symmetry

Article

Dynamic Dual-Threshold Virtual Machine Merging Method Based on Three-Way Decision

Jin Yang and Guoming Zhang *

School of Computer Science, Guangdong Business and Technology University, Zhaoqing 526020, China
* Correspondence: beike2004@126.com

Abstract: Cloud computing, an emerging computing paradigm, has been widely considered due to its high scalability and availability. An essential stage of cloud computing is the cloud virtual machine migration technology. Nevertheless, the current trigger timing of virtual machine migration in cloud data centers is inaccurate, resulting in insufficient virtual machine consolidation. Furthermore, the high and low workload fluctuations are also a potential symmetrical problem worthy of attention. This paper proposes a virtual machine energy-saving merging method based on a three-way decision (ESMM-3WD). Firstly, we need to calculate the load fluctuation of the physical machine and divide the load fluctuation into three parts. Furthermore, the corresponding mathematical model predicts the load according to the different classification categories. Then, the predicted load value is used to dynamically adjust the threshold to improve the virtual machine merge probability. Finally, the simulation experiment is carried out on the cloud computing simulation platform cloudsim plus. The experimental results show that the virtual machine energy-saving merging method based on the three-way decision proposed in this paper can better reduce the number of migrations, increase the number of physical machines shut down, better improve the probability of virtual machine merger, and achieve the purpose of reducing the energy consumption of the data center.

Keywords: three-way decision; cloud computing; dynamic thresholds; virtual machine migration; energy consumption optimization

check for updates

Citation: Yang, J.; Zhang, G. Dynamic Dual-Threshold Virtual Machine Merging Method Based on Three-Way Decision. *Symmetry* **2022**, *14*, 1865. https://doi.org/10.3390/sym14091865

Academic Editors: Weihua Xu, Yanhong She and Xiaoyan Zhang

Received: 11 August 2022
Accepted: 31 August 2022
Published: 7 September 2022

1. Introduction

Cloud computing is a service related to information technology, software, and the Internet, and it is centered on the Internet and provides fast and secure cloud computing services and data storage on websites so that everyone can use enormous computing resources and convenient services through the Internet [1–3]. With the vigorous development of cloud computing, users' requirements for cloud computing have increased dramatically, and the needs are also diverse [4–6]. However, the vast energy cost of data centers is becoming more and more prominent, and the utilization of computing resources of data centers is low. It is estimated that the average resource utilization of a data center is below 30% [7]. At the same time, the energy consumption of idle physical machines accounts for more than 70% of the peak energy consumption [8].

The utilization of computing resources is an important factor affecting the energy consumption of cloud computing. To make better use of the computing resources and reduce the energy consumption of the data center, an effective and commonly used method to solve the problem of huge energy consumption in data centers is virtual machine consolidation. Virtual machine consolidation refers to virtual machine migration, placing it on fewer servers according to the resource requirements of virtual machines and then shutting down some servers. However, frequent and large-scale migration also generates a lot of energy consumption. Therefore, minimizing the number of virtual machine migrations is also necessary to achieve balanced energy consumption optimization in virtual machine consolidation.

Note that reducing the number of virtual machine migrations is one of the most critical in the timing of the virtual machine migration. When the timing is accurate enough, it will improve the whole process of the virtual machine integration process. Judging the virtual machine migration trigger time requires setting an appropriate threshold. Threshold setting is generally divided into two types. One is static the threshold setting, which cannot be changed while the workload of the cloud datacenter is constantly changeable, and it also cannot be adjusted dynamically according to the dynamic change of load, making it difficult for the dynamic merging of virtual machines to improve the rate of computing resources. The other is the dynamic threshold setting, which effectively alleviates this problem and can adaptively adjust the threshold according to the change in the historical workload and achieve better resource allocation under the condition of complex resource requirements [9].

The problem of the frequent fluctuations and changes of cloud data centers is not suitable for a static threshold setting. The setting of dynamic thresholds with a single model is not efficient. There is a potential symmetrical problem with the high and low fluctuations, which is worthy of attention. We propose a virtual machine energy-saving merging method based on a three-way decision (ESMM-3WD), which firstly divides the fluctuation of historical load utilization into three parts based on the three-way decision. Then, using different mathematical models to predict the workload and adjust the threshold according to the predicted load for dynamically achieving virtual machine integration.

Three-way decision (3WD) is a native three-level decision theory proposed by Yao [10], essential elements including trisecting, acting, and outcome, i.e., the TAO model. In recent years, this theory has been widely used in cloud computing, and there are many "3 phenomena" in cloud data centers, such as cloud tasks, which include long-time, medium-time, and short-time tasks. The workload can be divided into low, medium, and high loads. The physical machine can be divided into active, sleep, shutdown three states, etc. This paper divides the fluctuation value of physical machine historical load into three types based on the three-way decision: high, medium, and low, and uses the corresponding mathematical model to predict the load. Finally, the predicted load is used to set the threshold dynamically, and then a virtual machine energy-saving combination method based on the three-way decision is proposed [11–14]. The main contributions of the current article are as follows:

1. We establish the correlation between virtual machine migration and granular computing theory to deeply analyze the widespread granularity phenomenon in the process of virtual machine migration. Then, we design the corresponding multi-granularity intelligent decision-making scheme to optimize the energy consumption;
2. Compared with other methods, this paper studies the performance of virtual machine migration under a particular granularity, that is, a magic number 3, and constructs the dynamic migration mechanism of the virtual machine and the analysis method of three-way decision;
3. Based on the three-way decision model, we propose a novel cloud resource prediction algorithm to study the quasi-periodic effect of cloud resource demand and dynamically adjust the virtual machine migration mechanism via predicting the cloud resource demand.

The remainder of this paper is organized as follows: Section 2 mainly introduces some related research on the threshold-triggered virtual machine. Section 3 describes the problems that need to be solved, establishes mathematical models, and describes the core algorithms. Section 4 puts forward the measurement indicators and compares and analyzes some algorithms. Finally, this paper is concluded with further work in Section 5.

2. Related Work

Threshold-based triggering methods are divided into two types: one is the static threshold triggering method, and the other is the dynamic threshold triggering method. As we all know, a static single threshold triggering strategy only determines that the physical machine is overloaded when the utilization rate of the physical machine is more significant than 85% to trigger the migration of the virtual machine. However, this single threshold static triggering of the migration of the virtual machine will lead to a surplus of physical machine resources, resulting in a significant waste of resources. Virtual machine migration will be triggered when the CPU utilization of physical machine load is lower than the lower threshold and higher than the upper threshold. However, static thresholds cannot adapt to the complex changes in workload, which leads to the low efficiency of resource integration.

For virtual machine migration, Arshad et al. [15] propose the Energy Efficiency Heuristic with VM Consolidation, which reduces power consumption while reducing SLA violations. Bahrami et al. [16] propose using the time series prediction method and double smooth development technique to predict the processor efficiency in the future and also propose the optimal relationship by the dynamic threshold that improves SLA performance, reducing VM migrations and optimizing energy consumption. Liu et al. [17] propose an enhancing energy-efficient and QoS dynamic VM consolidation method, which consists of four algorithms that correspond to different stages in VM consolidation, guaranteeing QoS, reducing VM migrations, and optimizing energy consumption. Khan et al. [18] propose a normalization-based VM consolidation strategy that places virtual machines in an online strategy while minimizing energy consumption, SLA violations, and the number of VM migrations.

For the study of dynamic thresholds, Seth et al. [19] propose that using dynamic thresholds has significant effects in reducing the number of VM migrations, improving SLA performance, and optimizing energy consumption. Yan et al. [20] propose a virtual machine migration algorithm based on dynamic threshold adjustment. The algorithm mainly takes the historical workload as the main parameter, analyzes the historical load, and predicts the load. Adjust the virtual machine migration threshold dynamically to achieve a more accurate timing for triggering virtual machine migration.Deafallah et al. [21] use the remaining capacity and the difference between the predicted value and the current value to predict the threshold. This algorithm can convert the physical machine load fluctuation data into stable data, calculate the difference between the predicted load and the current load, add or subtract the calculated value and the current threshold, and finally obtain the dynamic change of the threshold. In ref. [22], a distributed mathematical model calculates the intervals of physical machine loads to achieve dynamic adjustment thresholds. The above study of dynamic thresholds only uses a single mathematical model to predict load utilization, which is unsuitable for complex and variable workload change.

The basic theory of a three-way decision is to think with three, and the subsequent research has established the trisection-acting-outcome model (TAO model of three-way decision) [23,24], and there are numerous "3" phenomena in the cloud data center. According to the three-way decision model, the workload fluctuation is divided into three regions, and then different prediction models are used to predict the load. Finally, the threshold is adjusted dynamically to carry out the virtual machine consolidation for improving resource [2,25,26]. The related studies are shown in Table 1.

Table 1. Research comparison of dynamic thresholds.

Ref	Utilized Algorithm	Evaluation Metric	Strengths	Weaknesses	Year
Arshad et al. [15]	EEHVMC	SLA, energy consumption, VM migrations, Performance degradation, Execution time	Meet SLA, save energy consumption, reduce the number of VM migrations	I/O intensive tasks are not considered	2022
Bahrami et al. [16]	OMMD	SLA, energy consumption, VM migrations	The number of migrations of virtual machines, energy consumption and the rate of SLA violations are improved	On the real cloud infrastructure is not exactly clear	2021
Liu et al. [17]	EQVC	QOS, VM migrations, energy consumption	Reduce the amount of VMs migrations, low operating costs, meet SLA	Neglects the essential factors like workload, type of host, and temperature	2018
Khan et al. [18]	NVMC	SLA, energy consumption, VM migrations	Improvement in Quality of Service, Meet SLA	Performance degradation	2021
Seth et al. [19]	STA	SLA, power consumption, VM migrations	Meet SLA, save power consumption, reduce the number of VM migrations	Host type is not considered	2017
Yan et al. [20]	AOTS-VMDC	SLAV, energy consumption	Meet SLAV, save power consumption, reduce the number of VM migrations	The influence of memory, network adn disk is ignored	2016
Deafallah et al. [21]	DTFA	SLAV, energy consumption, VM migrations	The number of migrations of virtual machines, energy consumption and the rate of SLA violations are improved	Memory and bandwidth are not considered	2018
Beloglazov et al. [22]	A novel technique for dynamic consolidation of VMs based on adaptive utilization thresholds	SLA, energy consumption, VM migrations	Meet SLA, save power consumption,reduce the number of VM migrations	Multiple system resources are considered	2010

Given the shortcomings of the above research on the dynamic threshold, and combined with the three-way decision, this paper proposes a virtual machine energy-saving merger method based on the three-way decision. This method divides the workload fluctuation into three parts based on the three-way decision and uses different mathematical models to predict the load, respectively [27,28]. Finally, the predicted load adjusts the double threshold dynamically, making the virtual machine trigger more accurate. This method not only optimizes the number of computing nodes, but also dramatically reduces the number of migrations to achieve the purpose of optimal energy consumption optimization.

3. Preliminaries

This section mainly introduces some basic models, such as the process of VM consolidation, prediction model, workload model, and three-way decision model.

3.1. Process of the Virtual Machine Dynamic Consolidation

Virtual machine dynamic consolidation mainly includes three processes: (1) trigger of VM migration timing; (2) select the virtual machine to be; and (3) placement of virtual machines. Figure 1 depicts the process of virtual machine dynamically consolidation, and PM represents the physical machine.

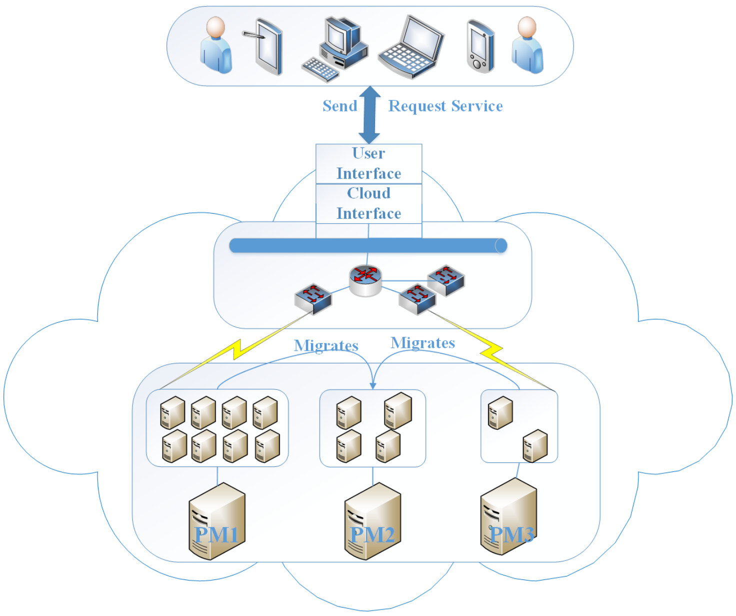

Figure 1. Virtual machine dynamic consolidation.

As shown in Figure 1, when a large number of user requirements are transmitted to the cloud, the cloud data center receives them through the interfaces and then allocates computing resources to users according to various decisions to meet the needs of users. However, due to users' complex and diverse requirements, VM migration is required to optimize resource allocation. The specific migration process includes the following three parts:

(1) Virtual machine migration trigger timing

Virtual machine migration trigger means that the system detects the resource utilization of the physical machine. It will trigger the migration if it finds that the resource utilization exceeds the set threshold. Virtual machine migration triggers are generally divided into two types, one is triggered by low load, and the high load triggers the other. The low load trigger is that when the load of the physical machine is lower than the lower threshold, all the virtual machines in the physical machine are migrated out, and the physical machine is converted into a sleep state. The high load trigger is when the workload is higher than the upper threshold, some suitable virtual machines are migrated out so that the workload status returns to normal, and the SLA quality of service is guaranteed. Static threshold triggered virtual machine migration cannot allocate computing resources according to the various load changes. In contrast, a dynamic threshold triggering virtual machine migration can dynamically adjust the threshold according to the workload, which can fully integrate computing resources and make more physical machines shut down to save energy consumption.

(2) Virtual machines selection

When it is detected that the resource utilization of a physical machine exceeds the predetermined threshold, the physical machine will trigger the virtual machine migration, which requires selecting some virtual machines for migration. There are many methods to select virtual machines, and different methods have different consideration angles. Including considering the cost of virtual machine migration and the minimum time of virtual machine migration;

(3) Virtual machines placement

The data center consists of hundreds or thousands of physical machines, and many virtual machines are waiting to be migrated simultaneously, so there are many ways to place virtual machines. Different placement methods also have different effects. It is necessary to use an appropriate way and method to achieve the best energy consumption optimization under the condition of ensuring QoS.

3.2. Prediction Model for Cloud Workload

(1) Exponential sliding model

$$s_i = \alpha x_i + (1-\alpha)s_{i-1}, 0 \le \alpha \le 1. \tag{1}$$

where, s_i is the smoothed value at time step i (understood as the ith time point), and x_i is the actual data at this time step. α can be any value between 0 and 1, which controls the balance between old and new information: when α approaches 1, only the current data point is retained; when α approaches 0, only the previous smoothing value is retained (the whole curve is flat). The recurrence relation is:

$$\begin{aligned} & s_i = \alpha x_i + (1-\alpha)s_{i-1} \\ & = \alpha x_i + (1-\alpha)[\alpha x_{i-1} + (1-\alpha)s_{i-2}] \\ & = \alpha x_i + (1-\alpha)[\alpha x_{i-1} + (1-\alpha)[\alpha x_{i-2} + (1-\alpha)s_{i-3}]] \\ & = \alpha[x_i + (1-\alpha)\alpha x_{i-1} + (1-\alpha)^2 x_{i-2} + (1-\alpha)^3 s_{i-3}] \\ & = \ldots \\ & = \alpha \sum_{j=0}^{i} (1-\alpha)^j x_{i-j} \end{aligned} \tag{2}$$

According to the recurrence relation, it can be seen that in the exponential smoothing method, all previous observations have an impact on the current smoothed value, but their role gradually decreases as the power of the parameter α increases. Those relatively early observations play a relatively minor role. The result of an exponential smoothing calculation can be extended beyond the dataset and range and thus can be used for forecasting. The prediction method is:

$$x_{i+h} = s_i \tag{3}$$

where, this s_i is the last value that has been calculated. When h is equal to 1, it represents the next value of the predicted.

(2) Holt's two-parameter exponential smoothing models

In Holt's two-parameter smoothing model, the prediction consists of two parts. One part is the horizontal part, which is updated by a simple exponential smoothing method based on the horizontal part of the previous period. The other part is the trend part, which is adjusted smoothly based on the trend part of the previous period and updated by a simple exponential smoothing method; the sum of the two will get the predicted value. The model is established as follows:

$$S_t = \alpha x_t + (1-\alpha)(S_{t-1} + b_{t-1}) \tag{4}$$

$$b_t = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1} \tag{5}$$

$$\hat{x}_{t+1} = S_t + b_t \tag{6}$$

where α and γ are smoothing coefficients, which can reflect the impact of recent data on the prediction results, x_t is the monitoring value of state data at time t, and $\hat{x}_{t+1}$ is the predicted value of state data at the next time. S_t is the smoothing value predicted at time t, reflecting the overall level of previous state data. b_t is the trend value, reflecting the changing state data trend. The state prediction model based on Holt Exponential Smoothing predicts the smooth value of state data in the sliding window and its trend value. Adding the trend value to the smooth value eliminates the disadvantage of prediction lag existing in other prediction models, and prediction accuracy is improved. Selecting optimal parameters by step acceleration method, the definition is as follows:

$$F = \alpha + n(\alpha - \gamma), n > 0 \tag{7}$$

where, n is the coefficient, this method mainly calculates the parameters of Holt's two-parameter exponential smoothing model to find the smoothing parameters α and γ that minimize the value F in the every step.

3.3. The Fluctuation Model of Workload

This subsection mainly introduces the workload fluctuation function, which can calculate the fluctuation of workload.

$$P_{load} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - x_{avg})^2} \tag{8}$$

where, n is the number of historical CPU utilization, x_i is the workload at the ith moment, and x_{avg} is the average historical load. The smaller the P_{load} value, the smaller the fluctuation of the workload, and on the contrary, the greater the fluctuation.

3.4. Three-Way Decision Model

The basic idea of the three-way decision is to divide the whole into three different regions according to a pair of thresholds (α, β) in the rough set, namely the positive region $POS(X)$, the negative region $NEG(X)$, and the boundary region $BND(X)$. Let OB be a set of objects, and $\Pr(X|[x])$ represents the conditional probability that an object in the equivalence class belongs to the set X. Then, we have:

$$\begin{cases} POS_{(\alpha,\beta)}(X) = \{x \in OB | \Pr(X|[x]) \geq \alpha\} \\ BND_{(\alpha,\beta)}(X) = \{x \in OB | \beta < \Pr(X|[x]) < \alpha\} \\ NEG_{(\alpha,\beta)}(X) = \{x \in OB | \Pr(X|[x]) \leq \beta\} \end{cases}. \tag{9}$$

The decision-making scheme is given as $D = \{a_P, a_B, a_N\}$, which represent that the decision-making is accepted immediately, the decision-making is delayed, and the decision-making is rejected, respectively. Let $\lambda_{ij}(i = P, B, N, j = P, N)$ represent the loss value for different actions, we can obtain the Formula (10) and Table 2.

$$\begin{aligned} R(a_P|[x]) &= \lambda_{PP}P(X|[x]) + \lambda_{PN}P(X^C|[x]) \\ R(a_B|[x]) &= \lambda_{BP}P(X|[x]) + \lambda_{BN}P(X^C|[x]) \\ R(a_N|[x]) &= \lambda_{NP}P(X|[x]) + \lambda_{NN}P(X^C|[x]) \end{aligned} \tag{10}$$

Table 2. Load period partition cost function.

	a_P	a_B	a_N
X	λ_{PP}	λ_{BP}	λ_{NP}
X^C	λ_{PN}	λ_{BN}	λ_{NN}

According to the Bayes minimization criterion, the optimal decision scheme is the action set with the minimum expected loss, and follows the following decision rules:

(P) If $R(a_P|[x]) \leq R(a_B|[x])$ and $R(a_P|[x]) \leq R(a_N|[x])$ are satisfied, it can obtain $x \in POS(X)$;

(B) If both $R(a_B|[x]) \leq R(a_P|[x])$ and $R(a_B|[x]) \leq R(a_N|[x])$ are satisfied, it can obtain $x \in BND(X)$;

(N) If both $R(a_N|[x]) \leq R(a_P|[x])$ and $R(a_N|[x]) \leq R(a_B|[x])$ are satisfied, it can obtain $x \in NEG(X)$;

The thresholds α and β are as follows:

$$\alpha = \frac{\lambda_{PN} - \lambda_{BN}}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}, \tag{11}$$

$$\beta = \frac{\lambda_{BN} - \lambda_{NN}}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}. \tag{12}$$

where, $\lambda_{ij}(i = P, B, N, j = P, N)$ represents the loss value for different actions a_P, a_B, a_N.

4. A Virtual Machine Energy-Saving Merging Method Based on Three-Way Decision

The specific process of the virtual machine energy-saving merging method based on the three-way decision is first to calculate the fluctuation of load history rate. Then, we divide the workload fluctuation into three parts based on the three-way decision and predict the workload using different mathematical models through different types of division. Finally, adjust the upper and lower thresholds dynamically through the predicted workload.

The block diagram of the steps of the proposed approach is shown in Figure 2. First, by analyzing the workload data of the cloud data center, the historical data fluctuations are divided into three parts, and then different prediction algorithms are used for each part to obtain the virtual machine trigger migration threshold. When the trigger condition of the VM migration is satisfied, VM integration is performed according to the predicted threshold to achieve the purpose of energy consumption optimization.

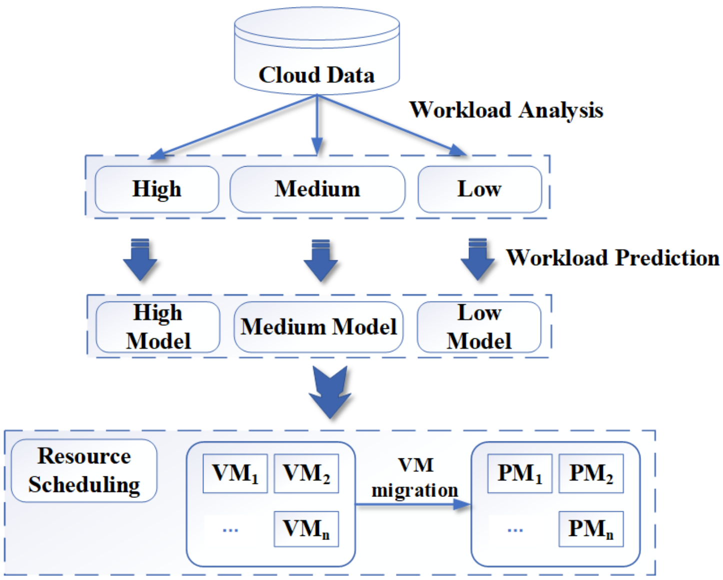

Figure 2. Block diagram of steps of the proposed approach.

The algorithm flow is shown in Figure 3, and this process mainly consists of two parts:

(1) Workload predictor

First, obtaining the historical workload usage from the cloud datacenter, and then calculating the fluctuation value δ by Formula (8). When the fluctuation value is low, the workload is relatively stable, and the change is small. When it is a high fluctuation value, the workload has changed significantly, but when it is not high or low, considering global changes.

(2) VM integration

When VM migration trigger conditions are met, the data center will give the decision to consolidate the VMs. The threshold is adjusted by predicting the workload, and the selected virtual machine is triggered to migrate, and then the virtual machine is placed in the suitable physical machine. The Algorithm 1 is inspired from [29] and described as follows:

Algorithm 1 Dynamic threshold adjustment algorithm.

Input:
Historical CPU utilization (CPU(x)) obtained from GoogleTrace (capacity:CPU(%)) = $\{x_1, x_2, x_3, \ldots, x_n\}$; The number of PM is m
Output:
UT , LT;
Calculate α and β from Equations (10) and (11);
Initialize UT $\leftarrow$ 70%, LT $\leftarrow$ 30%, Over_utilizedList, Low_utilizedList;
Put the load value less than 30% into Low_utilizedList;
Put the load value higher than 70% into Over_utilizedList;
if Over_utilizedList.Size < n/5 or Low_utilizedList.Size < n/5 **then**
  The threshold remains unchanged;
end if
Calculate δ from Equation (8);
if $\delta \leq \beta$ **then**
  Load predicting using exponential smoothing method;
  if | the predicted load $-$ the threshold | $\leq$ 5% **then**
    The threshold $\leftarrow$ the predicted load;
  else
    The threshold remains unchanged. The maximum threshold cannot exceed 90%, and the minimum cannot be lower than 10%.
  end if
end if
if $\delta \geq \alpha$ **then**
  Load predicting using Holt's double-parameter exponential smoothing method;
  if | the predicted load $-$ the threshold | $\leq$ 5% **then**
    The threshold $\leftarrow$ the predicted load;
  else
    The threshold remains unchanged. The maximum threshold cannot exceed 90%, and the minimum cannot be lower than 10%.
  end if
end if
if $\beta < \delta < \alpha$ **then**
  if (85% of the physical machine CPU usage exceeds the threshold) or (within 5% of the threshold) **then**
    Load predicting using exponential smoothing method;
    The threshold $\leftarrow$ the predicted load;
  else
    The threshold remains unchanged;
  end if
end if

There are n^2 historical loads in each physical machine of the algorithm, and the number of physical machines is $O(m)$. First, the time complexity of calculating the upper and lower thresholds is $O(n)$, and calculating the load fluctuation is also $O(n)$. The prediction of each physical host is $O(n)$. So, the time complexity of this algorithm is $O(m \times n^2)$.

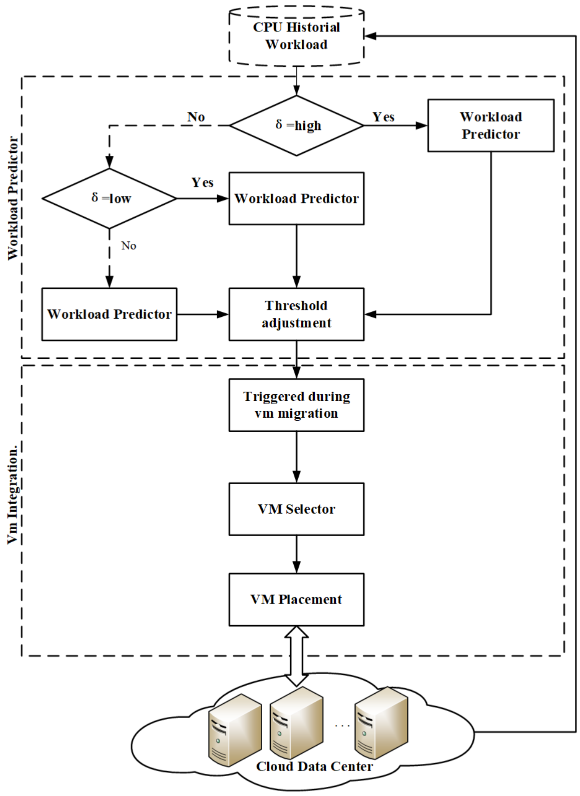

Figure 3. Energy saving combination method of VM based on workload prediction.

4.1. Virtual Machine Selection Strategy

MMT (minimum migration time) algorithm [29] is an effective single target virtual machine selection algorithm, which selects virtual machines based on the minimum migration time. Divide the amount of RAM used by the bandwidth required for the VM migration to get the migration time. The MMT formula is as follows:

$$v \in V_j \forall a \in V_j, \frac{RAM_u(v)}{NET_j} \leq \frac{RAM_u(a)}{NET_j} \tag{13}$$

where, $RAM_u(a)$ is the amount of ram utilized by virtual machine a, and NET_j is the bandwidth utilized by physical machine j.

4.2. Virtual Machine Placement Strategy

Best fit algorithm (best fit) [30], which always assigns the minimum free partition that meets the job's requirements. In order to speed up the search, the algorithm requires that all accessible areas are sorted by their size to form a blank chain in increasing order. This way, the first free area that satisfies the requirements found each time must be optimal, and the algorithm is very convenient.

5. Performance Analysis

This paper adopts the CloudSim Plus platform, which has more models and algorithms, which include heuristic algorithms, more VM migration algorithms, higher accuracy, scalability, and ease of use [31]. The algorithm in this paper compares the performance of the STA algorithm [19] and DTA algorithm [32] from five aspects: data center energy consumption, number of virtual machine migration, SLA violation rate, closing computer nodes, and virtual machine migration efficiency.The dataset uses Google clusters, which is a randomly-picked 1-s sample of CPU usage from within the associated 5-min usage-reporting period for that task, and using this data, it is possible to build up a stochastic model of task utilization over time for long-running tasks on a cluster of about 12.5 k machines [33]. The link to the dataset is https://github.com/google/cluster-data (access on 1 December 2021).

5.1. Experiment Setup

The experiment creates a simulated cloud datacenter on the cloudsim plus platform and constructs various resources such as physical machines and virtual machines of a specific scale. The experimental environment uses idea2021 2.2, jdk8, and cloudsim plus5 0.4. The symbols and definitions are shown in Table 3. The configuration of the physical machine and the virtual machine is shown in Table 4.

Table 3. Symbols and definitions.

Symbol	Definition
n	The number of historical CPU utilization
x	Workload value
m	The number of physical machine
α	Lower fluctuation
β	Upper fluctuation
δ	Workload fluctuation
UT	Upper threshold
LT	Lower threshold
VM	Virtual Machine
PM	Physical machine

Table 4. Configuration of PMs and VMs.

Parameters	Physical Machine	Virtual Machine
Memory	2 GB	1 GB
Bandwidth	100 Gbps	100 Mbps
Disk	1 TB	1 GB
CPU	1000 MIPs	1000 MIPs
PEs	4	1

5.2. Performance Metrics

(1) Number of virtual machine migrations: the number of virtual machines in the migration queue is NUM.

(2) Calculate the number of closed nodes: $B_c(i)$ is a binary decision variable indicating whether the physical machine i is powered on and working, i equal to 1 means that the physical machine i is in the active state,$B_c(i)$ equal to 0 means that the physical machine i is in the shutdown state, $PMList$ means the complete set of physical machines, $PMList'$ represents an active set of physical machines.

(3) Service Level Agreement Violation Rate (SLAViolations, SLAV) In order to test the performance of the algorithm, the default rate of SLA mainly depends on two aspects: one is the time that each active physical machine violates the SLA (SLATAH) [29], and the other is the performance degradation due to virtual machine migration (PDM).

SLATAH is defined as the percentage of the time that the active physical machine is under 100% CPU utilization and the service time:

$$SLATAH = \frac{1}{N}\sum_{i=1}^{N}\frac{T_{si}}{T_{ai}} \tag{14}$$

where, N is the number of active physical machines, T_{si} is the time of the i-th physical machine under 100% utilization of CPU; and T_{ai} represents the total time that the i-th physical machine provides services.

PDM is defined as the percentage of virtual machine performance degradation due to migration:

$$PDM = \frac{1}{M}\sum_{j=1}^{M}\frac{C_{dj}}{C_{rj}} \tag{15}$$

where, M represents the number of virtual machines. C_{dj} represents an approximation of the performance degradation caused by the migration of the jth virtual machine; C_{rj} represents the CPU performance required by the jth physical machine during operation.

Then, the SLA violation rate can be defined by *SLATAH* and *PDM* together as:

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

5.3. Experiment Results

This subsection includes two parts: the first part demonstrates the performance and hypothesis testing analysis of the cloud load prediction algorithm, and the second is the performance evaluation and analysis of the proposed virtual machine migration algorithm.

5.3.1. Prediction Analysis

This part presents the analysis of the proposed algorithm showing the prediction performance under various scenarios and hypothesis testing analysis. The prediction evaluation indicators under different scenes are shown in Table 5. The comparison of actual values and predicted values are shown in Figure 4.

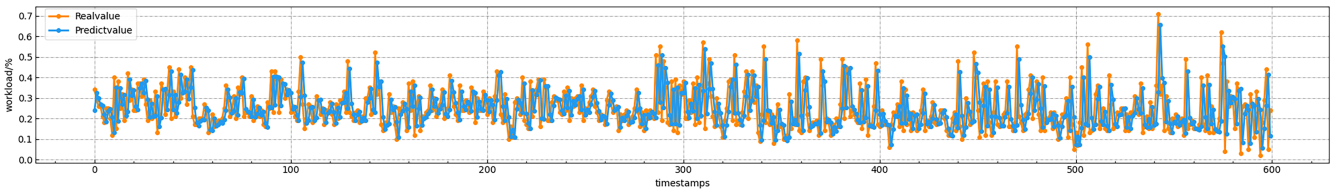

Figure 4. Comparison of actual values and predicted values.

Table 5. Prediction evaluation indicators under different scenes.

Numbers of Cloud Task	Evaluation Indicator			Numbers of Cloud Task	Evaluation Indicator		
	RMSE	**MAPE**	**MAE**		**RMSE**	**MAPE**	**MAE**
50	0.072	0.565	0.06	550	0.165	0.238	0.129
100	0.093	0.279	0.073	600	0.164	0.237	0.127
150	0.112	0.241	0.084	650	0.165	0.238	0.127
200	0.13	0.204	0.097	700	0.165	0.24	0.126
250	0.146	0.271	0.112	750	0.163	0.237	0.124
300	0.152	0.251	0.123	800	0.164	0.24	0.124
350	0.155	0.243	0.126	850	0.162	0.238	0.122
400	0.158	0.237	0.127	900	0.162	0.24	0.122
450	0.162	0.239	0.129	950	0.16	0.24	0.12
500	0.162	0.236	0.128	1000	0.159	0.238	0.118

5.3.2. Performance Analysis

Through cloud platform simulation experiments, the results will show in this subsection. The experiment is to compare the number of cloud tasks under different conditions. As seen from Figures 5 and Table 6, the proposed algorithm has fewer migrations times than the others, which are found via different mathematical models to obtain the double thresholds according to different load fluctuations, and the effect is noticeable.

Table 6. VM migration times.

Numbers of Cloud Task	Algorithm			Numbers of Cloud Task	Algorithm		
	DTA	STA	ESMM-3WD		DTA	STA	ESMM-3WD
50	23	15	12	550	265	195	145
100	45	33	26	600	275	216	162
150	75	55	42	650	295	228	172
200	110	85	64	700	310	244	184
250	150	111	84	750	330	258	194
300	190	132	98	800	350	271	204
350	210	113	114	850	375	281	212
400	215	162	122	900	390	297	212
450	225	166	125	950	405	312	235
500	245	180	136	1000	415	331	250

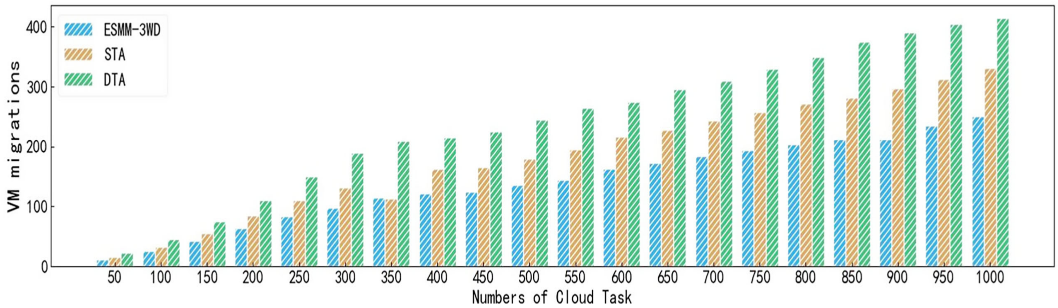

Figure 5. VM migration times.

Figure 6 and Table 7 show the comparison of the number of closed physical machines. The proposed algorithm is similar to the STA algorithm on the number of the closed physical machine. The STA algorithm is the best in static threshold situations. In comparing the number of closed physical machines, the performance of the STA algorithm is similar to the proposed algorithm.

Table 7. The number of closed PMs.

Numbers of Cloud Task	Algorithm			Numbers of Cloud Task	Algorithm		
	DTA	STA	ESMM-3WD		DTA	STA	ESMM-3WD
50	10	5	5	550	132	66	66
100	22	11	11	600	138	68	68
150	36	18	18	650	146	73	73
200	54	27	27	700	154	77	77
250	74	37	37	750	164	82	82
300	94	50	50	800	174	87	87
350	104	52	52	850	186	93	93
400	106	53	53	900	194	97	103
450	112	56	56	950	202	101	101
500	122	61	61	1000	206	103	103

Figure 7 and Table 8 show the comparison of SLAV. The DTA algorithm has certain fluctuations, ESMM-3WD and the STA algorithm is more effective in ensuring the SLA service quality and the SLAV of STA algorithm is similar to the ESMM-3WD algorithm.

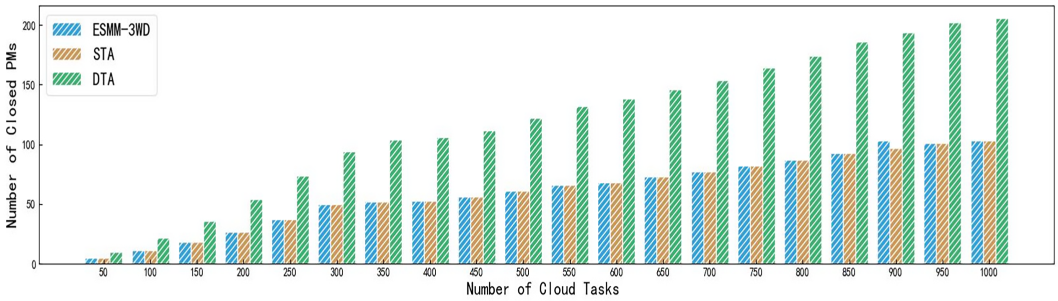

Figure 6. The number of closed PMs.

Table 8. Comparison results of SLAV.

Numbers of Cloud Task	Algorithm			Numbers of Cloud Task	Algorithm		
	DTA	STA	ESMM-3WD		DTA	STA	ESMM-3WD
50	0.33	0.06	0.05	550	0.91	0.06	0.07
100	0.44	0.06	0.06	600	0.58	0.06	0.07
150	0.54	0.06	0.06	650	0.65	0.06	0.07
200	0.54	0.06	0.06	700	0.57	0.06	0.06
250	0.70	0.06	0.06	750	0.62	0.06	0.06
300	0.88	0.06	0.06	800	0.64	0.06	0.06
350	0.98	0.06	0.06	850	0.77	0.06	0.06
400	0.69	0.06	0.06	900	0.71	0.06	0.06
450	0.80	0.06	0.07	950	0.67	0.06	0.07
500	0.90	0.06	0.07	1000	0.52	0.06	0.06

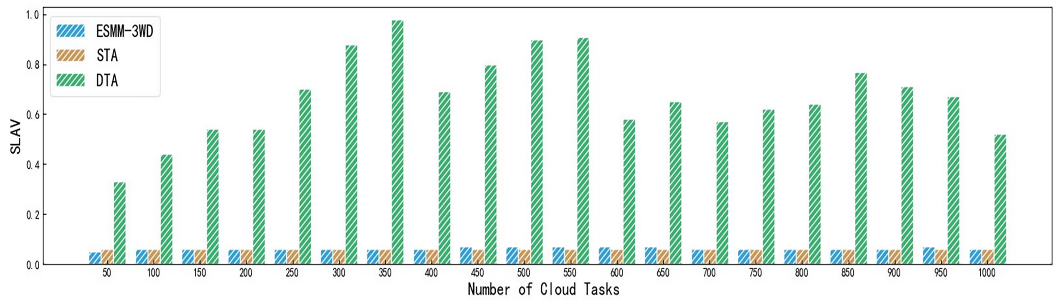

Figure 7. Comparison results of SLAV.

Figure 8 and Table 9 show the comparison of energy consumption. It can be seen from the figure that the energy consumption of ESMM-3WD algorithm is lower than that of the other two algorithms under various conditions of the number of virtual machines.

Table 9. Comparison of energy consumption.

Numbers of Cloud Task	Algorithm			Numbers of Cloud Task	Algorithm		
	DTA	STA	ESMM-3WD		DTA	STA	ESMM-3WD
50	283,798	238,931	235,523	550	3,400,377	3,010,653	2,746,698
100	594,089	526,726	501,743	600	3,654,381	3,330,206	3,036,504
150	959,948	849,061	783,741	650	3,917,211	3,537,579	3,226,389
200	1,427,668	1,296,176	1,182,015	700	4,149,745	3,777,007	3,448,343
250	1,904,193	1,697,087	1,547,134	750	4,404,565	3,983,717	3,625,197
300	2,377,261	2,005,019	1,819,867	800	4,653,212	4,202,847	3,827,010
350	2,627,820	2,309,121	2,091,915	850	4,897,536	4,356,494	3,993,306
400	2,758,826	2,471,964	2,254,208	900	5,137,882	4,615,868	4,004,202
450	2,901,305	2,554,697	2,361,617	950	5,378,853	4,859,284	4,448,659
500	3,146,223	2,781,776	2,548,082	1000	5,626,801	5,137,607	4,708,651

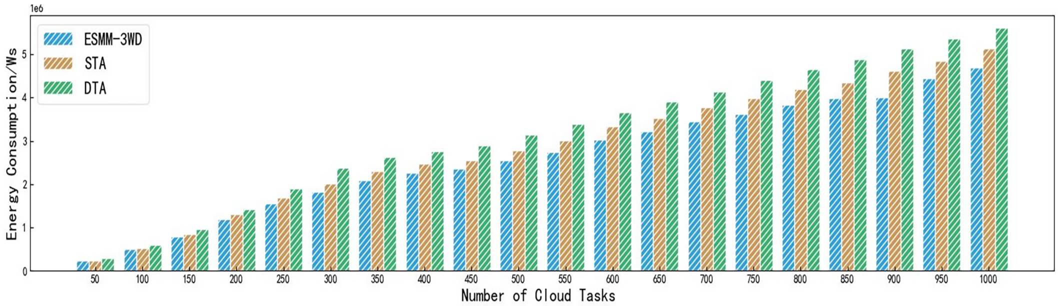

Figure 8. Comparison of energy consumption.

In order to test whether there are significant differences between different algorithms, we adopt Wilcoxon pairwise test to compare these experimental results. Given that the test threshold is 0.05, we can find that all the test P-values are smaller than the threshold in Table 10, where the units of p-values is $\times 10^{-4}$. Thus, we could consider that ESMM-3wd is significantly different from other comparison algorithms (i.e., DTA, STA) in most cases.

Table 10. The Wilcoxon test results and mean value of ESMM-3WD and other methods.

Methods	VM Migration Times	p-Value	Number of Closed PMs	p-Value	SLAV	p-Value	Energy Consumption	p-Value
DTA	244.90	0.8832	121.50	0.8857	0.67	0.8845	3,210,084.70	0.8857
STA	184.25	0.8857	60.85	0.8857	0.06	0.8820	2,877,091.00	0.8857
ESMM-3WD	139.65	-	61.15	-	0.06	-	2,619,540.20	-

Finally, the ESMM-3WD algorithm combines the theories of three-way decision, firstly dividing the workload fluctuation into three parts and then using different mathematical models to predict the load and dynamically adjust the threshold. The predicted load improves the integration probability of virtual machines to optimize energy consumption. From the effectiveness of the above indicators, the proposed algorithm has better performance than other algorithms in reducing the number of migrations and saving energy consumption, user service experience, and it is suitable for the dynamic and complex environment of cloud data centers.

6. Conclusions

At present, the energy consumption of cloud datacenters is increasing daily, but most of the waste of energy consumption is caused by the number of idle physical machines and the excessive migration of virtual machines. In order to solve the above problems, the proposed algorithm in this paper divides the load fluctuation into three categories: high, medium, and low based on the three-way decision and uses different mathematical models to predict the load according to different fluctuation categories. Finally, the predicted load value adjusts the double threshold dynamically for virtual machine integration and reducing energy consumption. In further research, we will focus on virtual machine selection and placement under the cloud environment and propose a more significant algorithm for energy savings.This paper does not consider the type of physical machines, such as compute-intensive, memory-intensive, or storage-intensive, when migrating virtual machines.

Author Contributions: Conceptualization, J.Y. and G.Z.; methodology, J.Y.; software, J.Y.; validation, J.Y., G.Z.; formal analysis, J.Y.; investigation, J.Y.; resources, J.Y.; data curation, J.Y.; writing—original draft preparation, J.Y.; writing—review and editing, J.Y. and G.Z.; visualization, J.Y.; supervision, J.Y.; project administration, J.Y. and G.Z.; funding acquisition, J.Y. and G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Special financial fund of Guangdong Province (2220062000038) and the National Natural Science Foundation of China (62003092).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. They are restricted to experimental results.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Armbrust, M.; Fox, A.; Griffith, R. *Above the Clouds: A Berkeley View of Cloud Computing*; Technical Report UCB/EECS-2009-28; EECS Department, University of California: Berkeley, CA, USA, 2009.
2. Foster, I.; Kesselman, C.; Tuecke, S. The anatomy of the grid: Enabling scalable virtual organizations. *Int. J. High Perform. Comput. Appl.* **2001**, *15*, 200–222. [CrossRef]
3. Mell, P.; Grance, T. *The NIST Definition of Cloud Computing*; Special Publication (NIST SP); National Institute of Standards and Technology: Gaithersburg, MD, USA, 2011; pp. 800–145.
4. Theja, P.R.; Babu, S.K. Evolutionary computing based on QoS oriented energy efficient VM consolidation scheme for large scale cloud data centers. *Cybern. Inf. Technol.* **2016**, *16*, 97–112. [CrossRef]
5. Armbrust, M.; Fox, A.; Griffith, R. A view of cloud computing. *Commun. ACM* **2010**, *53*, 50–58. [CrossRef]
6. Chu, Y.; Ma, T.; Zhao, L. Cloud Computing Resource Scheduling Policy and Algorithm. *Comput. Sci.* **2013**, *40*, 8–13.
7. Barroso, L.A.; Clidaras, J.; Hölzle, U. The datacenter as a computer: An introduction to the design of warehouse-scale machines. *Synth. Lect. Comput. Archit.* **2013**, *8*, 1–154.
8. Fan, X.; Weber, W.D.; Barroso, L.A. Power provisioning for a warehouse-sized computer. *ACM Sigarch Comput. Archit. News* **2007**, *35*, 13–23. [CrossRef]
9. Huang, J.; Wu, K.; Moh, M. Dynamic Virtual Machine migration algorithms using enhanced energy consumption model for green cloud data centers. In Proceedings of the International Conference on High Performance Computing & Simulation (HPCS), Bologna, Italy, 21–25 September 2014.
10. Yao, Y. Three-way decision: An interpretation of rules in rough set theory. In Proceedings of the 4th International Conference on Rough Sets and Know Ledge Technology (RSKT 2009), Gold Coast, Australia, 14–16 July 2009.
11. Yuan, K.; Xu, W.; Li, W.; Ding, W. An incremental learning mechanism for object classificationbased on progressive fuzzy three-way concept. *Inf. Sci.* **2022**, *584*, 127–147. [CrossRef]
12. Chen, X.; Xu, W. Double-quantitative multigranulation rough fuzzy set based on logical operations in multi-source decision systems. *Int. J. Mach. Learn. Cybern.* **2022**, *13*, 1021–1048. [CrossRef]
13. Li, W.; Xu, W.; Zhang, X.; Zhang, J. Updating approximations with dynamic objects based on local multigranulation rough sets in ordered information systems. *Artif. Intell. Rev.* **2022**, *55*, 1821–1855. [CrossRef]
14. Xu, W.; Yuan, K.; Li, W. Dynamic updating approximations of local generalized multigranulation neighborhood rough set. *Appl. Intell.* **2022**, *52*, 9148–9173. [CrossRef]
15. Arshad, U.; Aleem, M.; Srivastava, G.; Lin, J.C.W. Utilizing power consumption and SLA violations using dynamic VM consolidation in cloud data centers. *Renew. Sustain. Energy Rev.* **2022**, *67*, 112782. [CrossRef]
16. Bahrami, M.; Haghighat, A.; Gholipour, M. The Reduce energy consumption in cloud computing data centres by optimizing virtual machines. *Turk. J. Comput. Math. Educ.* **2021**, *12*, 5342–5350.
17. 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*, 31224–31335. [CrossRef]
18. Khan, M.A. An efficient energy-aware approach for dynamic VM consolidation on cloud platforms. *Cluster Comput.* **2021**, *21*, 1–18. [CrossRef]
19. Seth, S.; Singh, N. Dynamic threshold-based dynamic resource allocation using multiple VM migration for cloud computing systems. In Proceedings of the International Conference on Information, Communication and Computing Technology, New Delhi, India, 13 May 2017.
20. Yan, C.; Li, Z.; Yu, X. Virtual machine dynamic consolidation method based on adaptive overloaded threshold selection. *J. Comput. Appl.* **2016**, *36*, 2698–2703.
21. Deafallah, A.; Eidah, J.A.; Nasrin, S.; Zahir, T.; Albert, Y.Z. DTFA: A Dynamic Threshold-Based Fuzzy Approach for Power-Efficient VM Consolidation. In Proceedings of the 2018 IEEE 17th International Symposium on Network Computing and Applications (NCA), Cambridge, MA, USA, 1–3 November 2018.
22. Beloglazov, A.; Buyya, R. Adaptive tHreshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Datacenters. 2010. Available online: https://www.researchgate.net/publication/240319582_Adaptive_Threshold-Based_Approach_for_Energy-Efficient_Consolidation_of_Virtual_Machines_in_Cloud_Data_Centers (accessed on 10 August 2022).
23. Guo, D.; Jiang, C.; Wu, P. Three-way decision based on confidence level change in rough set. *Int. J. Approx. Reason.* **2022**, *143*, 57–77. [CrossRef]

24. Xu, W.; Yu, J. A novel approach to information fusion in multi-source datasets: A granular computing viewpoint. *Inf. Sci.* **2017**, *378*, 410–423. [CrossRef]
25. Guo, D.; Jiang, C.; Sheng, R. A novel outcome evaluation model of three-way decision: A change viewpoint. *Inf. Sci.* **2022**, *607*, 1089–1110. [CrossRef]
26. Xu, W.; Guo, Y. Generalized multigranulation double-quantitative decision-theoretic rough set. *Knowl.-Based Syst.* **2016**, *105*, 190–205. [CrossRef]
27. Xu, W.; Li, W. Granular computing approach to two-way learning based on formal concept analysis in fuzzy datasets. *IEEE Trans. Cybern.* **2016**, *46*, 366–379. [CrossRef]
28. Xu, W.; Yuan, K.; Li, W.; Ding, W. An emerging fuzzy feature selection method using composite entropy-based uncertainty measure and data distribution. *IEEE Trans. Emerg. Top. Comput. Intel.* **2022**, 1–13. [CrossRef]
29. Beloglazov, A.; Buyya, R. Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud datacenters. *Concurr. Comput. Pract. Exp.* **2012**, *24*, 1397–1420. [CrossRef]
30. Kou, L.; Markowsky, G. Multidimensional bin packing algorithms. *IBM J. Res. Dev.* **1977**, *21*, 443–448. [CrossRef]
31. Silva, F.; Oliveira, R.; Monteiro, C.; Inácio, P.; Freire, M. CloudSim Plus: A cloud computing simulation framework pursuing software engineering principles for improved modularity, extensibility and correctness. In Proceedings of the 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), Lisbon, Portugal, 8–12 May 2017.
32. Zhao, C.; Yao, L.; Cui, Y.; Xiang, H.; Feng, B. Dynamic adjusting threshold algorithm for virtual machine migration. *J. Comput. Appl.* **2017**, *37*, 2547–2550.
33. Dong, Z.; Zhuang, W.; Rojas-Cessa, R. Energy-aware scheduling schemes for cloud datacenters on google trace data. In Proceedings of the 2014 IEEE Online Conference on Green Communications (OnlineGreencomm), Online, 12–14 November 2014.