Sustainable Energy Technologies and Assessments 45 (2021) 101087

Contents lists available at ScienceDirect

Sustainable Energy Technologies and Assessments

journal homepage: www.elsevier.com/locate/seta

An approach towards development of new linear regression prediction model for reduced energy consumption and SLA violation in the domain of green cloud computing

Nirmal Kr. Biswas [a], Sourav Banerjee [b,*], Utpal Biswas [c], Uttam Ghosh [d]

[a] Department of Computer Science & Engineering, Global Institute of Management & Technology, Krishnagar, Nadia, W.B, India
[b] Department of Computer Science & Engineering, Kalyani Government Engineering College, Kalyani, Nadia, W.B, India
[c] Department of Computer Science & Engineering, University of Kalyani, Kalyani, Nadia, W.B, India
[d] Department of Electrical Engineering and Computer Science, Vanderbilt University Nashville, TN, USA

ARTICLE INFO

Keywords:
Energy consumption (EC)
Physical Machine (PM)
SLA violation (SLAV)
Smart City
Virtual Machine (VM)

ABSTRACT

With the increase of mega-cities, the demand for Smart Cities is overgrowing. The mega-cities can be smarter through the Cloud of Things (CoT). Efficient energy consumption in Smart Cities has a massive impact on the environment. But, computational power is increasing rapidly in the cloud computing environment. Enormous energy consumption (EC) and Service Level Agreement violation (SLAV) becomes a key concern. The Virtual Machine (VM) consolidation approach can significantly reduce EC, SLA violation (SLAV), and increase resource utilization. However, dynamic VM consolidation may produce performance degradation of Physical Machines (PMs) and SLAV. Therefore, it becomes essential to find a trade-off between EC and SLAV. Herein, a novel New Linear Regression(NLR) prediction model, host overload/underload, and VM placement policy have been proposed to reduce EC and SLAV. The NLR model's primary intention is to take that the model goes through a straight line and a mean point. Future CPU utilization is predicted based on the proposed NLR model. Evaluation of proposed algorithms has been accomplished by extending CloudSim Simulator. The experiment shows that proposed algorithms reduced EC and SLAV in cloud data centers and can be used to construct a smart and sustainable environment for Smart Cities.

1. Introduction

Energy consumption and its impact on the environment in cloud computing [1,2] have become a challenging issue to researchers. The data centers consume a massive amount of energy to provide the required services to cloud users. There are millions of servers in a cloud computing environment, which are consuming high energy. It causes a huge amount of carbon dioxide emissions to the environment. In 2016, the amount of electricity consumed by cloud data centers is larger than the total electricity consumed in Britain - 416.2 terawatt-hours [3]. The EC of cloud data centers will cause 3.2% of total carbon emission by 2025, which will be one-fifth of global electricity [4]. It can reach 14% of the total carbon emission by 2040 [4]. So, cloud data centers become modern frontier warfare toward environmental change.

An effective method must be adopted to increase the utilization of resources with reduced EC in cloud data centers. It can propagate a decisive impact on the environment. With the increasing population, the quality of living and the number of mega-cities are also increasing. These mega-cities' living qualities can be increases by using various technologies and computation approaches in a smart way. These technologies can be integrated into a Cloud of Things (CoT) [5] to build smart cities. In CoT, it is possible to interconnect too many aspects of smart cities into the IoT, such as hospitals, homes, stores, automobiles, etc. [6], and integrated with cloud computing systems to improve the smart city's services and sustainability. Efficient energy consumption in CoT can significantly reduce carbon emissions to the environment.

Virtual Machine (VM) consolidation is a process that can improve the utilization of resources with reduced EC in cloud data centers. VM consolidation consists of four different steps. In the first step, all overloaded physical machines (PMs) have been identified because these overloaded PMs have the potential for performance degradation and providing low Quality of Service (QoS) [7] to the cloud users. Some of

* Corresponding author.
E-mail address: mr.sourav.banerjee@ieee.org (S. Banerjee).

https://doi.org/10.1016/j.seta.2021.101087
Received 5 August 2020; Received in revised form 19 January 2021; Accepted 1 February 2021
Available online 24 February 2021

the VMs of overloaded PMs need to migrate into lesser loaded PMs [8] to improve the QoS. In the second step, all the underloaded PMs are detected to migrate all VMs of underloaded PMs into other PMs and put them into a sleep state. It can preserve unnecessary EC in inactive PMs. The third step is to determine suitable PMs where all the VMs are to migrate from the overloaded and underloaded PMs. The final step is to select the best candidate VM for the migration from the overloaded PM.

Cloud service providers do an agreement with cloud users to provide reliable QoS, called Service Level Agreement (SLA). But, dynamic VM consolidation may result in performance degradation of PMs and SLA violation [9]. So, it is very essential to find a balance in energy and performance i.e with minimum EC maximum meeting of SLA.

Live VM migration [10,11] is an approach to reduce EC and SLAV. To accomplish these objectives, it is very much essential to anticipate the future PM state precisely. The correlation between the EC of PMs and their CPU utilization is linear. So herein, a new prediction model has been proposed to detect whether PM is overloaded or underloaded. The proposed New Linear Regression (NLR) prediction model will predict the future CPU utilization.The key intention of the NLR prediction model is to reduce EC and SLAV. The novelties of this study are as follows:

- Propose an NLR prediction model for anticipating future CPU utilization.
- A Hosts Overload Detection (isHOD) algorithm using the NLR prediction model has been proposed to reduce EC and SLAV.
- A Hosts Underload Detection (isHUD) algorithm using the NLR prediction model has been proposed to reduce EC and SLAV. A lower utilization threshold using the interquartile range (IQR) also proposed.
- A novel Modified Power-Aware Best Fit Decreasing (MPABFD) algorithm using the NLR prediction model has been proposed to perform VM placement.

The remainder of the paper composed as follows: Section 2 gives a short concept about the literature survey and state of art. The proposed NLR prediction model is shown in Section 3. The details analysis of the NLR prediction model is shown in Section 4. Finally, Section 5 details the conclusions of the work and future exploration area.

2. Literature survey

In this section, related literature on VM consolidation to reduced power consumption with optimized SLA violation is reviewed. VM consolidation is consider as an optimization problem because of it NP-hard characteristics [12–14].

Beloglazov et al. [10,15,16] proposed Linear Regression (LR) prediction model to predict future CPU utilization of PMs. Using last few observation of CPU utilization, LR prediction model is predicted the future CPU utilization of PMs. he authors also proposed an adjusted utilization threshold-based approach, named the interquartile range (IQR). It is a model of mathematical dissemination which is being equal to the difference between the 3^{rd} quartiles and 1^{st} quartiles. Another methods proposed by authors is Median Absolute Deviation (MAD). It is also a model of dissemination which is a strong estimator of range compared to the standard deviation.

Gholipour et al. [17] proposed a joint container consolidation approach and a multi Criteria migration decision (JVCMMD) approach between VMs and containers; any one of them should be migrated concurrently to optimize EC and SLAV. Zhou et al. [18] introduces a survey on energy-efficient approaches and discussed different VM consolidation approaches in the cloud computing environment.

Alsadie et al. [19] proposed an approach named as a dynamic threshold based fuzzy approach (DTFA). DTFA is used to detect overloaded and underloaded PMs. The authors proposed a VM placement policy by using the Lowest Interdependence Factor Exponent Multiple Resources Predictive (LIFE-MP) approach. The DTFA approach can alter the threshold values of PMs and the LIFE-MP approach select the PM in which migrated VMs can be placed to decrease the performance degradation because of migration. The LIFE-MP approach maintained the QoS as per as SLA with minimum number of VM migration.

Li et al. [20] proposed SLA-aware and Energy-efficient consolidation approach based on Robust Linear Regression (RobustSLR) prediction model. The authors proposed different methods to predict the residual. Based on the predicted residual final CPU utilization is predicted. Using RobustSLR prediction model authors proposed an overloaded/underloaded host detection approach and a RobustSLR Power Aware Best Fit Decreasing (RPABFD) VM allocation policy also proposed. In [21] authors proposed an SLA-aware and Energy-efficient consolidation using the Host States Naive Bayesian Prediction (HSNBP) model in cloud data centers. HSNBP model can predict future host states using the host CPU utilization to reduce EC and SLAV. HSNBP model can predict overloaded hosts using The Naive Bayesian classifier.

Yadav et al. [22] proposed adjustable energy-aware algorithms to reduce EC and SLAV in cloud data centers. The authors proposed regression-based algorithms, named Gdr and MCP to produce a progressive upper CPU utilization threshold to detect overloaded PMs. Using these approaches the host resource utilization also been optimized. The authors also proposed a Bandwidth-Aware (Bw) dynamic VM selection approach to maintaining an optimal balance between power consumption and the number of VMs migration.

Lin et al. [23] proposed a new VM consolidation procedure to enhance the execution time and reduce the EC of server clusters in cloud data centers. The VM consolidation is based on Peak Efficiency Aware Scheduling (PEAS) approach. PEAS consists of two approaches. The first approach is the Peak Efficiency Aware Placement (PEAP) approach which detects the most suitable host for every VM. The second approach is Peak Efficiency Aware Cost-efficient Reallocation (PEACR) to redistribute from overloaded/underloaded hosts. In [24] the authors proposed an energy-efficient along with QoS dynamic VM consolidation approach in the cloud computing environment. The authors proposed new host overloaded, host underload, VM selection, VM placement, and Enhancing Energy-Efficient and QoS VM Consolidation (EQVC) algorithms. The host overload detection algorithm is based on an autoregressive integrated moving average (ARIMA) prediction model. A host may consider as an underloaded host if the host has not performed any VM migration and not overloaded. Then underload detection approach detects a host as an underloaded host which has the maximum proportion of the host's EC to the CPU capacity. In the VM selection approach, invalid VM migration is reduced. When a VM is chosen to migrate, the migrating VM's CPU capacity loss is compared with the loss caused by an overloaded host. If the comparison returns true then the VM migration becomes invalid. The VM placement policy based on reserve CPU utilization for PMs to block unexpected variations in the workload and keep the regular load status of hosts. The EQVC approach detects overloaded hosts and selects the VMs for migration using the VM selection method and using a VM placement policy to recognize a new and suitable placement for the VM migration.

A Security-aware Dynamic VM Consolidation (SDVMC) framework, including a Security Monitoring Module (SMM) and an SDVMC module, was developed by Elshabka et al. [25]. There are so many proposed Dynamic VM Consolidation approaches to reduce cloud data centers' energy consumption. The overall security risk of cloud data centers has increased due to the inadequate safety measurement techniques and the consolidation of VMs without any understanding of their safety risk level. The authors claimed that the proposed framework decreased the overall risk by 2%–5% without increasing the energy consumption and reducing the QoS. Z. Li et al. [26] proposed an energy-efficient and quality-aware VM consolidation (EQ-VMC) approach to optimize the energy efficiency and service quality. The proposed method was based on heuristic evolutionary algorithms. The authors developed a discrete differential evolution algorithm for the global optimum solution for VM placement. By combining the solution with a suite of algorithms, the authors proposed host overload detection, VM selection, and underload host detection approach to decrease the EC and increase the QoS. J.P.B. Mapetu et al. [27] developed a dynamic VM consolidation method based on load balancing and low time complexity to reduce the EC and SLA violations. The authors proposed the Pearson correlation coefficient to correlate the resources for optimal VM placement. An efficient VM selection approach has also proposed to reduce the EC and SLA violation. S.M. Moghaddam et al. [28] proposed and embedding individualized machine learning prediction models for energy-efficient VM consolidation within Cloud data centers. To detect the over and underutilized hosts, the authors developed a fine-tuned Machine Learning (ML) prediction framework to determine the best time of each VM for migration. VM selection approach has been proposed based on migration time and host CPU usage of each VM.

R. Mandal et al. [29] developed a power-aware VM selection policy to reduce the total EC and SLA violation of data centers in the cloud computing environment. The utilization of each VM computed by the ratio of current VM utilization and total allocated resource to that VM. The VM selection approach selects the maximum utilized VM for migration. The power-aware VM selection policy significantly reduces total EC with minimal SLA violations. Monil et al. [30] proposed QoS-aware VM consolidation in cloud computing environments. The authors divided the Dynamic Virtual Machine Consolidation (DVMC) problem into sub-problems and proposed a VM placement policy in which a sleeping host can not be put into an active mode except it is very much essential. An optimized phase was proposed in which rearranging of VMs between host is being performed so that sleeping hosts activation becomes the last option. In [10,31] the authors proposed a greedy algorithm for VM placement, named Power-Aware Best-Fit Decreasing (PABFD) algorithm. In the PABFD algorithm, the host is chosen with the minimum power increasing next to the VM migration to the host. In this paper, a novel VM placement policy is proposed by modifying the PABFD algorithm.

3. Proposed system model

In this segment, firstly a concise presentation of the Simple Linear Regression is explained, and then the details of the proposed NLR prediction model.

3.1. Simple linear regression

Simple Linear Regression (SLR) is based on the approach proposed by Cleveland [32]. In the statistical method predicts the estimations of numeric or continuous characteristic is known as regression. It permits us to sum up and observe the association between two uninterrupted variables. A variable denoted by x is called the predictor variable and another one is denoted by y is called response variable [20]. The SLR predicts *the best fitting line* by the Eq. 1.

$$\hat{y}_i = \beta_0 + \beta_1 x_i \tag{1}$$

where, x_i is the predictor values, $\hat{y}_i$ is the predicted values, β_0 is the intercept and β_1 is the slope.

Eq. 1 predicts the actual values of y_i. The difference between predicted values of $\hat{y}_i$ with actual values of y_i is denoted as residual error and it can be calculated by the Eq. 2.

$$e_i = y_i - \hat{y}_i \tag{2}$$

The Least Squares Method [33] is used to reduce the error and find *the best fitting line* as:

$$Error = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2 \tag{3}$$

The estimated value of β_0 and β_1 can be determined if Eq. 4 and Eq. 5 are fulfilled.

$$\frac{dError}{d\beta_0} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0 \tag{4}$$

$$\frac{dError}{d\beta_1} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)x_i = 0 \tag{5}$$

Now the value of β_1 and β_0 can be estimated as:

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{6}$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} \tag{7}$$

where, $\bar{x}$ is mean of x_i and $\bar{y}$ is mean of y_i.

3.2. New linear regression model

New Linear Regression (NLR) model is based on the approach proposed by Yang Yu [34]. Although the Least Square Method is the most ideal approach to solve linear regression, however, its absolute error [35] is not the least. The fundamental idea of the model is to accept that the model goes through a straight line and a mean point. The Eq. 8 and Eq. 9 are fulfilled for the new model.

$$\beta_1 = \frac{\sum_{i=1}^{n}\frac{(y_i - \beta_0)}{x_i}}{n} \tag{8}$$

$$\bar{y} = \beta_1 \bar{x} + \beta_0 \tag{9}$$

Now the value of β_1 and β_0 can be estimated as:

$$\beta_1 = \frac{\sum_{i=1}^{n}\frac{y_i}{x_i} - \bar{y}\sum_{i=1}^{n}\frac{1}{x_i}}{n - \bar{x}\sum_{i=1}^{n}\frac{1}{x_i}} \tag{10}$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x} \tag{11}$$

3.3. VM consolidation using new linear regression model

The fundamental idea of VM consolidation is to run the maximum number of VMs in a host to increase the performance and reduce the EC of the host by data centers in a cloud computing environment. The live VM migration accomplished by the four different parts together: (1) Overload Hosts Detection, (2) Underload Hosts Detection, (3) Selection of few VMs from hosts that are overloaded and selection of entire VMs from hosts that are underloaded, (4) VM placement policy for entire VMs selected from overloaded hosts and underloaded hosts. Fig. 1 shows the different parts of VM Consolidation. In the following section each part is described in details.

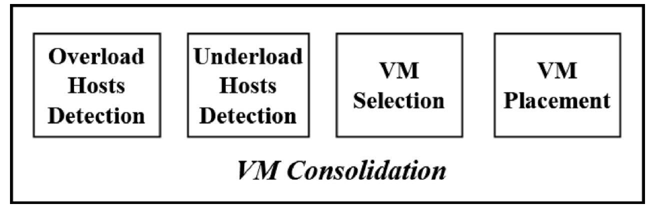

Fig. 1. Different parts of VM Consolidation.

3.3.1. Overload hosts detection

First of all, an Algorithm 1 is proposed to predict the CPU utilization of a host as per the reverse CPU utilization history by using New Linear

Regression (NLR) prediction model. The length of CPU utilization history is set as 10.

Now proposed Algorithm 1 has been used to detect the overload hosts. In Algorithm 1 mean of x as $\bar{x}$ and mean of y as $\bar{y}$ is computed using line number 7 and 8 respectively. The value of β_0 and β_1 is computed in line number 9 using Eq. 10 and Eq. 11 receptively. The migration interval is set in line number 10. Present utilization is calculated in line number 11 and final utilization is computed using line number 12. Line number 13 return the final utilization of the host.

Algorithm 1. New Linear Regression (NLR) prediction

Input: The set of reverse utilization history
Output: The prediction of utilization
1 **start**
2 $n \leftarrow$ $utilizationHistoryReversed.length$;
3 **for** $i \leftarrow 1$ *to* n **do**
4 $\quad x[i] \leftarrow i+1$;
5 $\quad y[i] \leftarrow utilizationHistoryReversed[i]$;
6 **end**
7 $\bar{x} \leftarrow \frac{\sum_{i=1}^{n} x_i}{n}$;
8 $\bar{y} \leftarrow \frac{\sum_{i=1}^{n} y_i}{n}$;
9 Compute the β_1 and β_0 using the Eq. 10 and Eq. 11 receptively;
10 $MigrationIntervals \leftarrow \frac{MaxVmMigrationTime}{SchedulingInterval}$;
11 $PrUtilization \leftarrow \beta_0 + \beta_1(\mathrm{n} + MigrationIntervals)$;
12 $FlPrUtilization \leftarrow SafetyParameter \times PrUtilization$;
13 return $FlPrUtilization$;
14 **stop**

Algorithm 2. Hosts Overload Detection (isHOD).

Input: Host
Output: Boolean
1 **start**
2 **if** *($utilizationHistory < 10$)* **then**
3 $\quad$ return $Host.getTotalRequestedMips() >$ $0.9 * Host.getTotalMips()$;
4 **end**
5 **else**
6 $\quad$ **for** $i \leftarrow 1$ *to* n **do**
7 $\quad\quad utilizationHistoryReversed[i] \leftarrow utilizationHistory[n-i+1]$;
8 $\quad$ **end**
9 $\quad Prediction \leftarrow NLR(utilizationHistoryReversed)$;
10 $\quad$ return $Prediction \geq 1$;
11 **end**
12 **stop**

Algorithm 2 represent the Overload Hosts Detection procedure. When the current CPU utilization is over ninety percent of the PM resources, and the host *utilizationHistory* length is less than 10, it will return line number 3 of Algorithm 2. Otherwise, compute the reverse utilization history using line numbers 5, 6, 7, and 8. The utilization of a host is predicted using Algorithm 1 by line number 9. Finally, a host is overloaded or not for the next time unit is predicted by line number 10.

3.3.2. Underload hosts detection

One of the crucial steps of VM consolidation is the identification of underloaded hosts. All VMs of underloaded hosts need to migrate to other hosts and shutdown all underloaded hosts to minimize EC and increase resource utilization. Algorithm 3 shows the proposed underload hosts detection approach. In Algorithm 3 underutilized hosts have been detected based on a lower threshold.

Algorithm 3. Hosts Underload Detection (isHUD).

Input: Host
Output: Boolean
1 **start**
2 **if** *($utilizationHistory < 10$)* **then**
3 $\quad$ return $Host.getTotalRequestedMips() <$ $0.1 * Host.getTotalMips()$;
4 **end**
5 **else**
6 $\quad$ **for** $i \leftarrow 1$ *to* n **do**
7 $\quad\quad utilizationHistoryReversed[i] \leftarrow utilizationHistory[n-i+1]$;
8 $\quad$ **end**
9 $\quad Threshold \leftarrow 0.6(1 - SafetyParameter \times IQR)$;
10 $\quad Prediction \leftarrow NLR(utilizationHistoryReversed)$;
11 $\quad$ return $Prediction < Threshold$;
12 **end**
13 **stop**

value. The lower threshold value has been set based on the interquartile range (IQR) approach. IQR is a contrast within the third quartiles and first quartiles; it can be represented by the Eq. 12.

$$IQR = Q_3 - Q_1 \tag{12}$$

Eq. 13 represented the lower threshold value to detect the underload hosts.

$$Threshold = 0.6(1 - SafetyParameter \times IQR) \tag{13}$$

When the current utilization of CPU is under ten percent of the PM resources and the host *utilizationHistory* length is less than 10 then it will return line 3 of Algorithm 3. Otherwise, prediction is computed using Algorithm 1 in line 10 and finally return underloaded host for next time unit in line 11.

3.3.3. VM selection

VM migration is accomplished to reduce the EC and increase the performance by selecting a few VMs from hosts that are overloaded and entire VMs from hosts that are underloaded. It also reduced the performance degradation because of VM migration. In this research work, the Minimum Migration Time (MMT) [10] policy has been chosen for VM selection. The MMT policy selects a VM that requires least an ideal opportunity to migrate contrast with different VMs. The migration time is computed by performing division between the measure of RAM utilization by a VM and the network bandwidth accessible for a host.

3.3.4. VM placement

The Power-Aware Best-fit Decreasing (PABFD) [10,31] policy is a classical approach used for VM placement in a cloud computing environment. In PABFD policy all the VMs are sorted decreasingly as per the CPU utilization. Then each VM finds the suitable hosts in which the EC is least. Finally, it assigns the VM to the minimum power consumed host to increase the energy efficiency.

Algorithm 4. Modified Power Aware Best Fit Decreasing (MPABFD).

(continued on next page)

Table 1
Characteristics of Hosts [10].

Host	Clock Speed	Cores	RAM	Bandwidth
HP ProLiant ML110 G4	1860 MHz	2	4 GB	1 Gbps
HP ProLiant ML110 G5	2660 MHz	2	4 GB	1 Gbps

(*continued*)

```
Input: host_list = set of Hosts, vm_list = set of VMs
Output: allocation
start
vm_list.sortDecreasingCPUUtilization();
for vm_list.contain(vm) is true do
    minPower← MAX;
    allocated_host← NULL;
    for host_list.contain(host) is true do
        if (isHOD(host)) then
            continue;
        end
        if (isHUD(host)) then
            continue;
        end
        if (host.isSuitableForVm(vm)) then
            power←getPowerAfterAllocation(host, vm);
            powerDiff←power−host.getPower();
            if (powerDiff < minPower) then
                minPower←powerDiff;
                allocatedHost←host;
            end
        end
        if alloctedHost is not NULL then
            allocation.add(vm, allocatedHost);
        end
    end
end
return allocation;
stop
```

A modified PABFD has been proposed by using NLR prediction Algorithm 1. Algorithm 4 depicted the proposed Modified Power-Aware Best-fit Decreasing (MPABFD) approach. In the MPABFD algorithm, the overload and underload hosts are eliminated from the *host_list* to perform a VM placement by predicting the future host load using the NLR prediction algorithm.

4. Result and analysis

In this section, overall performance and the effectiveness of the proposed algorithms have been illustrated. CloudSim simulation toolkit [36] has been used to assess the effectiveness of proposed algorithms in a viable cloud computing environment. CloudSim can be extended to handle the resources and then implementation and evaluation can be accomplished.

4.1. Experimental setup

Prior to continuing further, let's present the experimental setup used to evaluate proposed algorithms and also in comparison with some benchmark algorithms described in Section 2. One data center consisting of 800 different hosts is chosen, between them, 50% of hosts are HP ProLiant ML110 G4 servers clocked at 1,860 1860 MHz, and remaining are HP ProLiant ML110 G5 servers clocked at 2,660 MHz. Each one has 2 cores, 4 GB memory, 1 GB/s network bandwidth. The characteristics of hots have shown in Table 1.

Table 2
Characteristics of VMs [16].

VM Instances	Clock Speed	Cores	RAM
Micro Instance	500 MHz	1	613 MB
Small Instance	1000 MHz	1	1740 MB
Extra large Instance	2000 MHz	1	1740 MB
High-CPU Medium Instance	2500 MHz	1	870 MB

Table 3
Characteristics of Workload [10].

Date	No. of VMs	No. of Hosts	Mean (%)	St.dev. (%)
03-03-2011	1052	800	12.31	17.09
09-03-2011	1061	800	10.70	15.57
22-03-2011	1516	800	9.26	12.78
20-04-2011	1033	800	10.43	15.21
Random	50	50	–	–

The Amazon EC2 [37], types of VMs instances have been used for the analysis of the proposed algorithms. Four different types of VMs are available. Depending on the workload any one of the VM instances has been created into the hosts. The characteristics of VMs have shown in Table 2.

4.1.1. Workload

The analysis of the proposed NLR prediction model has been implemented by the real world workload as well as random workload:

- Real world workload: The real world workload is the data from PlanetLab [38]. PlanetLab collected the real data from an infrastructure monitoring framework, called CoMon [39]. It gathered utilization information at regular intervals from a huge number of servers situated over 500 places in the world. Between them, four days of data is chosen for the analysis of the proposed NLR prediction model.
- Random workload: In a random workload, the cloud users sends a request for fifty heterogeneous VM to the fifty hosts. The CPU utilization is produced on a random variable when a VM execute an application. Each application consists of three hundred bytes of input and output. The characteristics of each workload have shown in Table 3.

4.2. Performance evaluation

The performance measurement and comparison is an integral part of a research work. The performance evaluation of the proposed algorithms has done and compared with some renowned algorithms. The renowned Overloaded Hosts Detection algorithms are Linear regression (LR), Median absolute deviation (MAD), Inter-quartile range (IQR), Static Threshold (THR); the VM placement policy is Power-Aware Best-Fit Decreasing (PABFD) [10]. The effectiveness of the proposed prediction model is measured and compared by the metrics proposed by Beloglazov et al. [10].

4.2.1. SLA Time Performance Active Host (SLATAH)

SLATAH is the total time when an active host underwent a hundred percent CPU utilization and it can be represented by Eq. 14.

$$SLATAH = \frac{1}{N_{host}} \sum_{i=1}^{N_{host}} \frac{T_{CPU_i}}{T_{Active_i}} \tag{14}$$

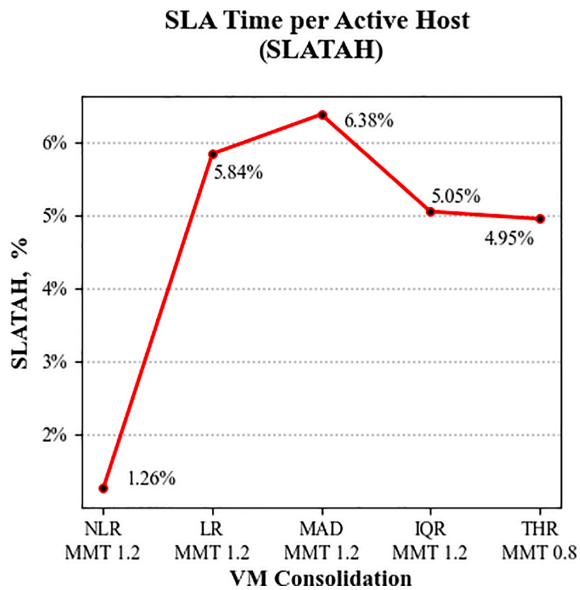

Fig. 2. SLA Time Performance Active Host (SLATAH) comparison.

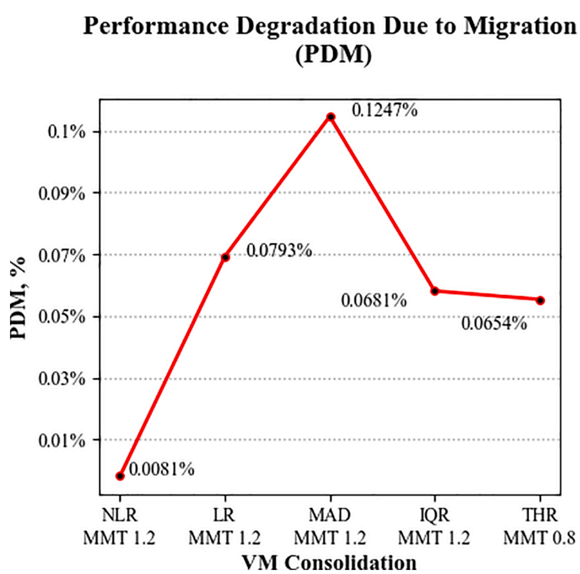

Fig. 3. SLA Performance Degradation due to Migration (PDM) comparison.

where,

- N_{host} is the number of Hosts.
- T_{CPU_i} is the total time accomplished 100% CPU utilization of hot i to an SLAV.
- T_{Active_i} is the time when host i is in active state.

Fig. 2 shows the comparative analysis of SLATAH of the proposed NLR prediction model with existing prediction models.

4.2.2. SLA Performance Degradation due to Migration (PDM)

SLA performance degradation due to migration (PDM) is an SLA-based metric. It is computed by Eq. 15.

$$PDM = \frac{1}{M_{vm}} \sum_{j=1}^{M_{vm}} \frac{C_{Deg_j}}{C_{CPU_j}} \tag{15}$$

where,

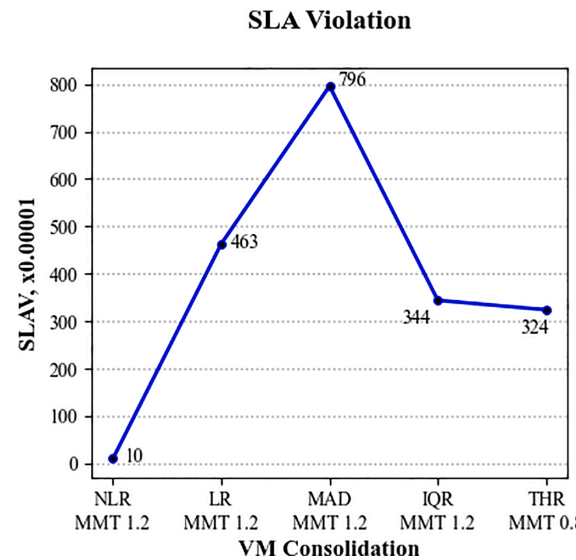

Fig. 4. SLA Violation comparison.

- M_{vm} is the number of Virtual machines.
- C_{Deg_j} is the performance degradation of VM j due to migration.
- C_{CPU_j} is the total capacity requested by VM j during its life time.

Due to live VM migration performance may be degraded. Fig. 3 shows the comparative analysis of PDM of the proposed NLR prediction model.

4.2.3. Service Level Agreement Violation (SLAV)

QoS is maintained by the cloud service provider through SLA. SLAs consist of several parameters to satisfy cloud users. So, SLAV becomes a very crucial metric to understand the level of QoS supplied by the cloud service provider and it can be computed by Eq. 16.

$$SLAV = SLATAH \times PDM \tag{16}$$

where,

- $SLAV$ is the service level agreement violation in percentage.

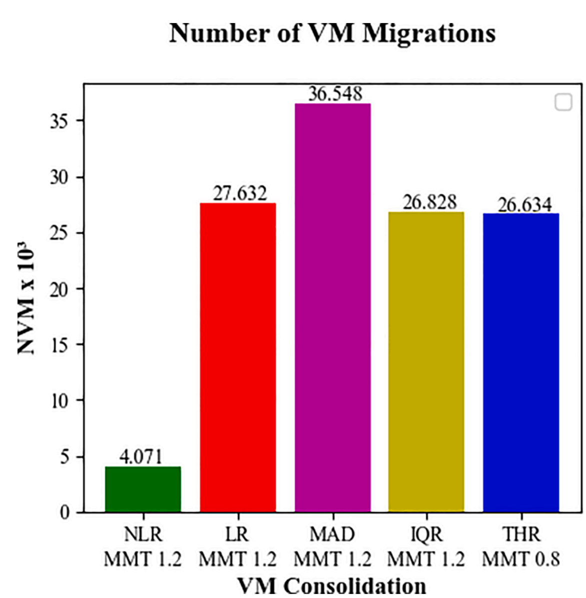

Fig. 5. Number of VM migrations (VMMG) comparison.

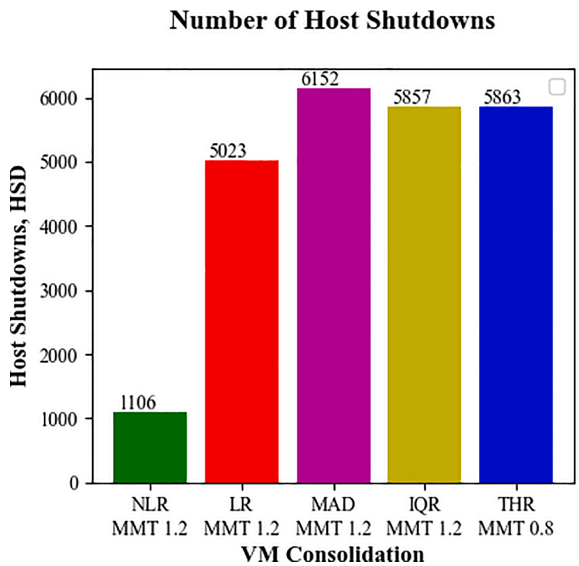

Fig. 6. Number of host shutdowns (HSD) comparison.

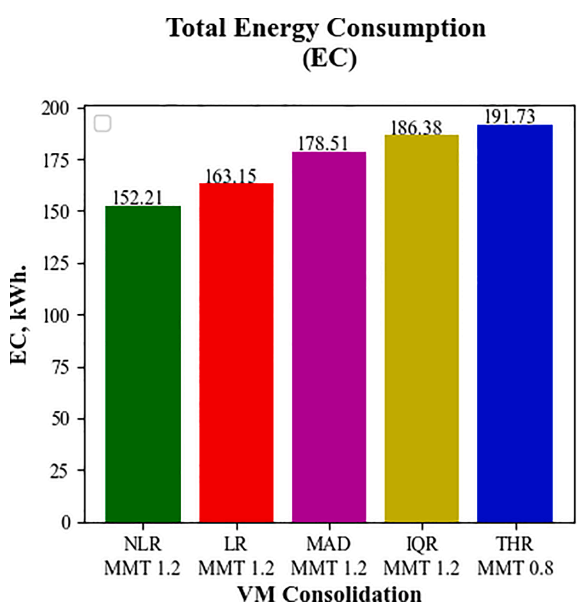

Fig. 7. Total Energy Consumption (EC) comparison.

- *SLATAH* is the time of an active host accomplished 100% CPU utilization.
- *PDM* is the performance degradation during VMs migration in percentage.

Fig. 4 shows the comparative analysis of SLAV of the proposed NLR prediction model.

4.2.4. Number of VM migrations (VMMG)

Live migrations is a technique utilized for load adjusting and

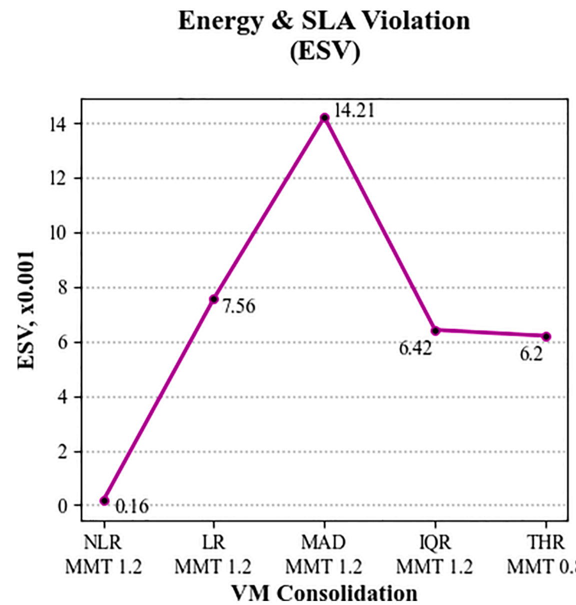

Fig. 8. Energy and SLA Violation (ESV) comparison.

enhancement of VM deployment in data centers. There is mainly two methods to perform live migrations i.e. (i) Pre-copy method (ii) Post-copy method [40]. In the Pre-copy method, memory is copied with several iterations to the destination host and then the execution is transferred and in the Post-copy method [41], VM's execution state is migrated and then each memory page is transferred only once. So, live migrations becomes an important aspect to provide QoS to cloud users. But too much migration can increase the cost of migration as well as total EC. Therefore, an optimal amount of VM migrations not only can improve QoS but also can decrease total EC. Fig. 5 shows the comparative analysis of VMMG.

4.2.5. Number of host shutdowns (HSD)

Number of host shutdowns (HSD) is a migration-based metric. Number of VMs migrations may increase number of HSD. The total EC may increase due to repeated turning on or off of hosts. Fig. 6 shows the comparative analysis of HSD. Fig. 6 shows the comparative analysis of HSD.

4.2.6. Total energy consumption (EC)

In modern cloud computing, total energy consumption becomes the key concern to the researchers. The main objective is to decrease total EC and SLAV with adequate QoS. It can be computed by Eq. 17 and Fig. 7 shows the comparative analysis of total EC. Table 4 shows the EC of hosts utilized in this research.

$$EC = \sum_{j=1}^{n} EC_{Host_j} \tag{17}$$

Where, EC_{Host_j} is energy consumption of $Host_j$.

4.2.7. Energy and SLA Violation (ESV)

Energy consumption is inversely proportional to SLA violation. So, the decrease in EC will increase SLAV. The balance between total EC and

Table 4
Energy consumption by hosts at different load levels in Watts [10].

PM	Energy Consumption (in Watts) at Different Load on Hosts										
	0%	10%	20%	30%	40%	50%	600%	70%	80%	90%	100%
G4	86	89.6	92.6	96	99.5	102	106	108	121	114	117
G5	93.7	97	101	105	110	116	121	125	129	133	135

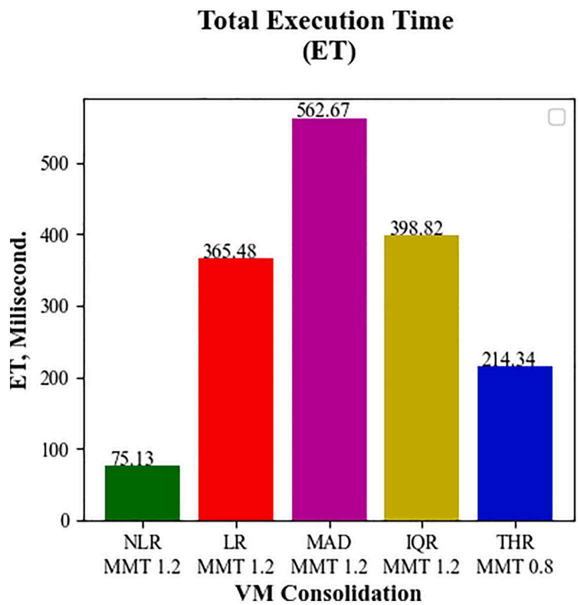

Fig. 9. Total Execution Time (ET) comparison.

SLAV is measured by Energy and SLA violation (ESV), and it can be computed by Eq. 18.

$$ESV = EC \times SLAV \tag{18}$$

Fig. 8 shows the comparative analysis of Energy and SLAV (ESV).

4.2.8. Total Execution Time (ET)

Total execution time (ET) is the time required to complete the execution of an algorithm for a given data set. Fig. 9 shows the comparative analysis of total execution time (ET) and it shows the proposed prediction model execute much faster than other benchmark prediction models. The median execution time of the proposed prediction model is 75.13 ms, which is much lesser than the other prediction models.The comparative analysis of all above mentioned metrics is shown in Table 5.

4.2.9. Safety parameter

The safety parameter is an important parameter for the proposed Algorithm 3. The safety parameter can deal with the energy performance trade-off. It defines the intense aggregation of VMs on physical servers in the cloud environment. Probabilities for energy conservation are going too low if the safety conditions are also tightened. On the contrary, if the safety parameter is too convenient, the rates of service level agreement violations appear to be disproportionately high. Hence, an experimental collection of safety parameters becomes significant that makes an optimal trade-off between EC and SLA violations.

In the proposed prediction model, the value of safety parameters varying from 0.5 to 2.5 with a different workload. Fig. 10 shows the total EC with different safety parameters. Fig. 11 and Fig. 12 show the analysis of SLAV and ESV respectively. Fig. 10 to Fig. 12 shows that lower value of safety parameter increases SLAV and ESV. Therefore, the optimal value of a safety parameter is chosen as 1.2 to execute and evaluate the above algorithms.

5. Conclusion

In this paper, a novel Linear Regression (NLR) prediction model has been proposed. A new overloaded and underloaded hosts detection algorithm based on the NLR prediction model and a new VM placement policy namely Modified Power-Aware Best-Fit Decreasing (MPABFD) algorithm has also proposed.

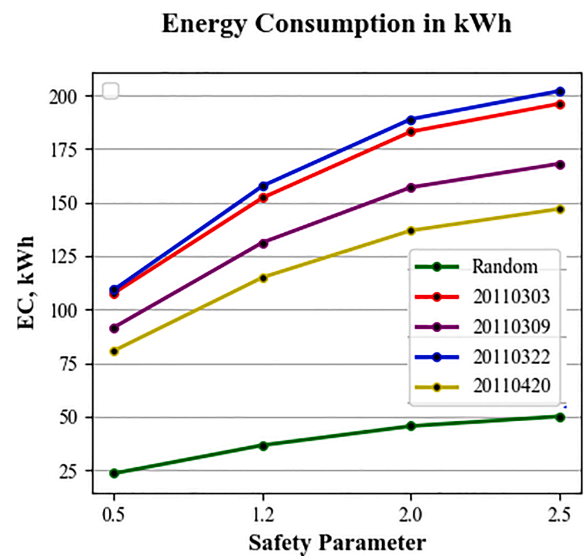

Fig. 10. Total Energy Consumption with different safety parameters.

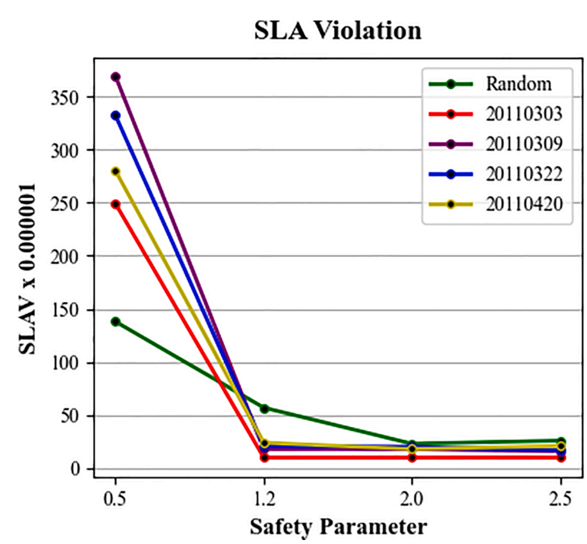

Fig. 11. SLA violation with different safety parameters.

Table 5
Comparative analysis of NLR prediction model with renowned prediction models

Policy	Energy in kWh.	SLATAH (%)	PDM (%)	SLAV $\times 10^{-5}$	Migration $\times 10^{3}$	Host Shutdw.	ESV $\times 10^{-3}$	Exe.Time in Millisec.
THR-MMT-0.8	191.73	4.95	0.0654	324	26.634	5863	6.20	214.34
LR-MMT-1.2	163.15	5.84	0.0793	463	27.632	5023	7.56	365.48
IQR-MMT-1.2	186.38	5.05	0.0681	344	26.828	5857	6.42	398.82
MAD-MMT-1.2	178.51	6.38	0.1247	796	36.548	6152	14.21	562.67
NLR-MMT-1.2	152.21	1.26	0.0081	10	4.071	1106	0.16	75.13

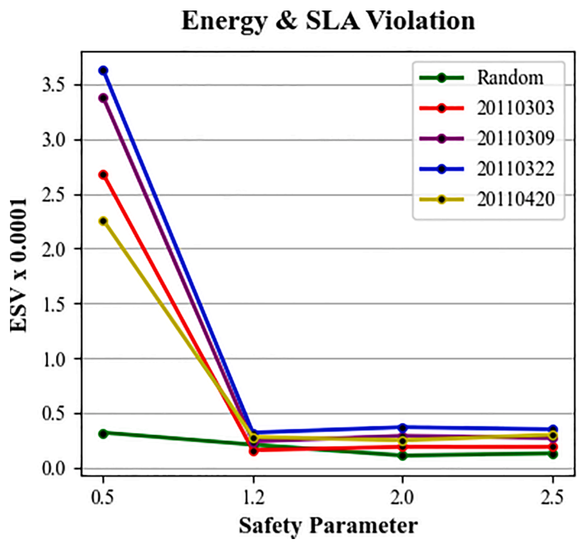

Fig. 12. Energy and SLA violation with different safety parameters.

The result and performance analysis has shown that the proposed prediction model significantly defeats the renowned prediction models. The proposed prediction model significantly reduces EC and SLAV which can be used to build a smart and sustainable environment for a smart city by CoT. The main drawback of the proposed NLR prediction model is that it assumes the model goes through a straight line and a mean point. In the future, we plan to improve the NLR prediction model so that it can overcome the drawback and efficiently decrease SLAV and EC in cloud data centers.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Nayyar A. Handbook of Cloud Computing: Basic to Advance research on the concepts and design of Cloud Computing. BPB Publications; 2019.
[2] Yin K. Cloud computing: Concept, model, and key technologies. ZTE Commun 2020;8(4):21–6.
[3] Trueman C. Why data centres are the new frontier in the fight against climate change,https://www.computerworld.com/article/3431148/why-data-centres-are-the-new-frontier-in-the-fight-against-climate-change.html (2019(accessed on 27th June 2020)).
[4] Vidal J. 'Tsunami of data' could consume one fifth of global electricity by 2025, https://www.climatechangenews.com/2017/12/11/tsunami-data-consume-one-fifth-global-electricity-2025/ (2019(accessed on 25th June 2020)).
[5] Parwekar P. From internet of things towards cloud of things. In: 2011 2nd International Conference on Computer and Communication Technology (ICCCT-2011), IEEE; 2011. p. 329–333.
[6] Qian Y, Wu D, Bao W, Lorenz P. The internet of things for smart cities: Technologies and applications. IEEE Network 2019;33(2):4–5.
[7] Heidari S, Buyya R. Quality of Service (QoS)-driven resource provisioning for large-scale graph processing in cloud computing environments: Graph Processing-as-a-Service (GPaaS). Future Gener Comput Syst 2019;96:490–501.
[8] Beloglazov A, Buyya R. Adaptive threshold-based approach for energy-efficient consolidation of virtual machines in cloud data centers, MGC@ Middleware 4.
[9] Upadhyay PK, Pandita A, Joshi N. Scaled Conjugate Gradient Backpropagation based SLA Violation Prediction in Cloud Computing. In: 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), IEEE; 2019. p. 203–208.
[10] Beloglazov A, Buyya R. 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 2012;24(13):1397–420.
[11] Liu H, Jin H, Xu C-Z, Liao X. Performance and energy modeling for live migration of virtual machines. Cluster Comput 2013;16(2):249–64.
[12] Farahnakian F, Pahikkala T, Liljeberg P, Plosila J, Tenhunen H. Utilization prediction aware VM consolidation approach for green cloud computing. In: 2015 IEEE 8th International Conference on Cloud Computing. IEEE; 2015. p. 381–8.
[13] Wang M, Meng X, Zhang L. Consolidating virtual machines with dynamic bandwidth demand in data centers. In: 2011 Proceedings IEEE INFOCOM, IEEE; 2011. p. 71–75.
[14] Joshi S, Kaur S. Cuckoo search approach for virtual machine consolidation in cloud data centre. In: International Conference on Computing, Communication & Automation, IEEE; 2015. p. 683–686.
[15] Beloglazov A, Abawajy J, Buyya R. Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future Gener Comput Syst 2012;28(5):755–68.
[16] Beloglazov A. Energy-efficient management of virtual machines in data centers for cloud computing [Ph.D. thesis]; 2013.
[17] Gholipour N, Arianyan E, Buyya R. A novel energy-aware resource management technique using joint VM and container consolidation approach for green computing in cloud data centers. Simul Model Pract Theory 2020;102127.
[18] Zhou Q, Xu M, Gill SS, Gao C, Tian W, Xu C, Buyya R. Energy efficient algorithms based on VM consolidation for cloud computing: comparisons and evaluations, arXiv preprint arXiv:2002.04860; 2020.
[19] Alsadie D, Tari Z, Alzahrani EJ. Online VM Consolidation in Cloud Environments. In: 2019 IEEE 12th International Conference on Cloud Computing (CLOUD). IEEE; 2019. p. 137–45.
[20] Li L, Dong J, Zuo D, Wu J. Sla-aware and energy-efficient VM consolidation in cloud data centers using robust linear regression prediction model. IEEE Access 2019;7:9490–500.
[21] Li L, Dong J, Zuo D, Liu J. SLA-Aware and Energy-Efficient VM Consolidation in Cloud Data Centers Using Host States Naive Bayesian Prediction Model. In: 2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom), IEEE; 2018. p. 80–87.
[22] Yadav R, Zhang W, Kaiwartya O, Singh PR, Elgendy IA, Tian Y-C. Adaptive energy-aware algorithms for minimizing energy consumption and SLA violation in cloud computing. IEEE Access 2018;6:55923–36.
[23] Lin W, Wu W, He L. An on-line virtual machine consolidation strategy for dual improvement in performance and energy conservation of server clusters in cloud data centers. IEEE Trans Services Comput.
[24] Liu Y, Sun X, Wei W, Jing W. Enhancing energy-efficient and QoS dynamic virtual machine consolidation method in cloud environment. IEEE Access 2018;6: 31224–35.
[25] Elshabka MA, Hassan HA, Sheta WM, Harb HM. Security-aware dynamic VM consolidation. Egypt Inform J 2020.
[26] Li Z, Yu X, Yu L, Guo S, Chang V. Energy-efficient and quality-aware VM consolidation method. Future Gener Comput Syst 2020;102:789–809.
[27] Mapetu JPB, Kong L, Chen Z. A dynamic VM consolidation approach based on load balancing using Pearson correlation in cloud computing. J Supercomput 2020: 1–42.
[28] Moghaddam SM, O'Sullivan M, Walker C, Piraghaj SF, Unsworth CP. Embedding individualized machine learning prediction models for energy efficient VM consolidation within Cloud data centers. Future Gener Comput Syst 2020;106: 221–33.
[29] Mandal R, Mondal MK, Banerjee S, Biswas U. An approach toward design and development of an energy-aware VM selection policy with improved SLA violation in the domain of green cloud computing. J Supercomput 2020:1–20.
[30] Monil MAH, Malony AD. QoS-aware virtual machine consolidation in cloud datacenter. In: 2017 IEEE International Conference on Cloud Engineering (IC2E), IEEE; 2017. p. 81–7.
[31] Beloglazov A, Buyya R. Managing overloaded hosts for dynamic consolidation of virtual machines in cloud data centers under quality of service constraints. IEEE Trans Parallel Distributed Syst 2012;24(7):1366–79.
[32] Cleveland WS. Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc 1979;74(368):829–36.
[33] Winner L. Simple Linear Regression I-Least Squares Estimation,https://users.stat.ufl.edu/winner/qmb3250/notespart2.pdf (2003(accessed on 2nd June 2020)).
[34] Yu Y. Research of a new method for solving linear regression. In: 2018 International Conference on Transportation & Logistics, Information & Communication, Smart City (TLICSC 2018), Atlantis Press; 2018.
[35] Hawley RW, Gallagher NC. On Edgeworth's method for minimum absolute error linear regression. IEEE Trans Signal Process 1994;42(8):2045–54.
[36] Calheiros RN, Ranjan R, Beloglazov A, De Rose CAF, Buyya R. CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. Software: Practice and Experience 2011;41(1): 23–50.
[37] A. EC. Amazon ec2 instance types,https://aws.amazon.com/ec2/instance-types/ (2019(accessed on 10th June 2020)).
[38] Peterson L, Bavier A, Fiuczynski ME, Muir S. Experiences building planetlab. In: Proceedings of the 7th symposium on Operating systems design and implementation; 2006. p. 351–66.
[39] Park K, Pai VS. CoMon: a mostly-scalable monitoring system for PlanetLab. ACM SIGOPS Operating Systems Review 2006;40(1):65–74.
[40] Clark C, Fraser K, Hand S, Hansen JG, Jul E, Limpach C, Pratt I, Warfield A. Live migration of virtual machines. In: Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation-vol. 2; 2005. p. 273–286.
[41] Hines MR, Gopalan K. Post-copy based live virtual machine migration using adaptive pre-paging and dynamic self-ballooning. In: Proceedings of the 2009 ACM SIGPLAN/SIGOPS international conference on Virtual execution environments; 2009. p. 51–60.