JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3021-3030
https://publishoa.com
ISSN: 1309-3452

A Hybrid Metaheuristic approach on Virtual Machine Allocation using Intelligent Backtracking and Simulated Annealing in cloud computing

T. Lavanya Suja[1] **B. Booba[2]**

[1]Research Scholar, Dept. of Computer Science, lavanyasujat@gmail.com

[2]Professor, Dept. of Computer Science, boobarajashekar@gmail.com

VISTAS

Chennai, India

Received 2022 April 02**; Revised** 2022 May 20**; Accepted** 2022 June 18**.**

Abstract

Background Cloud and its presence are ubiquitous more than ever after pandemic. Various domain has migrated from on-premises applications to cloud. Due to the increase in cloud adoption, infrastructure management has gained importance multi fold. Infrastructure refers to the servers in the datacenters. The servers are called as Physical Machines (PM) which is composed of logical cores called Virtual Machines (VM). Ultimately jobs are submitted to VMs so VM allocation one of the important tasks in job scheduling. Number of algorithms have poured into cloud market borrowed from various discipline such as nature, mathematics, physics etc. for VM allocation. The proposes approach of algorithms are of hybrid metaheuristic taken from mathematics, physics namely Intelligent Backtracking and Simulated Annealing respectively. Experimental studies show the parameters of concern such as execution time, Energy consumption and cost incurred are optimal when compared to the existing statistical approaches. There is a 48% decrease in the energy consumption and a 90% decline in cost of using the approach. **Objectives**: The objective of the research work is to bring down the Execution Time, Energy Consumption and cost incurred thereby obtaining an optimal solution for the VM allocation problem in cloud computing.

Methods: The proposed approaches namely Intelligent Backtracking for VM Placement and Simulated Annealing for VM migration are implemented with the parameters to study and analyze the working and feasibility of it. From the related works studied, the parameters of influential importance are Execution time, Energy consumption and cost incurred. Along with it, in the proposed approach Number of Hosts shutdown is also taken into consideration as it has an impact on time and energy parameters. These three parameters affect the cost parameter in a positive way. **Conclusions**: The proposed algorithm decreases the Mean and standard deviation Execution Time by 21.5% & 33.5% respectively. The proposed algorithm decreases the Number of Hosts shut down by 19%. As a result, the proposed algorithm's energy consumption is 68.30kWh and improves it by 48%. The research work concludes that hybrid metaheuristic approach on Virtual Machine Allocation using Intelligent Backtracking and Simulated annealing brings optimal results than the statistical, and simple metaheuristic algorithms proposed before. Future work will bring machine learning algorithms for load prediction and devising a new ML based approach for VM allocation in cloud computing.

Keywords: Metaheuristic approach, VM Allocation Algorithms, Intelligent Backtracking, Simulated Annealing, Cloud Computing

1. Introduction

Cloud computing is an extension of Service Oriented Architecture (SOA) and getting services through internet in pay per use basis. Cloud services have started from three basic types of viz. Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS) to Anything as a Service (XaaS). Among them Infrastructure as a Service stand as a foundation of other services. Infrastructure referred here are highly configured servers, humongous storage, modern databases and etc. In order to maintain the infrastructure various approaches for effective management is needed

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3021-3030
https://publishoa.com
ISSN: 1309-3452

as cloud jobs are submitted to them using any kind of services mentioned above. Right from the adoption of cloud and till today algorithms from different discipline are poured in and their implementation is tested by various simulation tools like CloudSim, Cloud Analyst, Fog Sim, etc. Among the simulation tools CloudSim is the pioneer one and have evolved a lot to accommodate IoT, edge and fog computing.

Adding value to the algorithms for Infrastructure management, this research paper proposes two hybrid metaheuristic approaches namely Intelligent Backtracking and simulated Annealing borrowed from Mathematics and Physics. The rest of the paper is organized like this. Section 2 a detailed and comprehensive literature review about the algorithms for infrastructure management in cloud market. Section 3 talks about the Proposed method by elaborating the system architecture, mathematical formulation, and pseudocode of the proposed algorithms. Section 4 explains the experimental setup and the graphs drawn from them. Section 5 brings the results and discussion, finally Section 6 gives the conclusion and future work followed by the references cited in the research paper.

2. Literature Review

The authors of [1] emphasize the importance of Load Balancing in Cloud Computing jobs scheduling. Their research problem is closely related to one of the problems proposed here as finding whether a PM is overutilized or underutilized. In their work, the authors have come up with an Improved Throttle Algorithm (ITA) which reduces the jobs processing time, limits the idling of cloud resources. They also claim this results in reduced cost of the datacenters which is an important factor. In Cognito with this research problem, in this proposed research approach of Intelligent Backtracking-Simulated Annealing, cost factor is introduced.

Another optimization algorithm [2] namely Whale Optimization is employed for workflow scheduling where load balancing is involved. They claim that the experimental results give more throughput and decrease in energy consumption. The energy consumption used in this paper [2] is one of the factors considered in our proposed approach.

In their work [3] , the authors proposed a combination of three approached borrowed from 3 different discipline, Non-dominated Sorting Genetic Algorithm (NSGAII), Chemical Reaction Optimization (CRO), Binary Particle Swarm Optimization (NBPSO), for finding an optimal solution for their Integer Linear Programming (ILP) model. This meta-heuristic approach, they claim to employ as it is very time consuming to arrive at a solution for an ILP. This is one of the reason this research work is included in this survey as the proposed approach is also of meta-heuristic nature. Another common ground here is the borrowing of approached from different discipline. They have borrowed from Genetic, Chemistry and Nature inspired approaches. Similarly in this research work successful algorithms like Backtracking from Mathematics, popular technique like Simulated Annealing from Physics is borrowed to bring an optimal solution for the research problem.

In Software Defined Data Centers (SDDC) of Connected and Automatic Vehicle (CAV) the energy efficiency and SLA are key factors to be followed at any cost. As new approaches are always welcome in these applications, the authors of [4] came up with EVCT which works on “Maximum Flow and Minimum cut Theory” on its directed graph created for VMs. The selection of VMs and VM cluster replacement achieved energy efficiency, high QoS, better scalability during variable workload which further decreased cost of energy utilization. They used a real time data like us, and inclusion of energy and cost factor are the reasons why this work is taken for survey. The fact that new algorithms are always welcomed in cloud IaaS domain is again reiterated by this research work too.

Another new algorithm based on Stochastic technique especially for edge clouds and where the load is variable most of the time. The authors of [5] mainly try to increase the revenue on resource utilization and decrease in time complexity. This work gave us an idea of making this work a multi-objective solution by incorporating revenue as a factor so as to maximize it while already other factors like tile, energy, cost are for minimization.

Related to our previously proposed approach of Hill Climbing, a recent work [6] employed Gradient Based optimization for Task Scheduling in cloud. The authors of [6] claims that their GBO avoid Local Minima and achieves in better makespan, great accuracy and improved speed in convergence. The idea is to emphasize the contribution of metaheuristic algorithms for Task Scheduling and Load Balancing in cloud market.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3021-3030
https://publishoa.com
ISSN: 1309-3452

The authors of [7] propose a novel thermal aware holistic model which considers the impact of node failure in cooling costs, energy needed for it. Their experimental results gave improved percentage on cutting down the cooling costs, minimizing the inlet temperature and improves the maximum air temperature. The results reiterate the fact why energy consumption is added as a factor of comparison in the proposed approach. Since clous uses more cooling for the datacenters, efficient energy usage is the need of the hour and will be always a concern for green computing.

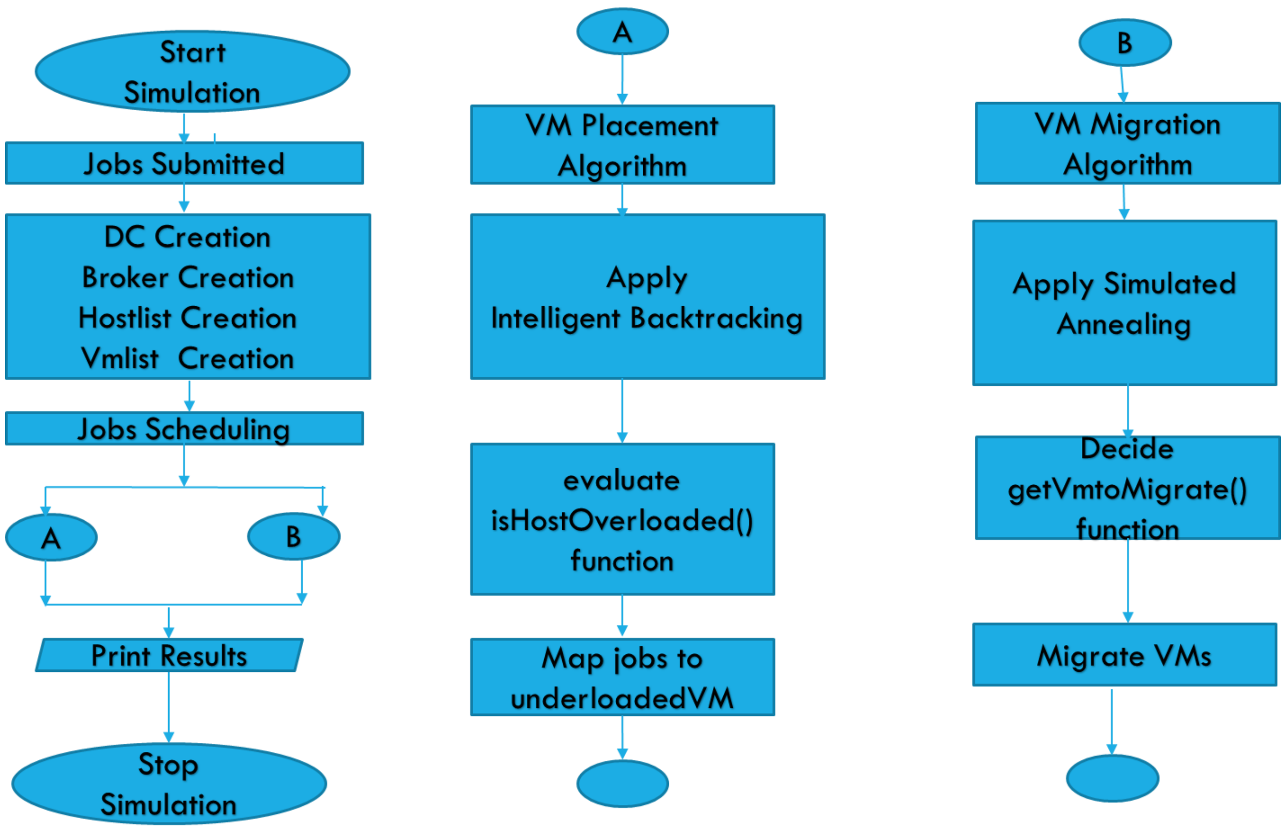
In a detailed comprehensive review [8] about meta-heuristic techniques for energy related problems, the authors presented the types of the meta-heuristic techniques and given the mathematical formulation too. They have strongly documented that meta-heuristic algorithm are borrowed from different domains such as Biology, Physics, and Ethology. Simulated Annealing is one among them from Physics they document in their review. Also, they divide the meta-heuristic algorithms into combinations of meta-heuristics with meta-heuristics and Hybrid meta-heuristics with other techniques. Another level of detail they record was that these types of branches into high level and low-level meta-heuristics algorithms. High level meta-heuristic algorithms don’t change the code of the borrowed algorithm in contrary to the low level meta-heuristic algorithms which change the code in the approach. After reading this review article [8] it is concluded that the proposed research problem’s solution is a hybrid low level meta-heuristic with meta-heuristic approach for the VM Allocation problem.

The authors of [9] in their energy efficient VM allocation approach come up with a look ahead module based on Holt Winters model for predicting the load for servers. The error rate for this model is 8.85% and it is better than previous models like ARIMA, Support Vector Regression and Non-linear Regression. Also because of this load prediction no VM migrations are happening between servers. Moreover, new servers are not turned on if there is space in the current sever hence decreasing the energy consumption and decreasing the number of hosts shutdown too. These experimental results resemble the results of the proposed one too. Hence the literature review of this work is appropriate and adds value to our results too. The common ground of simulation is done using Cloudsim which is the one used in the proposed research work too. The authors compare their work to LR_MMT and we compare our results starting with IQR_MMT.

3. Methods

After the detailed literature review, proposed approaches namely Intelligent Backtracking for VM Placement and Simulated Annealing for VM migration are implemented with the parameters to study and analyze the working and feasibility of it. From the related works studied above the parameters of influential importance are Execution time, Energy consumption and cost incurred. Along with it, in the proposed approach Number of Hosts shutdown is also taken into consideration as it has an impact on time and energy parameters. These three parameters affect the cost parameter in a positive way. The below flowchart gives the process flow of the proposed approach.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3021-3030
https://publishoa.com
ISSN: 1309-3452



After the simulation is started, jobs are submitted and default activities like DC, Broker, Hostlist and VMlist are created. Then Job Scheduling takes place. The area of concern here is VM allocation which consists of two sub tasks namely VM Placement for the jobs and VM migration if the host is overloaded. These two subtasks are carried out as and when needed and then the results are printed. This approach is a centralized one when compared to decentralized approach of the base work [12]

Mathematical formulation of the Algorithms

1. $Zmax = 1 - f(x) * f(y)$ *Where* $0.2 \le x, \le 0.4$, ***f(x) is the Safety Parameter and f(y) is the utilization Threshold***
2. $X = \begin{cases} \sum_1^n xi < Zmax, \ UnderUtilization \\ \sum_1^n xi > Zmax, \ OverUtilization \end{cases}$
3. $m = \frac{c_i}{C}$, *where* c_i = *current Mips of the ith VM and C =*

   Total Mips(Million instructions per second)of all VMs in the Host,

If m> Maxmetric2 then Maxmetric2 = m

The objective function is Zmax which is calculated as 1-f(x) * f(y), where f(x) is the safety parameter and f(y) is the utilization threshold which varies from 0.2 to 0.4. Then the sum of Xis is calculated to check for overutilization and underutilization. These calculations are done for VM allocation. The next variable m is the metric for measuring how much CPU utilization has been done for a VM. If it is greater than maximum number assigned to maxmetric then it becomes maxmetric's value and add this vm to the vmlist to be migrated, otherwise continue the iterations for vmlist.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3021-3030
https://publishoa.com
ISSN: 1309-3452

Pseudocode of Intelligent Backtracking for VM Placement

```
boolean isHostOverUtilized(PowerHost Host)
      initialize utilization=0.
     calculate upperThreshold;
      addHistoryEntry(host, upperThreshold);
      For Each VM in VMList of Host do
      Update totalRequestedMips & totalRequestedBw & totalRequestedRam;
     if(i>=countofvms/2) { //intelligent Backtracking
     if((totalRequestedBw> host.getBw())||(totalRequestedRam>host.getRam()))
     break;                  //backtrack }
      End For
      double utilization = totalRequestedMips / host.getTotalMips();
      if ((utilization > upperThreshold) ) return true;
      else      return false;
```

Pseudocode of Simulated Annealing for VM Migration

```
    Vm getVmToMigrate(PowerHost host)
    {
    Initialize vmList,MinMetric and MaxMetic
    Inside For loop
    Check min Metric > sizeofvm()
    Check maxMetric <  CpuUtilization(vm)
    If yes skip the iteration
    Otherwise minMetric = sizeofvm() and maxMetric= CpuUtilization(vm)
    Add vm to vmList
```

4. Experimental Setup

The experimental set up in Cloudsim3.0.3 has the following configuration. The configuration of the PM, Operating System is Linux with Architecture of x86 and virtual machine monitor as Xen. The DC models are HP ProLiant ML110 G4 servers and HP ProLiant ML110 G5 servers as per the system configuration given in [15]. The dataset is taken from Planet lab on the date “20110420” [10] is fed for all the existing, newly, and previously proposed algorithms to measure the performance characteristics. First the newly proposed Intelligent Backtracking and Simulated Annealing algorithm with metaheuristic approach is executed and the output is compared against existing statistical approaches and previously proposed approaches namely Backtracking with Minimum Migration Time, Backtracking with Hill Climbing and Backtracking with Simulated Annealing, is given below.

Table 1 Existing (1) and Proposed (2-6) Algorithms for comparison

S.No.	Name of the Algorithm
1	Inter quartile range & Min. Mig. Time [11]
2	Backtracking & Min. Mig. Time [13]
3	Backtracking & Hill Climbing [14]
4	Backtracking & Simulated Annealing
5	Intelligent Backtracking & Hill Climbing
6	Intelligent Backtracking & Simulated Annealing

Following are the illustration of influential parameters namely execution time- mean and standard deviation, number of hosts shutdown, energy consumption and cost incurred by existing and proposed algorithms.

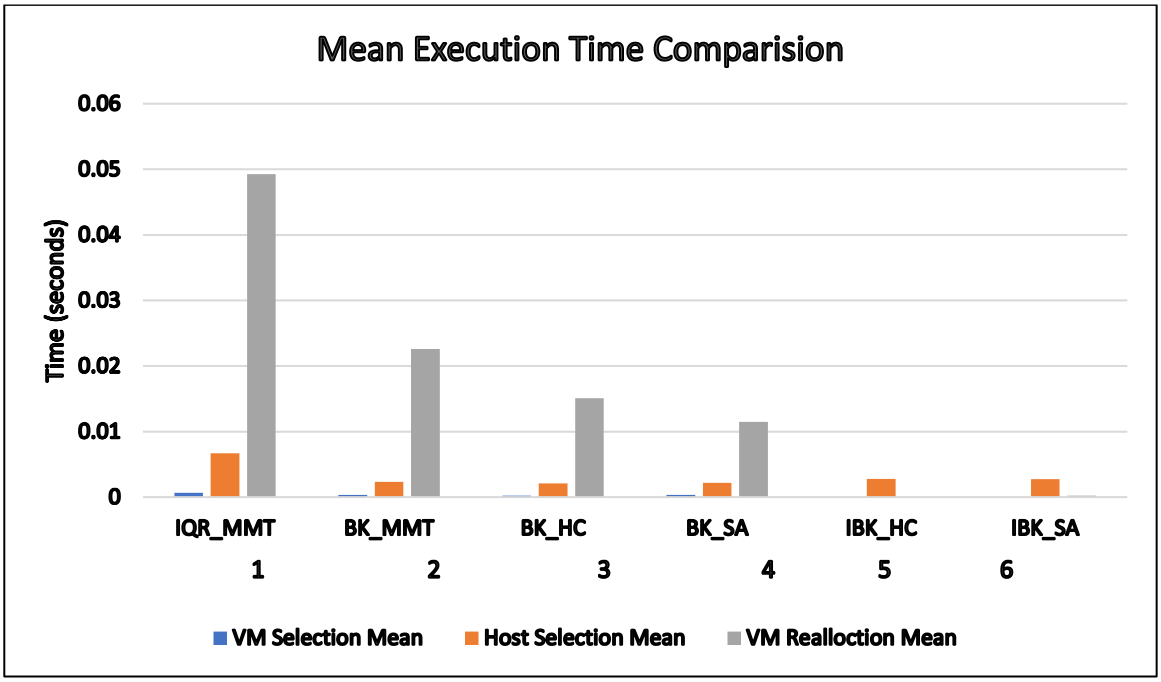


Fig. 1 Comparison of Mean Execution Time in Existing and Proposed Algorithms

In the fig. 1 above the average execution time in seconds are plotted as the bar chart where 1 refers to the first proposed Backtracking and Simple Hill climbing approach, 2 stands for Backtracking and Simulated Annealing, 3 stands for Intelligent Backtracking and Hill Climbing, 4 refers to the recently proposed Intelligent Backtracking and Simulated annealing approach. The bars in the graph clearly illustrates the fact that the mean execution time of VM Selection, Host Selection and VM Reallocation is minimal in approach 4 than the previous ones.

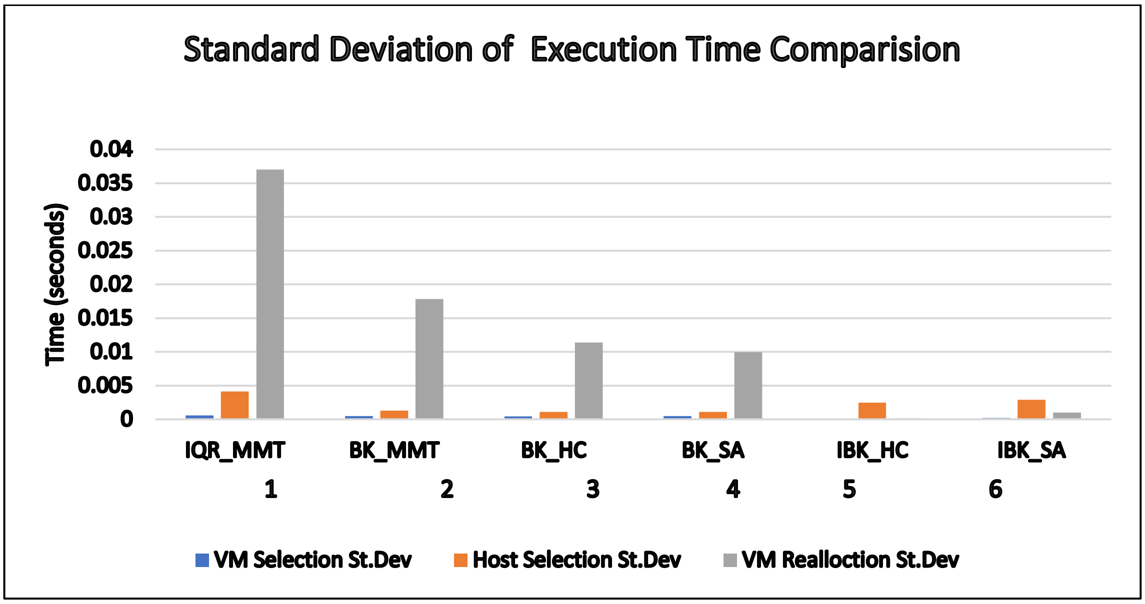


Fig. 2 Comparison of Standard Deviation of Execution Time in Existing and Proposed Algorithms

In the fig. 2 above the average execution time in seconds are plotted as the bar chart where 1 refers to the first proposed Backtracking and Simple Hill climbing approach, 2 stands for Backtracking and Simulated Annealing, 3 stands for Intelligent Backtracking and Hill Climbing, 4 refers to the recently proposed Intelligent Backtracking and Simulated annealing approach. The bars in the graph clearly illustrates the fact that the execution time's standard deviation of VM Selection, Host Selection and VM Reallocation is less in approach 4 than the previous ones.

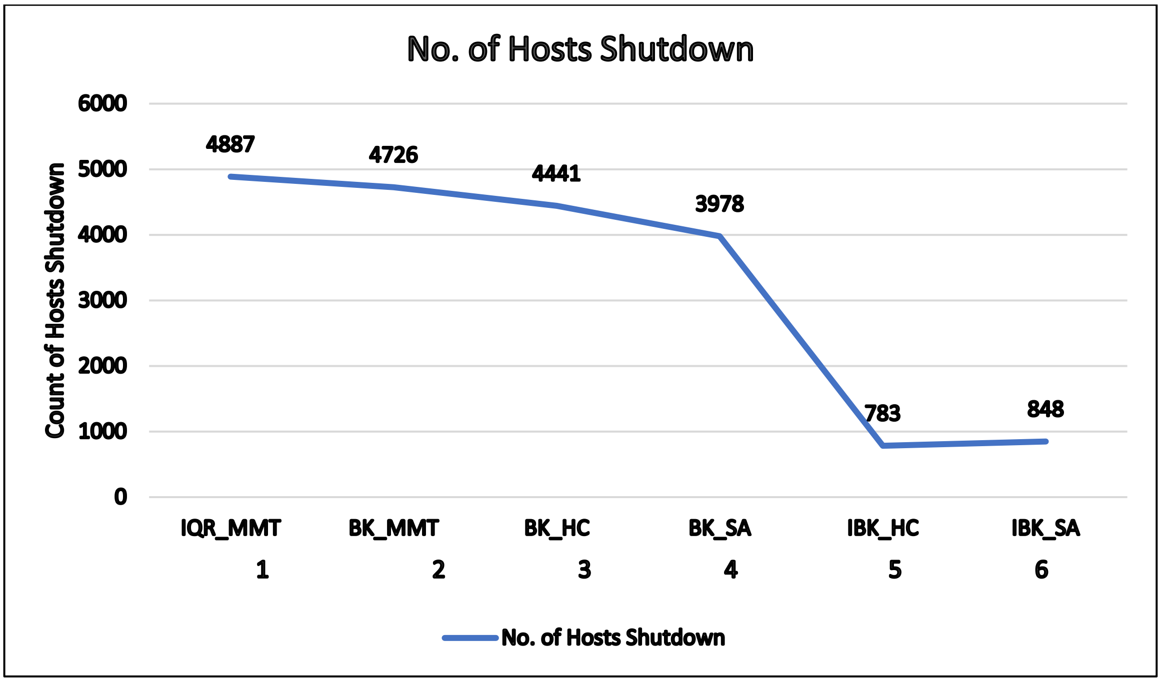


Fig. 3 Comparison of Number of Hosts Shutdown in Existing and Proposed Algorithms

In the fig. 3 above the number of hosts that are shutdown during the execution are plotted as the bar chart where 1 refers to the first proposed Backtracking and Simple Hill climbing approach, 2 stands for Backtracking and Simulated Annealing, 3 stands for Intelligent Backtracking and Hill Climbing, 4 refers to the recently proposed Intelligent

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 3021-3030
https://publishoa.com
ISSN: 1309-3452

Backtracking and Simulated annealing approach. The bar in the graph clearly illustrates the fact that count of hosts shutdown comes to an optimal number in approach 4 than the previous ones.

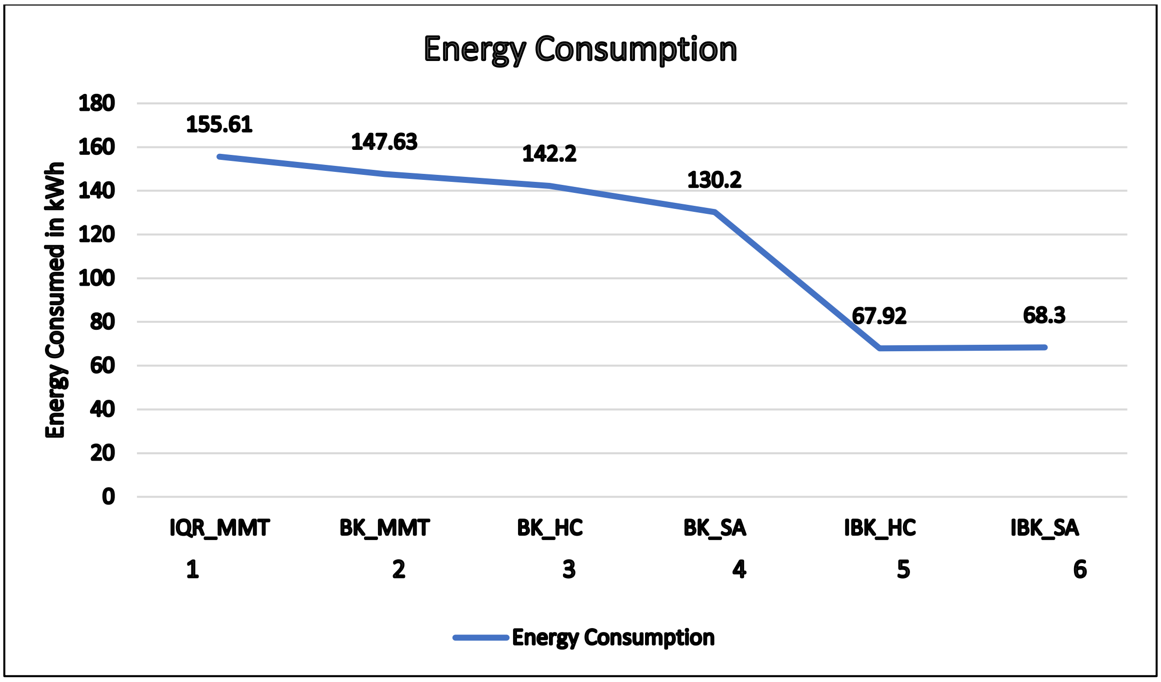


Fig. 4 Comparison of energy consumption in Existing and Proposed Algorithms

In the fig. 4 above the energy consumption in kilo watts per hour (kWh) are plotted as the line graph where 1 refers to the first proposed Backtracking and Simple Hill climbing approach, 2 stands for Backtracking and Simulated Annealing, 3 stands for Intelligent Backtracking and Hill Climbing, 4 refers to the recently proposed Intelligent Backtracking and Simulated annealing approach. The line in the graph clearly illustrates the fact that the energy consumed by the algorithms is minimal in approach 4 than the previous ones. Hence approach 4 in energy efficient.

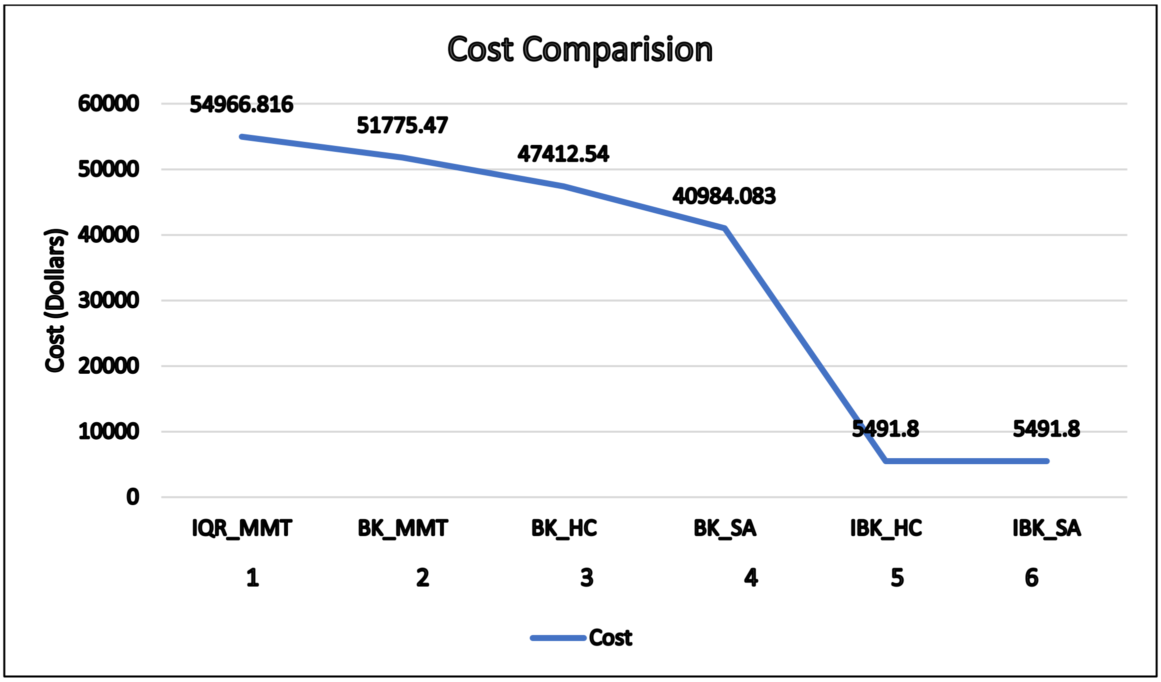


Fig. 5 Comparison of cost incurred in Existing and Proposed Algorithms

In the fig. 5 above the cost of usage of algorithms are plotted as the line graph where 1 refers to the first proposed Backtracking and Simple Hill climbing approach, 2 stands for Backtracking and Simulated Annealing, 3 stands for Intelligent Backtracking and Hill Climbing, 4 refers to the recently proposed Intelligent Backtracking and Simulated annealing approach. The sharp slope in the graph clearly illustrates the fact that the cost incurred by the algorithms is minimal in approach 4 than the previous ones. Hence approach 4 in cost efficient.

5. Results and Discussion

Following the experimental setup for various algorithms listed in Table 1, the observations of the parameters are plotted in the graphs from Fig. 1 through 5. After measuring the parameters of concern viz. Execution time, Number of Hosts Shutdown, Energy Consumption, and cost for the existing and proposed approach, it is evident that execution time, number of hosts shutdown are brought down considerably by 19% and as a good side effect of it, and because of the correlation, energy consumption and cost are brought down by 48% and 90% respectively.

The results are tabulated in the Table 2 below. This is a positive sign for energy aware solution needed nowadays to reduce the carbon footprints and a way for green computing. In this era of cloud computing, there is a concern that huge datacenters built in important cities around the globe contributes to global warming. Moving to unconventional and renewable energy resources is one side of a solution. Besides reducing the energy consumption in the datacenters is other side of the solution and it advocates to be a smart solution too.

Table 2. Parameters of concern in existing and Proposed Algorithm

Name of the Proposed Algorithm ------------------------- **Parameters (in %)**	**Intelligent Backtracking & Simulated Annealing**
Number of Times Host Shutdown	**19.09% (Decreased)**
Mean Execution Time	**2.25% (Decreased)**
Standard Deviation of Execution Time	**8.87% (Decreased)**
Energy Consumption	**48% decrease**
Cost of Utilization	**90% decrease**

6. Conclusion

The proposed algorithm decreases the Mean and standard deviation Execution Time by 21.5% & 33.5% respectively. The proposed algorithm decreases the Number of Hosts shut down by 19%. As a result, the proposed algorithm's energy consumption is 68.30kWh and improves it by 48%. The research work concludes that hybrid metaheuristic approach on Virtual Machine Allocation using Intelligent Backtracking and Simulated annealing brings optimal results than the statistical, and simple metaheuristic algorithms proposed before. Future work will bring machine learning algorithms for load prediction and devising a new ML based approach for VM allocation in cloud computing.

References

[1]. Le, H. N., & Tran, H. C. Ita: The Improved Throttled Algorithm Of Load Balancing On Cloud Computing.
[2]. Sobhanayak, S., Mendes, I. K. A., & Jaiswal, K. (2022). Bi-objective Task Scheduling in Cloud Data Center Using Whale Optimization Algorithm. In Advances in Data Computing, Communication and Security (pp. 347-360). Springer, Singapore.

[3]. Awad, M., Kara, N., & Edstrom, C. (2022). SLO-aware dynamic self-adaptation of resources. Future Generation Computer Systems.
[4]. Zhou, Z., Shojafar, M., Li, R., & Tafazolli, R. (2022). EVCT: An efficient VM deployment algorithm for a software-defined data center in a connected and autonomous vehicle environment. IEEE Transactions on Green Communications and Networking.
[5]. Wei, W., Wang, Q., Yang, W., & Mu, Y. (2022). Efficient stochastic scheduling for highly complex resource placement in edge clouds. Journal of Network and Computer Applications, 103365.
[6]. Huang, X., Lin, Y., Zhang, Z., Guo, X., & Su, S. (2022). A gradient-based optimization approach for task scheduling problem in cloud computing. Cluster Computing, 1-17.
[7]. Li, J., Deng, Y., Zhou, Y., Zhang, Z., Min, G., & Qin, X. (2022). Towards Thermal-Aware Workload Distribution in Cloud Data Centers Based on Failure Models. IEEE Transactions on Computers.
[8]. Bouaouda, A., & Sayouti, Y. (2022). Hybrid Meta-Heuristic Algorithms for Optimal Sizing of Hybrid Renewable Energy System: A Review of the State-of-the-Art. Archives of Computational Methods in Engineering, 1-35.
[9]. İlksen, Ç., & Altılar, D. T. (2022). Look-ahead energy efficient VM allocation approach for data centers. Journal of Cloud Computing, 11(1).
[10]. beloglazov/planetlab-workload-traces. (n.d.). GitHub. Retrieved November 01, 2021, from https://github.com/beloglazov/planetlab-workload-traces
[11]. Beloglazovy, A., & Buyya, R. (2011). Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. Concurr. Comput. Pract. Experience, 24(13), 1-24.
[12]. Wang X, Liu X, Fan L, Jia X. A decentralized virtual machine migration approach of data centers for cloud computing. Mathematical Problems in Engineering. 2013 Jun; 2013:1–10.
[13]. Lavanya Suja, T., Booba, B., Meta Heuristic Backtracking algorithm for Virtual Machine Placement in Cloud Computing Migration, Manuscript submitted for publication.
[14]. Lavanya Suja, T.,Booba, B., A Hybrid Approach of Backtracking and Hill Climbing algorithms for Virtual Machine Allocation in Cloud Computing, Manuscript in Press
[15]. SPECpower_ssj2008. (n.d.). Http://Www.Spec.Org. Retrieved November 01, 2021, from http://www.spec.org/power_ssj2008/results/res2011q1/power_ssj2008-20110127-00342.html
[16]. What Does Standard Deviation Show us About Our Data? (n.d.). Http://Science.Halleyhosting.Com. Retrieved November 01, 2021, from http://science.halleyhosting.com/sci/soph/inquiry/standdev2.htm
[17]. Learn About Skewness. (2021, April 30). Investopedia. https://www.investopedia.com/terms/s/skewness.asp