Balanced Allocation Method of Physical Education Distance Education Resources Based on Linear Prediction

Hui Ren1 and Mishal Sohail2

1Zhengzhou Health Vocational College, Zhengzhou 450000, China
2Department of Computer Science, Superior University Lahore, Pakistan

Correspondence should be addressed to Mishal Sohail; mishalsohail53@gmail.com

Received 20 April 2022; Revised 16 May 2022; Accepted 19 May 2022; Published 7 June 2022

1. Introduction

General trend and multiple polarisation are the new normal in the current distant education sector. The term “general tendency” refers to the fact that contemporary distant education is expected to play an increasingly important role in educational reform in the years to come. [1] The state uses a variety of media, including radio, television, the Internet, and other carriers, to actively promote the popularisation of higher academic continuing education, nonacademic continuing education, and the building of a learning society. Investment in education should continue to rise from an economic standpoint. Multimedia, digitalization, networking, and intelligence should be realised progressively at the technological level. We should progressively actualize the socialisation and globalisation of resources and carry out individual learning, collaborative learning, simulation practice, etc., from an educational viewpoint. The market for modern distance education has the potential to develop enormously due to the vast number of individuals interested in it. Contemporary distant education has grown well beyond the original TVU, aopeng, and 68 colleges and institutions because of the so-called “multipolarization.” An increasing number of firms and organisations are adopting the modern distance education sector, in addition to public service systems. This growing growth momentum has substantially extended and reinforced the social resources of contemporary distant education on the original premise. Distant learning is undergoing a rearrangement and restructuring in response to the new normal for contemporary remote education [2], which is one of survival of the fittest. When it comes to the market and society, colleges and educational institutions have to establish their own place. Two outcomes are possible: either progress and growth in the competition or elimination. This is an opportunity, a challenge, and a once-in-a-lifetime chance for growth for both schools and businesses. The current distance education market is maturing as a result of the new normal. Sports remote education has grown fast as a component of contemporary distance education because of the general
advancement of the modern distance education sector [3, 4]. At the same time, people in China are becoming increasingly aware of the importance of physical fitness and a healthy lifestyle as a result of the rapid growth of the country’s social economy, political landscape, scientific community, and sports sector. As a new kind of physical education, remote learning has given national sports and long-term sports new life. Sports information is accessible to anyone, and it is also convenient for most sports employees to continue their education [5]. The contemporary distance education sector now has one of the most promising new areas to explore: sports remote education. There is a lot of room for growth and expansion in the current distance education sector when it comes to sports. A balanced allocation method for physical education distance education resources is to use information technology and an Internet platform to further optimize physical education teaching resources, promote quality improvement, adapt to current educational trends, and adapt to education and teaching reform development needs. Theory and practice are deeply intertwined. It is not essential to add data algorithm modeling technology in order to more intuitively display, process, and analyze data and increase data storage and interaction capabilities when faced with a huge number of sports distance education resource data. Data may be compared using linear prediction technology, which is able to identify patterns and randomness. A new approach to physical education teaching and learning is proposed in this paper based on linear prediction. This paper proposes a balanced allocation method for distance education resources in physical education that can meet the needs of different levels and groups of students in physical education.

2. Balanced Allocation Method of Physical Education Distance Education Resources Based on Linear Prediction

2.1. Design Resource Monitoring and Configuration Architecture. Users use various terminals to send user tasks or service requests to the agent center of the sports distance education platform through the network. According to the real-time status information in the data center sent by the monitoring center, the scheduling center decides to select a computing center, execute the specified scheduling algorithm in the scheduling domain, and then allocate the corresponding resources. After successful scheduling, the status information of resources will be updated to the monitoring center. At this time, users can use their own sports distance education resources. At the same time, the monitoring center will obtain the status information of all nodes in the cluster in real time and transfer it to the dispatching center, and then the dispatching center will dynamically adjust the resource location through resource migration according to the optimization objective function and idle resource information in different computing centers so as to balance the load of the overall sports distance education platform. The resource management system primarily consists of two key modules: resource monitoring and resource scheduling. According to the kind of information or information threshold, the scheduling module passively or actively gets monitoring information from the monitoring module and dynamically conducts load balancing, maximum utilization, or performance optimization algorithms in all scheduling domains [6]. There are three factors causing the scheduling algorithm: the early warning signal of the monitoring module, the timing monitoring of the scheduling module, and the task request of new users. After the execution of the scheduling algorithm, a series of virtual machine operations (migration, creation, suspension, and modification of virtual machine configuration) will be formed. This series of operations will be sent to the virtual machine operation driver module in the form of operation command information to realize the deployment of the final adjustment. The virtual machine operation driver module encapsulates the call interfaces of different virtualization software. Its main function is to call the encapsulated interface after receiving various virtual machine operation requests sent by the scheduling module, classify and batch the list of requests for parallel execution, increase the efficiency of execution, and feedback the final execution results and status information to the scheduling framework and sports distance education platform [7]. The two modules complement each other. The monitoring module provides real-time status information, historical prediction information and corresponding alarm information of various resources to the scheduling module. The scheduling module selects a certain load balancing algorithm according to the corresponding resource state information to achieve the specified purpose in a specified scheduling domain, or makes reasonable resource allocation according to the tasks submitted by the user, so that the user’s task execution meets the specified SLA requirements. Finally, the virtual machine operates the driver to perform the final action, returns the corresponding status, and timely notifies the updated information of virtual resources to the monitoring module. The overall framework of resource monitoring and allocation proposed in this paper is shown in Figure 1.

The monitoring center node and the monitored node are the two most significant nodes in conventional monitoring. Push mode (the monitoring node transmits monitoring information to the monitoring center node on a regular basis), pull mode (the monitoring center node pulls monitoring information from the monitored node on a regular basis), or a mix of the two modes are the two basic data transfer modalities. These methods have their own advantages and disadvantages. The common and typical problem is that pulling the monitoring information regularly is a great burden on the overall network performance of the sports distance education platform. There are many monitoring nodes in the cloud platform, coupled with the monitoring of virtual machine resources, which will bring great pressure to the monitoring center nodes in the model of traditional single monitoring center and multiple monitored nodes, and even lose information, so as to make wrong configuration strategies. This paper proposes a stepped monitoring structure, as shown in Figure 2.
Each monitoring agent node is deployed on a specific physical machine node or virtual machine node that needs to be monitored. Each local monitoring center is an independent and autonomous domain, which can deal with the problems in the domain internally. Only a small amount of monitoring information needs to be transmitted to the general monitoring center. In this way, the overall network load information can be reduced, and the effect of fault isolation can also be achieved. Through the process of self-learning, the monitoring master node saves all results in the master node and provides various interfaces: resource use display (the interface provided to the web monitoring display end), allocation scheduling interface (for the dispatching center to pull the specified requirements to pull the real-time monitoring information), early warning alarm interface (the interface that makes an alarm according to the prediction result or the value of the monitoring result exceeds the set threshold), the resource billing interface (called by monitoring the user’s flow and the service worker billing module used), and the monitoring prediction configuration interface (the maintenance interface that dynamically sets the learning parameters and threshold information, which is used internally). The monitoring agent node only needs to deploy the basic monitoring program and provide the monitoring data call interface and monitoring mode configuration interface to the local monitoring center. The local monitoring agent center deploys the adaptive manager and realizes the prediction mechanism to dynamically transmit control information to the collector and obtain resource monitoring information.

2.2. Estimation of Resource Demand Based on Linear Prediction. The sports distance education platform realizes flexible and scalable resource sharing services through resource management. Resource monitoring and prediction is the basis for realizing resource automation and high-performance management in the network environment. This chapter studies the resource monitoring and prediction under the network environment. According to the correlation between sports distance education resources, a resource prediction mechanism based on linear prediction is proposed to estimate the resource load and predict the resource use more effectively. The linear prediction model is used to highlight the intricate link between many internal dependant variables in the interaction between various system elements [8]. The model is solved using multivariable and multiple equations joint regression, which means that each endogenous variable regresses all of the other endogenous variables’ lag term equations. This solves the problem of comprehensive modeling of the lag value of each internal dependent variable to all other internal dependent variables in the system in structural modeling so as to obtain the estimated dynamic relationship of all internal dependent
variables [9]. The model is mainly used to predict and analyze the interconnected time series model and analyze the impact of nonfixed interference term on the system variable model [10]. The main principle of the linear prediction model is to list the functions of the lag values of all endogenous variables in the model relative to all variables except their own variables so as to transform the traditional univariate autoregressive model into a multivariable autoregressive model. Through this multivariable autoregressive model, the accuracy of short-term prediction and the data trend in the long term can be improved [11]. Let \( x_t = (x_{1t}, x_{2t}, \ldots, x_{nt}) \) be the basic variable of a \( 1 \times n \)-dimensional time series, then the linear prediction model is defined as follows:

\[
x = a + \prod_{1} x_{t-1} + \prod_{2} x_{t-2} + \cdots + \prod_{b} x_{t-b} + \delta_t.
\]  

(1)

In formula (1), \( x \) represents a linear predictive variable; \( a \) is the constant; \( \prod_{1} x_{t-b} \) represents the parameter matrix of \( n \times n; t \) represents time; \( b \) represents order; \( \delta_t \) represents the anti-interference parameter vector and has nothing to do with the time state. All variables of the linear prediction model are internal dependent variables. In addition, unidirectional causality variables and trend variables determined by external factors and external dependent variables can be added to the model as external dependent variables. Therefore, the following formula is provided:

\[
x = \prod_{1} x_{t-1} + \prod_{2} x_{t-2} + \cdots + \prod_{b} x_{t-b} + D_t + G_t + \delta_t.
\]  

(2)

In formula (2), \( D_t \) represents an exogenous variable and \( G_t \) represents a deterministic variable. Exogenous and deterministic variables may each have their own lag factor. The equation representing the link between numerous endogenous variables is reflected in the linear prediction model built in advance, and the most recent monitoring data is replaced into the equation to forecast monitoring data in the following cycle. According to the weight relationship between various variables, a comprehensive value is obtained to check whether the value reaches the specified threshold so as to determine whether to carry out new monitoring and the pull cycle of monitoring information [12]. Normalize and average the sample data and save it to an intermediate array. Each time point is processed in this way (the logarithm is mainly used to eliminate heteroscedasticity). Finally, the four numbers obtained through mean processing at each time point are assigned to the linear prediction model as the final input sample. The AIC information criterion is selected to determine the lag order, and the corresponding value is obtained through the following formula:

\[
\text{AIC} = -2 \left( \frac{\log \lambda}{T} \right) + \frac{2b}{T}
\]  

(3)

In formula (3), AIC represents the AIC value, \( \lambda \) means likelihood estimation, and \( T \) represents the time period. Through the sample, the above estimation parameter matrix will be predicted to obtain the linear model. By substituting the latest data of the sample into the equation, the predicted value at the next time point can be obtained. According to the importance of each parameter, a comprehensive value can be obtained by using the weight. The CPU utilization is relatively important, and the weight setting should be high to obtain the comprehensive value. If the comprehensive value reaches the specified threshold, the monitoring module actively sends a monitoring command to the collectors of all servers, clears to start timing, and moves to the next step. Thus, the resource demand of physical distance education is estimated.

2.3. Establish Educational Resource Matching Model. Facing the resource demand of sports distance education, the resource allocation algorithm should allocate resources fastest and optimally. In the heterogeneous cloud computing environment of the sports distance education platform, it is necessary to match the performance requirements of the virtual machine with the performance of a physical server and select the appropriate physical server to deploy the virtual machine. Therefore, the performance matching distance between virtual resources and physical resources is established based on the virtual machine and server performance vector. The smaller the matching distance is, the better the performance of physical server resources matches the demand performance of the virtual machine. Because the virtual machine and the physical server are heterogeneous, the virtual machine has varying needs for various resources when it is formed on the physical server, resulting in a shifting percentage of various resources in the server [13, 14]. The server can no longer deploy new virtual machines when a given kind of server resource has been used, even if there are still a significant number of other types of server resources available. This results in resource fragmentation and waste. Because of memory resource depletion, the server will be unable to establish a new virtual machine, wasting CPU resources. In order to reduce resource fragmentation, it is necessary to consider the proportion between the requirements of various types of resources of the virtual machine and the proportion between the remaining types of resources of the server, and deploy the virtual machine to the server close to the two, so as to balance the changes of various types of resources of the server, so as to reduce the probability of resource fragmentation. The problem of virtual machine placement is transformed into looking for servers with the same or similar proportion of remaining resources as the resource required by the virtual machine, that is, the ratio of CPU, memory, and hard disk requested by the virtual machine is close to the ratio of various remaining resources of the server [15]. The closer the two are, the fewer the remaining resource fragments and the higher the resource utilization rate. Therefore, a resource matching distance model is established:

\[
d_{ij} = \sqrt{\sum_{z \in \{c,m,v\}} \left( \frac{y_{pz} \phi_{iz}}{y_{iz} - \phi_{iz}} \right)^2}.
\]  

(4)
In formula (4), $d_{ij}$ represents the resource matching distance; $i, j$ represent two resources; $z$ represents the virtual machine; $\{c, m, v\}$ indicate the amount of CPU, memory, and hard disk resources applied; $\gamma$ represents the ratio of various remaining resources of the server; $\varphi$ indicates that the virtual machine requests the ratio of CPU, memory, and hard disk. The smaller the resource matching distance, the more balanced the server resource utilization. At the same time, in order to make full use of various types of resources and minimize the fragments of unavailable resources, if the remaining amount of some resources of the server is not enough to create any type of new virtual machine, the remaining rate of other types of resources of the server shall be less than a preset minimum [16]. In order to ensure the timeliness of resource provision, the balanced allocation method of sports distance education resources should actively predict and configure a certain number of virtual machines in advance to deal with the sudden growth of resource demand in the future. The resource allocation framework constructed in this paper is shown in Figure 3.

Assuming that the number of virtual machine requests at a certain time is $P(t)$ and the request queue is $x_i = (x_{1i}, x_{2i}, \ldots, x_{mi})$, the demand for main types of virtual machines at the next time is predicted by using the short-term resource demand prediction algorithm based on the linear prediction proposed above. The calculation formula is

$$Q(t + \tau) = M_i(t + \tau) + \cdot \cdot \cdot + M_j(t + \tau) + M_j(t + \tau). \quad (5)$$

In formula (5), $Q(t + \tau)$ represents the resource demand, $t + \tau$ indicates the next time, $M_i$ represents the demand for main types of virtual machines, $s$ represents the total quantity, and $i$ indicates the serial number. At this time, the total resource allocation of the sports distance education platform is $U(t)$, which should be the sum of the current virtual machine requests and the actively predicted virtual machine demand, that is,

$$U(t) = P(t) + Q(t + \tau)\delta\eta. \quad (6)$$

In formula (6), $\eta$ represents the proportion of the number of virtual machines configured in advance to the total predicted number. For example, 30% of the predicted number of virtual machines can be configured in advance to prevent excessive number of virtual machines configured in advance, resulting in high load pressure or waste of resources. $\varphi$ indicates the adjustment parameter. If $Q(t + \tau)$ is higher than a certain threshold number, you need to actively configure the virtual machine in advance, then $\varphi = 1$. If $Q(t + \tau)$ is lower than a certain threshold, there is no need to configure the virtual machine in advance, then $\varphi = 0$. After the number of actively predicted virtual machines is determined, the predicted virtual machine request queue is constructed for different types of virtual machines according to the demand. Predict the virtual machine request queue. Finally, the virtual machine types with the same demand can be arranged randomly. The demand urgency of the predicted virtual machines is set to the lowest, the user priority is set to the lowest, and the request order of the same type of virtual machines can be arranged randomly.

2.4. Design the Balanced Allocation Algorithm of Sports Distance Education Resources. The data center of sports distance education has a large number of heterogeneous servers, providing a large number of virtual machines for external resource services. In the face of sudden and urgent resource demand, a good resource allocation algorithm should allocate resources fastest and optimally, actively predict future resource demand, configure a certain number of virtual machines in advance to deal with the sudden nature of future resource demand, set priority allocation principles to deal with emergency resource demand, and establish resource performance matching to ensure the optimization of resource allocation [17, 18]. At the same time, the goal is to occupy the least number of servers, improve resource utilization and balance the use of various resources. The virtual machine is matched to the proper physical server during the resource allocation procedure. If a virtual machine is allocated to a physical server, the mapping element should be 1; otherwise, it should be 0. Assuming that the problem’s solution is a two-dimensional matrix, the mapping element should be 1. The matrix depicts the mapping relationship between the virtual machine request queue and the platform physical server in this manner. The total number of servers, the total distance of resource performance matching, and the total distance of resource matching are the optimization goals of the balanced allocation of physical education distance education resources. Therefore, an optimized resource balanced allocation model aiming at the above results is established. The set constraints are as follows: (1) the total demand of virtual machines deployed to a server for various types of resources is less than its idle resources [19]. (2) If the remaining resources of a server are insufficient to create a new virtual machine, that is, the remaining resources of this type of server are less than the required amount of all types of virtual machines, and the remaining rate of other types of resources should be less than the predetermined threshold [20]. Therefore, the problem of balanced allocation of sports distance education resources is transformed into the solution of the multiobjective optimization mathematical model. This problem is an NP-hard problem, which needs to be solved by the multiobjective evolutionary algorithm to obtain a set of Pareto optimal solutions. One of the optimal solutions is selected as a mapping solution from virtual machine queue to the physical server group. The number of individuals in the population is large, and the calculation time of individual
fitness function is too long. We use the objective function as the fitness function, and use multicore processors and multithreads to calculate and evaluate the fitness value in parallel so as to speed up the convergence speed of the algorithm. For the merged parent and child populations, the nondominated sets of individuals are sorted. For each nondominated set, the duplicate is removed first, the solutions with the same objective function and variable values are found, and one individual is retained. Then, the Euclidean distance between two adjacent individuals is calculated to judge whether the distance is less than the threshold. If it is less than the threshold, it is optimized according to the strategy. Calculation of the distance between nondominated individuals \( u \) and \( v \) is as follows:

\[
\text{dr}(u, v) = \sqrt{\sum_{w} \left( f(u) - f(v) \right)^2}.
\]

(7)

In formula (7), \( \text{dr}(u, v) \) represents the distance between nondominated individuals \( u \) and \( v \); \( w \) represents the number of individuals; \( f \) represents Euclidean distance. Calculate the threshold according to the maximum distance between two individuals in the current nondominated set:

\[
\sigma = \frac{\text{dr}(u, v)_{\text{max}}}{2R}.
\]

(8)

In formula (8), \( \sigma \) represents the threshold and \( R \) indicates the size of the population. After the weight removal is completed, the adjacent individuals are optimized. The preferred schematic diagram of adjacent individuals is shown in Figure 4.

As shown in Figure 4, if the distance between \( u \) and \( v \) is less than the threshold, find out the position of the center point \( o \) of \( g \) and \( h \) individuals adjacent to them, calculate the Euclidean distance between \( u \), \( v \), and \( o \), and compare the one closer to the center point. In Figure 4, the distance between two adjacent individuals \( u \) and \( v \) is less than the threshold, and individual \( v \) is close to the center point \( o \) so individual \( v \) is retained and \( u \) is eliminated. The steps of the balanced allocation algorithm of physical education distance education resources are as follows: establish a multiobjective optimal resource allocation model and set the initial parameters of the algorithm. Merge the parent and offspring populations, calculate the individual fitness value, and rank them. Delete duplicate individuals on each nondominated set and optimize adjacent individuals. Nondominated individuals are sorted from large to small according to the crowding distance and selected to form a new parent population. Select, cross and mutate the new parent population to generate a new offspring population. Judge whether the termination conditions are met. If so, output the Pareto optimal solution set of resource allocation. So far, the design of the balanced allocation method of physical education distance education resources based on linear prediction has been completed.

3. Experiment

3.1. Experimental Preparation. The goal of this experiment is to see whether the balanced allocation approach of physical education and remote education resources based on linear prediction is useful and performs well. If the experiment is conducted in a large-scale school data center, there may be certain unknown hazards, resulting in the insecurity of school distant education platform resource services and significant financial losses to schools and platform operators. As a result, we use simulation experiments to evaluate the method’s effectiveness under various sizes of virtual machine requests and physical servers, and we build a simulation experiment environment using Clouds cloud computing simulation software and the jmetal multiobjective algorithm framework. The Alibaba cloud public data collection was utilised in this experiment. Currently, the number of the virtual machine requesting users is set to 5. The number of virtual machines requested by each user is generated randomly, with a total of 410. The virtual machine type is the actual request type of the Alibaba cloud. The user priority is randomly generated within the range of \([1, 5]\), the user’s current virtual machine request emergency level is randomly generated within the range of \([1, 5]\), the predicted management user to which the virtual machine belongs is set to the lowest level 0, the predicted virtual machine emergency level is set to the lowest level 0, and the threshold of each type of resource fragment residual rate is set to 5%. Build a cloud data center with the CloudSim cloud computing simulation software and run 400 physical servers, involving four server types, as shown in Table 1.

Different types of resource ratios have different dimensions, and the hard disk resource ratio is quite different from the memory and CPU. In order to prevent the resource ratio distance from being dominated by the hard disk, the hard disk resource data is divided by 10.

3.2. Experimental Results and Analysis. In order to verify the superiority of this method, this method is compared with the balanced allocation method of sports distance education resources based on the BF algorithm and RR algorithm. Firstly, evaluate the advantages and disadvantages of different algorithms in resource performance matching. Under the condition of the different number of virtual machines, calculate the resource performance matching distance. The results are shown in Tables 2–4.
According to the findings in Tables 2–4, the resource performance matching distance of the three resource balanced allocation techniques increases as the number of virtual machines increases. The resource performance matching distance of this technique is much less than the balanced allocation method of sports distance education resources based on the BF algorithm and RR algorithm under the same number of virtual machines. Taking the test with 1000 virtual machines as an example, the average resource performance matching distance of this method is 765, which is 284 and 465 lower than that based on the BF algorithm and RR algorithm. It can be seen from the above results that the RR algorithm adopts the polling sequential allocation method and does not consider any resource matching problem, so the distance value is the largest. The BF algorithm considers placing the virtual machine on the server with the smallest remaining CPU resources, and its distance value decreases slightly. This method aims at the minimum number of servers, the minimum resource performance matching distance, and the minimum resource matching distance, so the matching distance is smaller than the two comparison methods. Therefore, this method has more advantages in resource performance matching and reducing resource fragments than the two comparison methods.

4. Conclusion

Due to the limitation of research time and research environment, the research on the balanced allocation method of sports distance education resources proposed in this paper is not perfect and needs to be improved. The linear prediction model can effectively solve the problem of prediction accuracy, but there is a problem. Because the prediction effect of multiple factors on a single factor is considered, the time complexity will be improved with the increase of the number of factors. One of the main directions of our future work will be to determine how to optimize performance on this basis, select several other variable factors that have a greater impact on a single variable factor based on the specific scenario,
remove the variable factors with less impact, and reduce the overall number of variable factors. Not only does this minimize the complexity of algorithm execution, but it also enhances forecast accuracy. As cloud computing applications become more and more complex and resources become more and more heterogeneous, there are more and more factors causing the change in resource demand, which affect the accuracy of prediction. Therefore, in the next step, we will study more factors affecting load prediction, establish more effective prediction strategies, and propose more effective methods to predict server load.

Data Availability
The data used to support the findings of the study are available on request.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

References