R1$ and $R2$ are a random number between 0 and 1. W is immigration rate. $C1$ and $C2$ are learning factor and it is set as 2. The Pseudo code of Hybrid BBO is given below:

Algorithm 1 Pseudo code of Hybrid BBO

Set maximum migration rate (immigration and emigration);
Set maximum mutation rate and island size S_{max};
Initialize the island;
Compute G_{best} of island;
Compute immigration and emigration rate of each habitat;
Select habitat for emigration with emigration probability;
repeat
    Select SIV with immigration probability;
    Select SIV' with emigration probability;
    Replace SIV with SIV';
until $I < S_{max}$;
$I \leftarrow 0$
Compute G_{best};
if $G_{best} < G_{best'}$ **then**
    $G_{best'} \leftarrow G_{best}$;
end if
repeat
    Compute P_{best}
    Perform mutation equation 1
until $I < M$;
I=0;

Hybrid BBO algorithm first initializes immigration and emigration rate of the island after that it computes the global best, immigration and emigration probability as given in [7]. It selects habitat for immigration and emigration and using a roulette wheel. Then, it randomly replaces SIV of immigration and emigration habitat. After that it computes G_{best} and performs mutation on small number of times (M) using equation 1.

The flowchart figure 2 shows, how hybrid BBO algorithm works in cloud data center.

IV. Experimental Setup

For our experiment, we chose google cloud dataset [8] which contains trace logs of job submitted, executed and information about the requested resources such as the amount of RAM, CPU, Hard disk and the Priority of the task.

Two days trace data are used to train the model for learning about the VM configurations. After that, we tested the model with the task from the following days.

The main criteria that we tested was the power consumption of CPU, RAM, HARD DISK in the data center using Without Technique (Wit.Tec), BBO, hybrid BBO and K-means algorithms. The power consumption results are compared with full consumption resources for the whole duration of execution.

We have also measured convergence rate of the Basic BBO and Hybrid BBO technique. Here, the cost value is the total

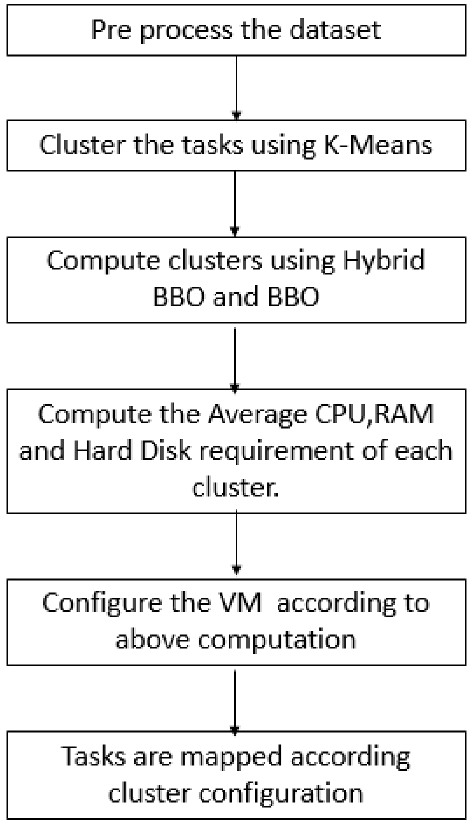

Fig. 2. Flow Chart

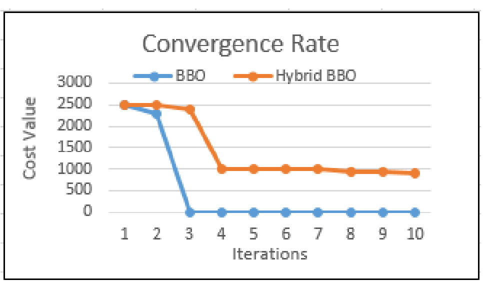

Fig. 3. Convergence Rate

SSE (Sum of squared error) of the clustering in each iteration. From the Figure 3 we can see that our hybrid BBO technique converges quickly to the optimal solution than BBO. This property is very important because the model building gets faster which directly relates to how quickly the algorithm will start producing results.

From the Figures 4 to 6, it can be seen that our hybrid BBO technique out performs every other technique at all instance of time. The readings for each unit that is CPU, RAM, Hard Disk have been normalized in the range of [0,1]. The reduced power consumption can lead to huge savings in power.

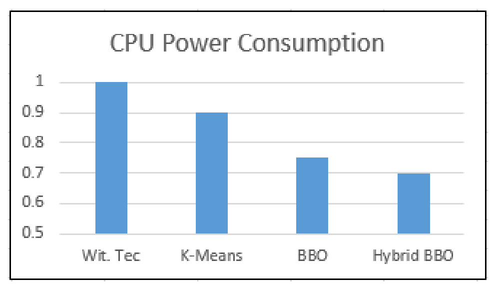

Fig. 4. CPU Power Consumption

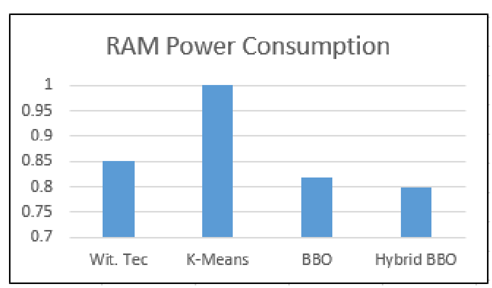

Fig. 5. RAM Power Consumption

V. Conclusion

We have improved the already given framework for lowering the power consumption by incorporating nature-inspired

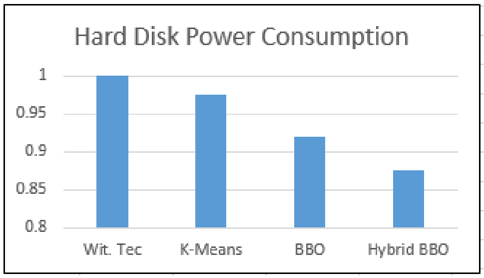

Fig. 6. Hard Disk Power Consumption

machine learning technique Hybrid BBO to allow globally optimized clustering of tasks.

From the results, we can conclude that the convergence rate is greatly improved and the optimal clusters are produced in the smaller number of steps than the BBO algorithm.

Also, the Power consumption is greatly reduced by using the BBO and hybrid BBO approach rather than using the K-means approach for clustering the task.

In the future work, we can analyze the effect of assigning the task to predefined VMs and how this affects their average response time and if the frequency of tasks being killed is increased.

References

[1] K. H. Kim, A. B Beloglazov and R. Buyya, *Power aware provisioning of virtual machines for real-time cloud services*, Concurrency and Computation: Practice and Experience, Addison-Wesley, pp. 1-7, 2011.

[2] J. S. Liao, C. C. Chang, Y. L. Hsu and X. W. Zhang, *Energy-efficient resource provisioning with SLA consideration on cloud computing*, 41st International Conference on Parallel Processing Workshops, 2011.

[3] S. Fotuh, R. N. Calheiros, J. Chan, A. Vahid and R. Buyya, *Virtual machine customization and mapping architecture for efficient allocation of cloud data center resources*, The Computer Journal, vol. 59, issue 2, pp. 208-224, Feb. 2016.

[4] K. F. Man, K. S. Tang and S. Kwong, *Genetics algorithms: concepts and applications*, IEEE Transactions on Industrial Electronics. vol. 43, issue 5, pp. 519-534, Oct. 1996.

[5] V. Rao and V. Patel, *An improved teaching learning based optimization algorithm for solving unconstrained optimization problems*, Scientia Iranica, vol. 20, issue 3, pp. 710-720, June, 2013.

[6] E. Rashedi and H. N. Pour, *Gravitational search algorithm* Information Sciences, vol. 179, issue 13, pp. 2232-2248, June 2009.

[7] D. Simon, *Biogeography-based optimization*, IEEE Transactions on Evolutionary Computation, vol. 12, issue 6, pp. 702-713, Dec. 2008.

[8] C. Reiss and J. Wilkes, *Google cluster-usage traces: format and schema, Google Inc., version 2, 2013.*

[9] H. Ma and D. Simon, *Blended biogeography-based optimization for constrained optimization*, Engineering Applications of Artificial Intelligence, vol. 24, no. 6, pp. 517-525, April 2010.

[10] D. Du, D. Simon and M. Ergezer, *Biogeography-based optimization combined with evolutionary Strategy and immigration Refusal*, Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, pp.997-1002, October 2009.

[11] Z. Pooranian, M. Shojafar, J. H. Abawajy and M. Singhal, *Gloa: a new job scheduling algorithm for grid computing* International Journal of Interactive Multimedia and Artificial Intelligence, vol. 2, no. 1, pp. 59-64, 2013.