

Research Article

The Model Framework of College Students' Entrepreneurial Team Formation Based on Relevance Algorithm

Guisheng Hu¹ and Xiuqing Li² 

¹Student Affairs Office, Xi'an Siyuan University, Xi'an, Shannxi 710038, China

²School of Business, Xi'an University of Finance and Economics, Xi'an, Shannxi 710100, China

Correspondence should be addressed to Xiuqing Li; 2021010027@xaufe.edu.cn

Received 10 March 2022; Revised 12 April 2022; Accepted 19 April 2022; Published 10 May 2022

Academic Editor: Sheng Bin

Copyright © 2022 Guisheng Hu and Xiuqing Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Reasonable realization of the task relevance of the entrepreneurial team formation model is a research hotspot in the field of cluster computing. Based on the theoretical basis of the correlation algorithm, this article constructs a parallel computing framework for college students' entrepreneurial formation model and realizes the application of the team formation model in the entrepreneurial model. This article replaces the traditional structure with the multi-entrepreneurial formation mode structure to avoid premature convergence and initializes the entrepreneurial formation mode with an algorithm, which solves the problem of the first-come, first-served entrepreneurial association degree algorithm of the entrepreneurial formation mode. During the simulation process, the multilevel entrepreneurial relevance algorithm based on the relevance algorithm was experimentally verified and performance tested. In the performance comparison test, the optimal solution search efficiency increased to 89.7%, the heterogeneous fault tolerance rate of resources reached 0.713, and the task set scheduling time was reduced to 0.221s. The experimental results show that the multilevel entrepreneurial relevance model based on the relevance algorithm can effectively reduce the task set scheduling time and relevance cost, and can take into account the load balance of the system, effectively promoting the diversity and dynamics of resource requirements.

1. Introduction

With the increasing proportion of college students' admissions year by year, college students' entrepreneurship has developed rapidly. The expansion of enrollment in colleges and universities meets the requirements of the transformation, upgrading, and development of the socialist market economy, effectively improving the national cultural literacy, delaying the time for initial employment, and relieving the employment pressure of the whole society, but at the same time, it has also caused a series of social problems [1]. This is mainly reflected in the serious disconnect between the goals of talent training in colleges and universities and social needs, which greatly reduces the employment quality of college students, with low initial employment, low employment level, and poor employment stability, and also highlights the social employment problem dominated by

structural employment contradictions. College students are unwilling to choose grassroots and central and western regions for employment and entrepreneurship, unwilling to go to rural areas for employment and entrepreneurship, and unwilling to go to regions and jobs with difficult conditions to train and grow, which is a real problem before us. Practice shows that the level of college students' entrepreneurship affects the level and quality of college students' employment and entrepreneurship, and they are two issues that are endogenously related [2–4].

As a kind of aggregation of scattered computing resources, the entrepreneurial establishment mode is to integrate different computers across regions into a virtual high-performance computer and then provide this high-performance computing power user. In the entrepreneurial establishment mode, multiple scattered computing nodes form a huge grid and then make full use of idle computing

resources to divide computing tasks into a limited number of “small pieces” and deliver them to each node in the grid. Each node chooses to process one or more “small pieces” according to its own processing capability. In this article, based on the theory-constructed case study method, we first use the method of theoretical sampling to select four typical college students’ entrepreneurial teams newly created Internet entrepreneurial formation models as cases [5–7]. We then use primary data collection methods such as semi-structured interviews and secondary data collection methods such as internal documents, journals, and magazines to collect and analyze data. The key factors and variable indicators in the strategic selection and growth process of the entrepreneurial formation mode of each case are obtained by analyzing the collected data by categorizing and coding. Then, through the intracase analysis, each case is regarded as an independent individual, so that the unique mode of each case emerges, paving the way for the subsequent cross-case analysis. In this article, we analyze four typical cases of college students’ entrepreneurial teams’ new Internet entrepreneurial formation model and then find out the strategic choice and growth model of each case, providing a data basis for the subsequent cross-case analysis of the strategic choice and growth model [8–11].

Aiming at the problem of real-time relevance of independent tasks under limited resources, this article proposes a real-time task relevance method based on the deep relevance algorithm. Starting from the cost-effectiveness of improving the task relevance, this article proposes a cost-effective improved relevance task relevance algorithm and converts the relevance of the relevance task to the processing of large-scale graph data. In order to explore more high-quality solution sets that may be ignored by the deterministic algorithm, the algorithm adopts the multi-entrepreneurial formation mode architecture to expand the search range of the optimal solution and designs a fitness function based on the relevance time and relevance cost of the task set. In a dynamic and uncertain cloud environment, how to automatically and intelligently allocate user task requests that continue to arrive at the entrepreneurial establishment mode application to appropriate resources for execution is a difficult online scheduling problem. To solve this problem, this article proposes an intelligent QoS-aware real-time task scheduling algorithm based on the deep correlation algorithm. By dynamically analyzing the characteristics of user tasks and the load status of resources, the algorithm can allocate user tasks to the best resources for execution, which not only improves user satisfaction but also enables balanced and efficient use of leased virtual machines. In the experiment, compared with the other five task correlation algorithms, this correlation method shows significant advantages in terms of average task response time, resource utilization, task success rate, etc. and can adapt to different application load scenarios.

2. Related Work

Aiming at the problem of combination resource selection and relevance degree of related tasks, from the two aspects of

supporting task parallelization and adapting to resource dynamics, based on the architecture and relevance degree algorithm, two efficient relevance degree scheduling algorithms are proposed. For the provider of the entrepreneurial formation model, the goal of the relevance degree is to select an appropriate IaaS resource combination scheme for a given relevance degree consisting of multiple related tasks and make the virtual machine rental costs low. Although the problem of correlation degree has always been a hot issue in the field of cloud computing, most studies default that each task in the correlation degree can only be completed independently by one virtual machine, and each task is executed on different types of virtual machines. Time is known and determined. In fact, this assumption is not reasonable [12–14].

In addition to the traditional SWOT analysis and Porter’s Five Forces analysis, Hu [15] proposed the customer value innovation strategy and studied the difference between this new strategy and the competitive strategy and how to implement it in the Internet entrepreneurial formation model, and then put forward the customer-oriented standard competition strategy, a new Internet entrepreneurial formation mode strategy, and conducted in-depth research on its following main strategies-competitiveness-oriented management strategy, alliance strategy in standard competition, and product mainstreaming strategy. When analyzing the strategic management of the Internet entrepreneurial establishment model, it is found that in the era of big data, “data” are the core advantage and capital of Internet companies. Its strategic goals should focus on the acquisition of “data” resources rather than short-term profitability goals. Tawafak [16] believes that the Internet entrepreneurial formation model needs to learn the ability to quickly adjust strategies in practice to cope with changes in the market and competitive environment, because the adjustment cycle of the Internet entrepreneurial formation model is much shorter than that of traditional industries. In addition, for the characteristics of the Internet entrepreneurial formation model that is different from the traditional entrepreneurial formation model, starting from the analysis of the characteristics of the two-sided market of Internet enterprises, it is pointed out that the establishment mode of Internet entrepreneurship not only has the two-sided market structure of platform and bilateral users, but also has the characteristics of synergy between basic platforms and value-added services. From the perspective of the industrial structure of the Internet industry, it is pointed out that the market concentration of service providers in my country’s Internet industry is very high and the product differentiation is small. Therefore, the key to the success of the Internet entrepreneurial establishment model is the rapid reflection and learning of the market, and it is necessary to strengthen service levels.

On the basis of the entrepreneurial formation model, Tittel [17] proposed an improved quadratic correlation degree algorithm. The algorithm first pre-allocates tasks through the entrepreneurial formation model and then performs secondary correlation degree on this basis, which effectively reduces the correlation. However, when there are too many small tasks in the task set, the effect degenerates to

the Max-Min algorithm, and the universality of the algorithm needs to be improved. Moscoso [18] found that the advantages of the two algorithms of the comprehensive entrepreneurial formation model and Max-Min are dynamically selected according to the standard deviation of the expected completion time of each task, which is more scalable than other algorithms in the scheduling process. The entrepreneurial formation mode algorithm is a widely used algorithm with low time complexity and high efficiency. The main idea of the algorithm is to always select the smallest task in the task set and assign it in each task allocation process to the machine with the strongest processing power to obtain a shorter computing time. The disadvantage is that it is likely to cause the entrepreneurial formation mode with relatively slow processing speed to remain idle, while the entrepreneurial formation mode with stronger processing power is always busy. Moscoso [19] analyzed that Max-Min is similar to the entrepreneurial formation model in terms of implementation principle, but it prioritizes the tasks with the largest task concentration to the entrepreneurial formation model with the strongest computing power, which will also cause the load in the server cluster. Most of the above heuristic-based task relatedness algorithms can achieve better relatedness performance, so they are widely used in independent task relatedness problems, but most of them have the problem of complicated implementation process. Compared with the GA and ant algorithms, the PSO algorithm has simpler rules and easier implementation in the iterative process of nodes, so it is widely used in the heuristic-based task correlation problem, but it is affected by the neighborhood position and the global maximum. Due to the influence of the optimal location, it is easy to fall into the local optimal solution [20–25].

3. Hierarchical Construction Mode Based on Correlation Degree Algorithm

3.1. Relevance Decision Hierarchy. Context-level tasks refer to tasks that have data transfer or logical dependencies with other tasks. The execution of associated tasks has a sequence, and each task needs to wait for the tasks of all its predecessor nodes to be executed before it can start executing. The related degree of association problem is called the degree of association degree. The main research is on how to treat multiple related tasks with time-series relationship as a whole to carry out the degree of association, select appropriate resources for each related task, and ensure the task. On the premise of the constraints between the two, it can achieve the goals of reducing the completion time of the entire association degree and reducing the cost of resource rental:

$$f(s, t, r) - f(s, t) \times f(s, r) = Q(s, t, r) - Q(s, t), \quad (1)$$

$$\{s, t, r \in \text{Real}(c)\}.$$

During a superstep at a certain level, if the value of a vertex does not change, the vertex will be marked as inactive. Vertices in the inactive state will be re-marked as active after receiving the message. Pregel does not call its user-defined

functions for vertices that are in the inactive state. Therefore, the entire computation process ends when all vertices are in the inactive state and the message queue is empty. The static task correlation degree (i.e., the compilation correlation degree) obtains some prior information (such as the dependencies between tasks and the operation time of tasks in each entrepreneurial formation mode) through the application program before the system is compiled. Relevance strategy. The dynamic task correlation degree algorithm mainly occurs in the compilation process of the system. With the running of the program, the system can dynamically add new tasks and dynamically allocate the newly added task nodes according to the current operating conditions of the system. Most of the application services in cloud computing are composed of several subtasks, and there are different degrees of association between the subtasks:

$$\text{Totaltime}(p, q) = \sum_{i=1}^t F(Y_i + x + yt)^2 - \sum_{i=1}^t F(Y_i + x + yt)^2. \quad (2)$$

Depending on whether there is an association between tasks, the association problem can be divided into two types: independent task association and dependent task association. The correlation degree of independent tasks mainly refers to assigning unrelated tasks to different processors, and the correlation degree processes of different tasks are independent of each other. There is a certain correlation between each subtask that depends on the task correlation degree, and a certain execution sequence needs to be satisfied in the correlation degree process. The IaaS cloud data center is modeled as a G/N/N queuing system with N as a variable, and then, a feedback resource controller is constructed based on this model, which can change the number of virtual machines used according to the change of the number of tasks. The experimental results show that although this method will cause a certain degree of waste of resources, it significantly reduces the waiting time of task requests and the probability of violating QoS constraints. The algorithm first determines the uplink weight of each task according to the communication overhead and average calculation overhead between dependent tasks, and sorts the uplink weights of the task set in descending order to obtain the execution order of each task. Tasks are allocated to the server that completes the task first in the server cluster to achieve the goal of the shortest total system completion time.

3.2. Balanced Classification Mode. According to the different scaling methods of balanced classification, the methods for realizing elastic resource allocation can be divided into vertical scaling and horizontal scaling. Vertical scaling refers to the ability to change system resources or services by increasing or weakening the capabilities of current computing resources (such as the number of CPU cores and memory size). A prediction model (HGE) is composed of multiple steps of load feature extraction, classification, and prediction based on techniques such as fixed-size

overlapping sliding windows, KFCM clustering, and architecture-optimized Elman neural network. Based on the model in Figure 1, an adaptive cloud resource elastic configuration framework based on resource monitoring and load prediction is constructed to improve the utilization of cloud resources.

During the operation of each entrepreneurial establishment mode node, it will communicate with the relevance node through the information cycle and report to the relevance node the number of its currently idle task slots, the running status of the task it is responsible for, and receive new task instructions. There is also a very important component in the entrepreneurial establishment mode node—BSPPeer. BSPPeer is the core component to realize the BSP model in the entrepreneurial establishment mode. The parallel computing framework of the entrepreneurial establishment mode provides the communication function in the BSP model through the BSPPeer. When the startup is running, a BSPPeer will be generated for each task of the startup, and these tasks will communicate with each other through the BSPPeer. The configuration file dynamic loading test has two test purposes: one is to verify whether the configuration file monitor ConfiguraitonFileMonitor can work normally, and the other is to verify whether the reloadConfiguration method in ResourcesManager can work normally:

$$\begin{cases} PI[\text{action}(i)] = \text{round}[\text{sigma}(i)], \\ PI(i, j) = PI(\text{sigma}(i + j)). \end{cases} \quad (3)$$

The choice of entrepreneurial formation mode is fixed and the system is overloaded, so a dynamic clustering task correlation algorithm is proposed. The algorithm establishes an elastic correlation model according to the granularity of the task and the number of entrepreneurial formation modes. In this model, the scale of task nodes can be dynamically changed, and the heavy-load entrepreneurial formation mode is added according to the load of the current entrepreneurial formation mode. Node clustering, or being separated from light-loaded node clustering, can effectively allocate task nodes dynamically according to the load situation of the entrepreneurial formation mode, expand the flexibility and scalability of the system, and effectively reduce the overall relevance time. The correlation degree of independent tasks can be described as follows: assigning n mutually independent tasks to m virtual machines with a reasonable strategy, the execution processes of the tasks are independent of each other, and the tasks are inseparable. Its goal is to execute all tasks in parallel in the shortest time, and the correlation system is in a load-balanced state. Assuming that the strategy for assigning tasks to virtual machines is defined as a matrix X , according to a given n tasks and m virtual machines, the mapping scheme of tasks to virtual machines is shown in Table 1.

After the above process is successful, the files related to running the startup (including the JAR file of the startup, the configuration information of the startup and the shard information of the input file) will be copied to the directory of the association node on HDFS, and the ID of the startup

will be used as the file name identifier. The backup degree of the startup JAR is usually high (the default is 10), so that the JAR executable files can be dispersed enough in the cluster, so that the startup formation mode can quickly read startup resources when running startup tasks. The applications based on the parallel computing framework of the entrepreneurial formation mode all need to inherit the abstract base class BSP. The implementation of the user algorithm needs to be defined in the BSP method, which accepts a BSPPeer object as a parameter, and the BSPPeer is responsible for providing input and output functions for tasks, and implementing communication and synchronization between tasks. In addition to the BSP method, there are two optional methods, namely, the setup method and the cleanup method. Users can choose to implement these two methods to achieve program preparation (setup) and cleanup.

3.3. Cluster Correlation Algorithm. In practical applications, in order to reduce the execution time of relevance, many relevance tasks such as scientific computing, image processing, and other tasks can be divided into several subtasks for parallel execution; in addition, due to the load fluctuation of IaaS cloud data center, there are changes in the network bandwidth. Due to the influence of factors, the performance of virtual machines is not fixed. Considering the above two problems, this article proposes a relevance relevance algorithm that supports task parallelization and a relevance relevance algorithm in a dynamic IaaS environment, which can help the entrepreneurial establishment model provider to better deploy relevance on the IaaS public cloud. System performance requires that the minimum processing time can be obtained with the least overhead in the case of processing the same tasks, so as to achieve the optimization of system performance. A good relevance algorithm can satisfy a small time complexity at the same time. In the process of task relevance, a large number of task inputs are often faced. Only an efficient algorithm can achieve a good relevance effect:

$$\begin{aligned} & \text{argument}\{\max(s, a) - \min(s, a)\} \\ & - \text{abs}\{\max(s, a) - \min(s, a)\} = 0, \end{aligned} \quad (4)$$

where k represents the number of iterations of the PSO algorithm, $t-1$ and $t-2$ are the weight coefficients, and they all take values between $[0, 1]$. It represents the current optimal position experienced by the i -th node, that is, the position where the best objective function value is obtained, also known as the local optimal solution or the historical optimal solution; g_{best} represents that among all nodes, the optimal objective can be obtained. g_{best} is also called the global optimal position as it represents the position of the function value. R_1 and R_2 are random numbers, which can ensure that the PSO algorithm always maintains the randomness of the solution during the iterative optimization process, so the understanding space is expanded. That is, we choose the action with the largest Q value to execute, so that the current best reward expectation value can be obtained. However, this may only be a local optimum. Because the relatedness nodes have not accumulated enough experience in the early stage

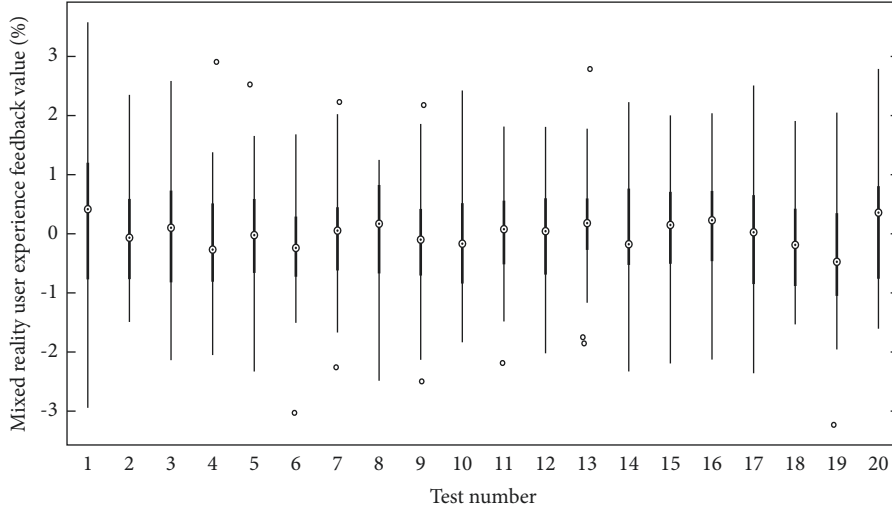


FIGURE 1: The distribution of the feedback rate of the balanced classification of the relevance degree.

TABLE 1: College students' entrepreneurship mapping scheme of virtual machine.

Mapping index	Cluster(1,1)	Cluster(1,0)	Cluster(0,1)	Cluster(1,2)	Cluster(2,2)	Cluster(2,1)
Cluster(1,1)	1.42	1.48	1.54	1.6	1.66	1.72
Cluster(1,0)	0.77	0.79	0.81	0.83	0.85	0.87
Cluster(0,1)	0.12	0.1	0.08	0.06	0.04	0.02
Cluster(1,2)	0.53	0.59	0.65	0.71	0.77	0.83
Cluster(2,2)	1.18	1.28	1.38	1.48	1.58	1.68
Cluster(2,1)	1.83	1.97	2.11	2.25	2.39	2.53

of learning, some actions with high rewards may not have been performed, and the Q estimated value obtained from the existing knowledge is not the real Q value. On the contrary, if the relevance node abandons the current optimal strategy and tries to perform other actions, although there is a certain probability that a better action can be found, it is also likely to lose a part of the reward because the selected action is not good.

In the cooperative entrepreneurial formation mode, the participants of the entrepreneurial formation mode in Figure 2 pursue the global optimal income, and the participants will reach an agreement that is binding on all parties, and the participants will carry out the entrepreneurial formation mode within the scope of the agreement. The noncooperative entrepreneurial formation mode mainly studies how the participants in the entrepreneurial formation mode, as independent individuals, choose the best strategy only from their own interests when there is conflict or competition in their interests. Therefore, the cooperative entrepreneurial formation mode emphasizes group rationality, while the noncooperative entrepreneurial formation mode emphasizes individual rationality. The main process is divided into two stages: in the first stage, the Max-Min algorithm is used to perform pre-association, and the initial mapping scheme is obtained; in the second stage, the secondary association is carried out, and the mapping scheme obtained in the first stage is heavily loaded. The tasks on the resources are redistributed to lightly loaded resources, thus effectively balancing the load. A heavily loaded resource is

defined as the virtual machine that achieved the maximum execution time in the first phase.

3.4. Real-Time Task Relevance Framework. The interaction between the relevance node and the environment is a continuous cycle: the relevance node changes the state of the environment by performing actions, and the environment feeds back some information based on the action performed by the relevance node, which reflects the pros and cons of the action; the relevance node uses the feedback information to update the policy to make better behavioral decisions in the future. Specifically, at a certain time t , the relevance node will select and execute an action according to the current state of the environment and the adopted strategy. After the action is executed, the environment will feed back a reward signal to the relevance node, and the state of the environment will change accordingly, migrating from s to $Ns+1$. Relevance nodes will adjust their strategy according to the reward signal. If positive feedback (reward value, usually positive) is received, then the tendency of connectedness nodes to perform this action in the future will increase. Conversely, if you get negative feedback (penalty value, which is usually negative), the tendency for the action to be performed is weakened.

$$\begin{cases} r_{i+j}^{i-j}|s.t(r(i, j) = s, x(s)) = \exp(i, j), \\ s_{i+j}^{i-j}|s.t(s(i, j) = \text{sigmod}(i, j)). \end{cases} \quad (5)$$

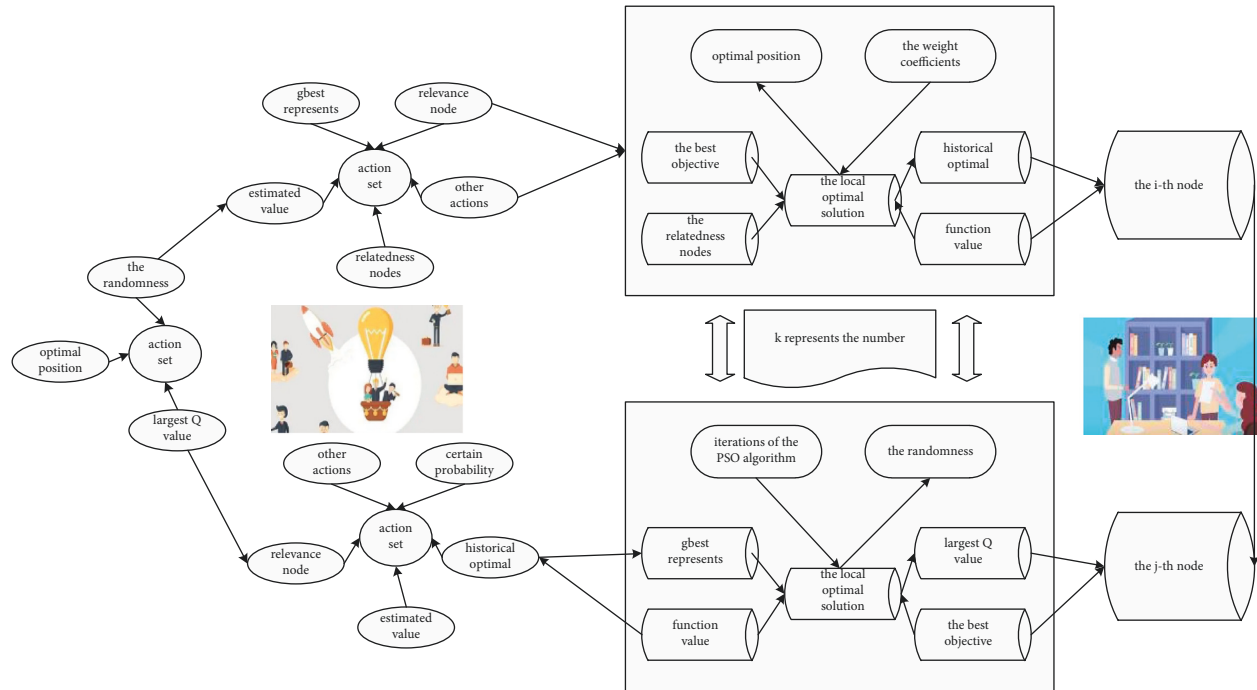


FIGURE 2: Hierarchical dependency of cluster correlation degree algorithm.

This article firstly analyzes the data of the four entrepreneurial formation models in separate cases, draws on the content analysis method of grounded theory, categorizes the collected data on a large number of fragmented entrepreneurial formation models, classifies them into their respective categories, and as shown in the text, summarizes the key factors and indicators in the strategic selection of the entrepreneurial formation mode of each case. Then, through the comparative analysis of multiple cases, to find the similarities and differences of each category in the group, analyze the data from a structured and diversified perspective, and finally summarize the strategic selection model of the new Internet entrepreneurial formation mode of college students' entrepreneurial teams. By calculating the upward weight value of each task in the task set, the larger the value, the higher the priority of the task; that is, the earlier the task should be executed; the smaller the value, the later the execution order of the task. Arrange each task in descending order according to the value of V to obtain the priority order of the tasks; that is, the order of execution is T1, T3, T4, T2, T5, T6, T9, T7, T8, T10. The stages in Table 2 mainly allocate the corresponding virtual machine to each task according to the task priority order obtained in the first stage and use the greedy strategy to allocate the task to the virtual machine that can be completed earliest each time, so as to achieve the total task completion time, shortest.

Since the entrepreneurial formation mode algorithm itself adopts a greedy strategy, when the Max-Min algorithm is used for pre-association in the first stage, the algorithm has obtained a preliminary local optimal solution; through the second stage, the tasks on the virtual machine with heavy load will be reassignment to the light-loaded virtual machine further reduces Makespan, makes the load more balanced,

and improves the resource utilization of the system. A total of 635 questionnaires were returned from the formal survey questionnaire, including 21 questionnaires with unknown information, and 108 invalid questionnaires whose answering time deviates from the normal time. In addition, because respondents who have not participated in the entrepreneurial training of Company Z cannot accurately answer the training questions, the total number of valid questionnaires recovered is 457, and the recovery rate of valid questionnaires is 72%. Therefore, the result of the correlation degree of the entrepreneurial formation mode algorithm is used as the global optimal solution when the node group algorithm is initialized, which has already achieved a good correlation degree. In the subsequent node update process, it can further optimize the entrepreneurial formation mode algorithm. Result. In this way, the quality of the entrepreneurial formation mode during initialization can be guaranteed, thereby effectively improving the node's ability to search for optimization.

4. Research on the Model Framework of College Students' Entrepreneurial Team Formation Based on Correlation Algorithm

4.1. Offline Training of Correlation Algorithm. The relevance algorithm environment model imitates the behavior of the environment. In short, when an action is performed in any state, the environment model can predict its next state and reward value. The goal of the relevance algorithm is to find an optimal mapping from the state space to the action space, that is, thereby maximizing the total cumulative reward value of the target task. Because the relevance algorithm

TABLE 2: Description of task assignment of the relevance algorithm.

Description unit name	Makespan value (%)	Speedup time (ms)	Average error
Formation mode	63.4820	222.5781	0.0724
Task of the system	60.5651	141.2305	0.0753
Utilization of the system	68.4246	525.8724	0.0608
Execution mode	66.4185	720.9400	0.0637
Priority of the task	63.2855	131.2542	0.0666
Deformation mode	61.3231	690.7826	0.0695

solves the problem of sequential decision-making, if only the immediate reward in the current state is considered and the future influence is ignored, the problem will fall into a local optimal solution, and the optimal strategy cannot be obtained. Therefore, the relevance algorithm uses the value function to evaluate, which can measure a certain state or state well. Unlike before, these four tasks are no longer independent individuals, nor can they be simply assigned to the virtual machine with the best computing performance for parallel execution only from the perspective of minimizing the time span. Dependencies between tasks may be such that some tasks must be executed before other tasks:

$$\sum_{y,k}^{y+k} \left(\frac{dr}{dk} + \frac{dr}{dj} + \frac{dr}{di} \mid (i, j, k) \longrightarrow (1, 1, 1) = \lim 1 - x \right). \quad (6)$$

When the relevance node receives the signal of the completion of the last task of the startup (the cleanup task mentioned above), it will set the status information of the startup to "success." Then, when the client invokes the entrepreneurial information, the "success" identifier is returned to the client. After the client knows that the startup has been successfully run, it will print startup statistics on the console. Finally, the connection degree node will clean up the related resources occupied by the startup and notify the startup formation mode to do the cleanup work. The relevance node is also responsible for receiving the startup submitted by the client and initializing the startup, and then adding the startup to the waiting queue of the first-come, first-served startup correlation degree. In this process, the subclass of JobInProgressListener (JobListener in the first-come, first-served entrepreneurial correlation detector) acts as a communication bridge between the correlation node and the first-come, first-served entrepreneurial correlation detector, and the addition of entrepreneurship is finally completed by the JobListener:

$$\cap_{i,j}^{i+j} (x_1^i + x_2^i + x_3^i + \dots + x_{i+j}^i) - \cap_{i,j}^{i+j} (y_1^i + y_2^i + y_3^i + \dots + y_{i+j}^i) > 0. \quad (7)$$

After the TaskRunner loads the main program submitted by the user, it first initializes an instance of the BSPPeer class for the task according to the configuration file and then uses it as a parameter of the setup method to start executing the setup method. The setup method generally only completes some relatively simple tasks, and the BSP method executed after the setup is where the algorithm logic really runs. When the BSP method is executed, it first reads the data input shard

information it is responsible for. At this stage, each task of the job is independent of each other. The processing of data follows the process of "parallel computing-communication-barrier synchronization" of the BSP model. The fitness of each node is calculated, and the node with high fitness is selected as the node required for the next iteration. When the newly generated particles have the same fitness function value as the current node, the newly generated individuals are preferentially selected to ensure the diversity of entrepreneurial formation patterns.

4.2. Load Analysis of College Students' Entrepreneurship Model. During the experiment, the main comparison parameters in this chapter are the total completion time of the load task in the entrepreneurial mode, the degree of system load balance, and the convergence of the algorithm. Set the number of tasks = 10, 20, 30, 40, 50, 80, and the number of iterations to 100. The length of the task is randomly generated, and the total task completion time of different algorithms is recorded. Aiming at the problem that the fluctuation of the application load of the entrepreneurial establishment mode under fixed resources will lead to insufficient or excessive supply of resources, a load-based elastic resource allocation method is proposed. Because the resources of the IaaS cloud data center are provided on-demand and pay-per-use, if the entrepreneurial establishment model provider can dynamically allocate resources according to changes in the application load, the computing power provided by the rented virtual machine can be close to that of the user at different times. It can reduce the rental cost on the premise of ensuring the quality of service. The adaptive cloud resource leasing algorithm proposed in this article fully considers the fluctuation of application load and the diversity of IaaS resources:

$$W(x_1^i + x_2^i + x_3^i + \dots + x_{i+j}^i) \leftarrow w(x_1^i) + w(x_2^i) + w(x_3^i) + \dots + w(x_{i+j}^i), \text{ for } \{\text{number}(x) > w\}. \quad (8)$$

Questionnaire validity usually refers to the correctness and validity of the questionnaire. In this study, factor analysis was used to test the characteristic validity of the questionnaire. If KMO (the proportion of variance that can be extracted by factor analysis) is above 0.6, the closer it is to 1, the better the correlation between the questionnaire items and the higher the quality of the questionnaire. Using the KMO sample test method and Bartlett's sphere test provided

in SPSS23.0 software for testing, the expansion of the target market is generally divided into two steps: first, through the technology research and development and improvement in the vertical market, or the development and expansion of new vertical markets, so that the company can enter from the original strategy, usually the vertical advance strategy, to the creative imitation strategy. Then, from the strategic model of the platform market, such as the early market entry strategy and the platform emerging industry innovation strategy, choose the type of platform market strategy suitable for your own entrepreneurial formation model, and finally gradually become a real platform. While the initial strategic choice is the creative imitation strategy, the entrepreneurial formation model skips the first step and directly enters the platform market through changes in the market and technology, expanding the development space of the entrepreneurial formation model.

Before using the structural equation model in Figure 3 for confirmatory factor analysis, it is necessary to check whether the factor loading of each observed variable in the exploratory factor analysis results is significant in the parameter estimation in the structural equation model, and to test the reliability and reliability of the measurement model of the latent variables. To measure the internal consistency of the model, the composite reliability of latent variables, that is, the C.R value, is generally used to measure. It is mainly used to evaluate the consistency of the latent variable measurement indicators in the model. The higher the combined reliability value, the higher the internal correlation between the measurement indicators. When the path

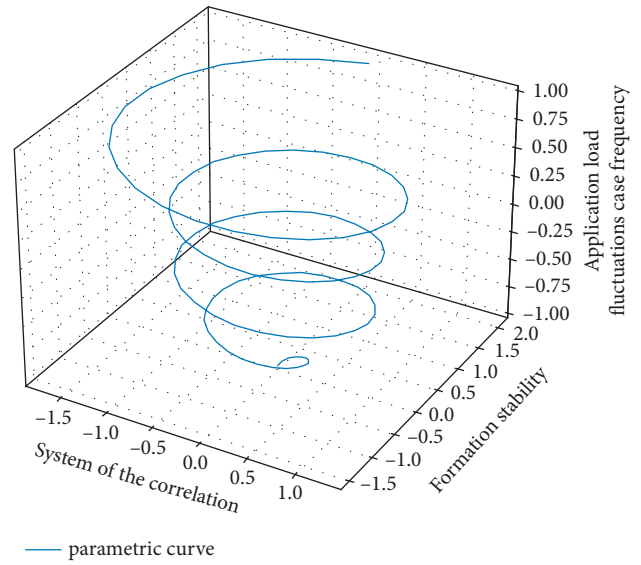


FIGURE 3: Application load fluctuation of the entrepreneurial formation model based on relevance.

between the latent variable and the indicator variable is significant, that is, the P value is less than 0.05, the absolute value of the path coefficient of the standardized observed variable is greater than 0.5, and the C.R. value is greater than 0.7, indicating that the internal consistency of the measurement model is good:

$$\text{tasks}(x, t) = \overbrace{tas\left(\text{arrival}\left(\sqrt{x_1^i}\right)\right) + tas\left(\text{arrival}\left(\sqrt{x_2^i}\right)\right) + tas\left(\text{arrival}\left(\sqrt{x_1^i + x_2^i}\right)\right)}^{i+j} . \tag{9}$$

Assume that whenever a new task arrives, the task associator will immediately assign it to a virtual machine in the resource pool. The task will first enter the waiting queue of the specified virtual machine and then execute in the order of first come, first served. It is assumed that the waiting queue space of each virtual machine is large enough to store all pending tasks. Additionally, each virtual machine can only execute one task at a time, and tasks cannot be interrupted by other tasks during execution. According to the above rules, if a new task is assigned to a virtual machine whose waiting queue is empty, then the task can be executed immediately. Otherwise, the new task needs to wait for all the tasks that arrived before it to start executing. This study tested the reliability and validity of the measurement model. The results show that the standardization factor loadings are all above 0.704, the combined reliability of each latent variable is greater than or equal to 0.820, and the average variance extraction AVE value is greater than or equal to 0.59. It shows that the measurement model of 10 latent variables has good reliability and convergent validity.

4.3. *Forming Pattern Framework Factor Recursion.* According to the dependent task correlation model described by the entrepreneurial mode load, fully consider the hierarchy of task nodes and the differences in the performance of the entrepreneurial formation mode, optimize the task priority determination process on the basis of the existing table correlation degree, and recursively determine from the exit task. The uplink weight of each task is used to determine the execution sequence of each task; on this basis, the insertion strategy and the strategy of the earliest completion time with the smaller the priority are used to allocate the tasks to the corresponding resources, and the effectiveness of the algorithm is verified by experiments. The cumulative expected return is obtained by the associated degree node performing action in states. Obviously, no matter what kind of value function it is, the above calculation formula is a function related to multiple subsequent states. In the Markov decision-making process, the state at the next moment and the reward obtained are only related to the current state and the selected action. After the association

node does not execute the action port according to the policy in the states, the probability of pAP(s) will be transferred:

$$w(u, v) - \frac{\sum \text{adervange}(pi \times s|u + vt) - \sum \text{adervange}(pi \times s|u - vt)}{1 - \sum \text{adervange}(u + vt)} = 0. \quad (10)$$

When making task relatedness decisions, it is necessary to select the best action for the current task according to the current state and the Q estimated value obtained from the deep neural network, that is, to decide which virtual machine to assign the task to. The algorithm adopts a t -greedy strategy for selection, that is, randomly selects an action from the set of actions with a probability of t , or selects the action with the highest Q estimated value with a probability of $1-e$, which is a variable in the range of $[0, 1]$, which has a large initial value and then gradually decreases as the decision progresses. This means that at the beginning of the algorithm, the virtual machine will be randomly selected with a high probability to explore more possibilities. But as more and more decision-making experience is accumulated, the algorithm will make more use of the existing experience to choose actions with high rewards, rather than exploring and trying those bad actions.

Further test the discriminant validity of latent variables and compare the correlation coefficient of each latent variable in Figure 4 and the square root of the average variance extraction (AVE); the calculation results show that the square root of the average variance extraction AVE value of each latent variable is greater than that of the latent variable. The correlation coefficient with all other latent variables indicates that the model has high discriminant validity. The research on the entrepreneurial choice of college students in this article is to regard entrepreneurial choice as a discrete variable with sorting characteristics: the field that is combined with one's own major is 1, the field of interest is 2, the current popular field is 3, and it is easy to start a business with less startup capital. The low industry is 4; therefore, the multivariate logit model should be used in this article. Because the random utility distribution form assumed by the logit model is more suitable for the distribution selection when the utility is maximized, the logit model is the most widely used. Since the results obtained in this article are all discrete data, the multivariate discrete model is selected as

the model, and the most widely used multivariate classification logit model is selected for regression analysis, and the parameter estimates of the respective variables and their marginal contributions are obtained.

5. Research on the Model Framework of College Students' Entrepreneurial Team Formation Based on Correlation Algorithm

5.1. Feature Extraction of Relevance Data. The characteristic validity of the correlation data reflects the actual measurement degree of the observed variable to the latent variable to be measured, including convergent validity and discriminant validity. When the standardization factor loading value of the observation index is greater than 0.5, the average variance extraction AVE value is greater than 0.5, and the combined reliability C.R. value is greater than 0.7, it indicates that the model has good convergent validity. When the square root value of the average variance extraction AVE of each latent variable is higher than the absolute value of the correlation coefficient between this variable and all other latent variables, it indicates that the latent variable has better discriminant validity. The P value of the path analysis result between training courses and quality improvement is 0.003, the standard β value is 0.286, and the CR value is 2.982; the path analysis result between training courses and social benefits is 0.021, the standard β value is 0.234, and the CR value is 2.302; the P value of the path analysis results between supervision and feedback and quality improvement is less than 0.001, the standard β value is 0.420, and the CR value is 4.886. The P -values between the other path relationships were all above 0.05, and the normalized beta values were between -0.078 and 0.114. The path analysis results of the measurement model are shown in the text. The dashed connecting line indicates that the path relationship is not significant:

$$\text{arguser} \left(\text{maximum}(u, s) \prod \sqrt{\frac{u+s}{u-s}} + \sqrt{\frac{u+s-1}{u-s}} + \sqrt{\frac{u+s-n}{u-s}} \right) \longrightarrow \left\{ (u, s) \mid \sqrt{\frac{u+s-n}{u-s}} \right\}. \quad (11)$$

For the application of the entrepreneurial formation mode that requires immediate response, the user task request load it receives is unpredictable and changes with time, so it is more suitable to use the method of online correlation. The method based on reinforcement learning is very suitable for such scenarios, because it is not necessary to know the user's task load, system operating environment, and other

information in advance. With the operation of the entrepreneurial formation mode application, the best task relevance strategy can be learned online.

Using these data, the task associator can improve the associativity policy in real time. In order to avoid this deficiency, the entrepreneurial formation mode algorithm is used to pre-associate the task, and the result is used as the

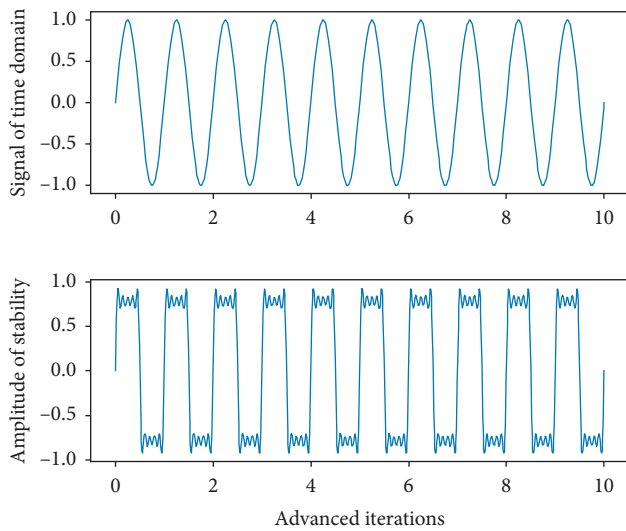


FIGURE 4: Iterative distribution of task relatedness decision-making.

global optimal solution of the initial node group, so as to optimize the node initialization operation; redesign the fitness function, and introduce it into the architecture (GA). The crossover mutation operation of it expands the solution space of the nodes, so that the nodes in Table 3 can search for the optimal solution with a greater probability, and the improved algorithm is compared with similar algorithms through experiments.

The a value of each dimension of the questionnaire is between 0.716 and 0.922, which shows that there is a good consistency between each dimension and each item in the questionnaire. The split-half reliability of each dimension is between 0.783 and 1.000, and the split-half reliability of the total questionnaire is 0.882, reaching a good level, indicating that the split-half reliability of each questionnaire is good. From the comprehensive point of view of various indicators, this questionnaire has good reliability. This article conducts a questionnaire survey on college students of different grades, majors, and educational levels in some colleges and universities in the region, effectively verifies the data after the survey, then identifies the key influencing factors of college students' entrepreneurial choices, and discusses these factors in depth. It can be seen that the difference in the number of tasks allocated to the two virtual machines is getting smaller and smaller. This is because assigning tasks to the oldest idle virtual machine reduces task wait time, but increases task execution time if it is not the best matching virtual machine type. For the sensible algorithm, because it is similar to the earliest algorithm, it is also a time greedy strategy, so the two perform similarly:

$$\begin{aligned} \forall t \in T (s > 0, u < 1, v < 1), \\ \exists \text{arguser}(\text{maximum}(1, 1)|(u, s)) \subset A(u, s). \end{aligned} \quad (12)$$

Among them, the strength of the independent variable confidence and decisiveness has the most significant impact. It can be seen from the data that for each unit of increase in this variable, the possibility of college students' choice of

field (2) is increased by about 21% compared with the original possibility. The influence of innovation and teamwork is also more significant. With each additional unit, college students are more inclined to teamwork, so the possibility of engaging in their own areas of interest will be relatively weakened. At the same time, the test of the individual trait variable group of college students in the field (3) is also significant, where $\chi^2(3) = 8.15$, $P = 0.027$, the result rejects the null hypothesis that these three variables have "all zero effects," and the alternative hypothesis can be accepted; that is, compared with the reference group, the three individual trait variables of college students have an impact on their choice of entrepreneurial behavior in today's popular fields. At the 5% confidence level, the internal control variable of college students is significant, and each additional unit increases the probability of college students choosing field (3) by about 19%. However, the assertion that the stronger the self-confidence and the decisiveness, the more inclined to choose the field (3) for entrepreneurship is not supported by the empirical results of this article. Domain (4) failed the significance test, $\chi^2(3) = 3.71$, $P = 0.163$, and the result accepted the null hypothesis of "all effects are 0" for the three variables.

5.2. Simulation Realization of College Students' Entrepreneurial Team Formation. This research adopts two methods of manual editing and software processing for the questionnaire information collected from the formation of entrepreneurial teams. First, the collected questionnaires need to be sorted manually, and invalid questionnaires with errors or missing important information are eliminated. At the same time, the valid questionnaires are numbered and archived, and the questionnaire data are coded into the EXCEL software to establish a database. In order to ensure the objective authenticity of the results, the questionnaires were answered anonymously. All questions were factor-analyzed accordingly, using KMO and Bartlett's test for sphericity. The KMO value is 0.961 and the chi-square value is 2501.72, $P < 0.01$, so the questions in the whole questionnaire are valid. This study standardized the dimensions of the questionnaire according to eigenvalues, the principle of steepness, and the number of factor items. Statistical results show that the questionnaire in this article is divided into three dimensions (influencing factors such as social statistical characteristics, college students' entrepreneurial choices, entrepreneurial individual characteristics, entrepreneurial environment, and entrepreneurial preparations are set in the text), and each dimension contains 2–5 questions, explaining 72.061% of the total variance. All questions were analyzed for reliability and validity, and the reliability of the questionnaire was tested using the internal consistency index (α) in Figure 5.

This article uses the multivariate classification logit model and the independent sample t -test method. The author uses the independent sample t -test in the research on personal experience to investigate whether there is a difference in the entrepreneurial choices of students with different genders and differences between urban and rural areas. The relationship between college students'

TABLE 3: Task relevance algorithm flow.

Task relevance data	Algorithm content test
Input result as $ai + bj$	The global optimal solution
The result is used as $\exp(-x/t)$	Import java.awt.event.actionevent;
Search for the $r!(n - r)!$	Import java.awt.event.actionlistener;
The crossover mutation operation	Import java.io.File;
Private jbutton = new $1 - f(s, t, r)$;	$\sum \sqrt{y(i, j)}$ into the architecture (GA)
Private but2 = new jbutton("delete");	Setbounds(150, 310, 60, 25) $\sqrt{y(t - x, j)}$
Optimal solution with $ai + bj$	Hashtable < String, Person >
For data with $a(i - 1, \sqrt{i + 1})$	Objectoutput tanh($x - 1$) = new object
OutputStream out = new object	To pre-associate the log(d), $d > x$ task
Clear.setbounds(250, 310, 60, 25);	($e, e(x, t')$) with a Greater probability
Output result as $\Delta m(m - n)$	Address.setbounds(230, 135, 90, 25);

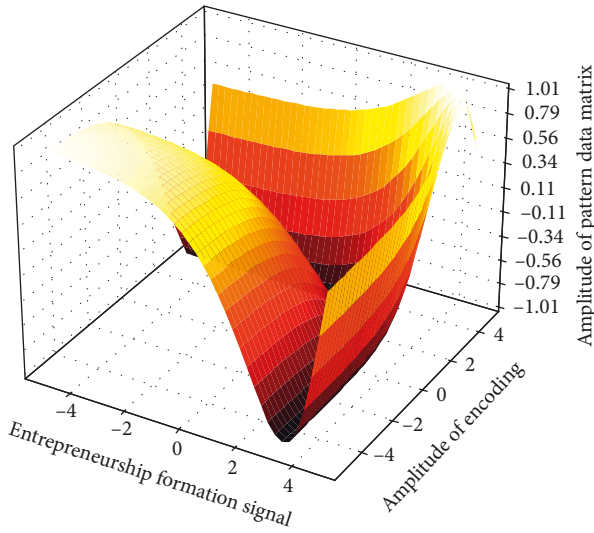


FIGURE 5: Entrepreneurial formation pattern data coded three-dimensional distribution.

entrepreneurial choices and personal experiences can be further studied. When the number of tasks is small and the number of iterations is large, the optimal solution can be found quickly, and these three algorithms all use the genetic algorithm in the process of association degree, so they can achieve good results. With the increasing number of tasks, the advantages of the improved algorithm proposed in this chapter are more obvious. This is because the PSO algorithm and the JLGA algorithm cannot guarantee the quality of the initial entrepreneurial formation model, and it is easy to fall into a local optimal solution in the later stage, while the improved algorithm uses the entrepreneurial formation model algorithm to pre-associate the tasks, which has already achieved good results. The correlation effect is further optimized by the crossover mutation operation of the PSO algorithm and GA in the later stage. Therefore, with the increase in the task scale, it can achieve a better correlation effect than the other two algorithms:

$$y(r, x) = \begin{cases} \sqrt{r(s, a, b)} + a \times \sqrt{r(s, a, b)}, & a > b, \\ r(s, a, b) + a \times \sqrt{r(s, a, b)} - b \times r(a, b), & a < b. \end{cases} \quad (13)$$

It can be seen that the DQN algorithm performs better than all other comparison algorithms. No matter how the probability changes, the average task response time obtained under the DQN algorithm has been stably maintained at a very low level. In contrast, the performance of other algorithms fluctuates greatly. When the probability changes from 0.1 to 0.5, the average task response time of the earliest algorithm and the sensible algorithm becomes longer and longer, and the average task response time of the other three algorithms becomes shorter and shorter. To explore the reason behind it, this experiment also studies the distribution of tasks on two types of virtual machines and the occupancy of two types of virtual machines. It should be noted that only the case where the probability is less than or equal to 0.5 is drawn in the figure, because it is less than 0. The experimental results of 5 and greater than 0.5 are basically symmetrical.

After standardizing the 457 samples data in Figure 6, import it into AMOS23.0, and match the data to the corresponding variables according to the model settings, and perform path coefficient estimation and significance testing. The calculation results of relevant parameters show that the P value of the path analysis results between training needs and quality improvement is 0.011, the standard β value is 0.192, and the CR value is 2.55; the path analysis obtains the relationship between training needs and social benefits. The P value is less than 0.001, the standard β value is 0.348, and the CR value is 3.891; the P value of the path analysis result between knowledge and technology foundation and social benefits is 0.002, the standard β value is 0.199, and the CR value is 3.117; the P value of the analysis results was 0.017.

5.3. Example Application and Analysis. The execution result of the entrepreneurial establishment mode node will be sent to the virtual machine where the successor task is located as the input data necessary for the execution of its successor task. Correspondingly, the export task can only be used as the successor node of other tasks, its relevance and execution are after other tasks, and the priority is the lowest. AMOS23.0 calculates the path coefficient of the measurement model, and when testing the significance level, P is less than 0.05, and the corresponding T value is greater than 1.96. The normalized path coefficients listed in the table ranged from -0.078 to 0.420 for H1 to H16, respectively, and the corresponding C.R.

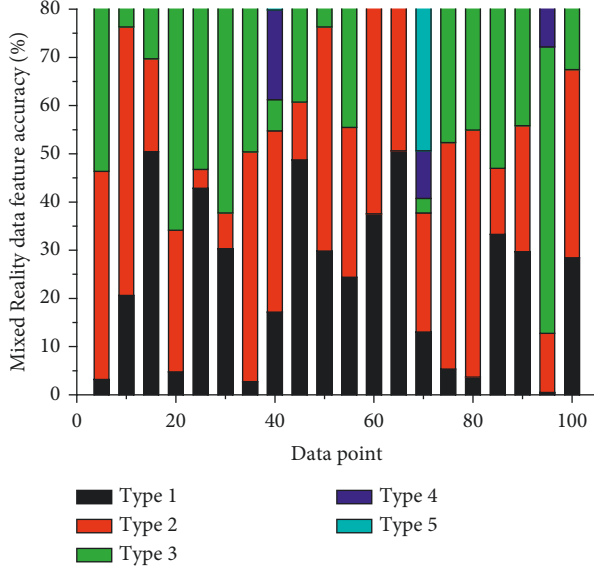


FIGURE 6: Feature extraction distribution of the entrepreneurial formation mode algorithm.

values ranged from -0.794 to 4.886 . Among them, in terms of the impact of training background on training effect, the impact of training demand on quality improvement has a path coefficient of 0.192 ($P = 0.011$); the impact of training demand on social benefits has a path coefficient of 0.348 ($P < 0.001$). The influence of foundation on social effect for the path coefficient is 0.199 ($P = 0.002$). In terms of the influence of training investment on training effect and the influence of teacher strength construction on quality improvement, the path coefficient is 0.218 ($P = 0.017$):

$$\frac{\sum \partial(p(s,a)|u(s,a) + v(t))}{\sum \partial(u + vt)} - \frac{\sum \partial(p(s,a)|u(s,a))}{\sum \partial(u + vt)} - \frac{\sum \partial(p(s,a)|v(t))}{\sum \partial(u + vt)} = 0. \quad (14)$$

The first stage calculates the priority of each task, defines the uplink weight of each task, and then sorts the uplink weight in ascending order to obtain the execution order of the tasks in the process of relevance; the second stage is the virtual machine allocation stage; through each case, assign the highest priority task to the virtual machine that completes it first, and execute all tasks in the shortest time possible. When calculating the uplink weight of each task, the HEFT algorithm adopts a bottom-up approach, according to the average computing cost of each task on different virtual machines and the communication cost of different virtual machines when switching tasks and the successor for the uplink weight of the task node determines the uplink weight of the task. When T-i and T-k are executed on the same virtual machine, the communication overhead between the two is 0, because compared with the communication overhead when different tasks are executed on different virtual machines, the switching overhead when executing different tasks on the same virtual machine is negligible. Before the task relatedness of Figure 7 starts, the

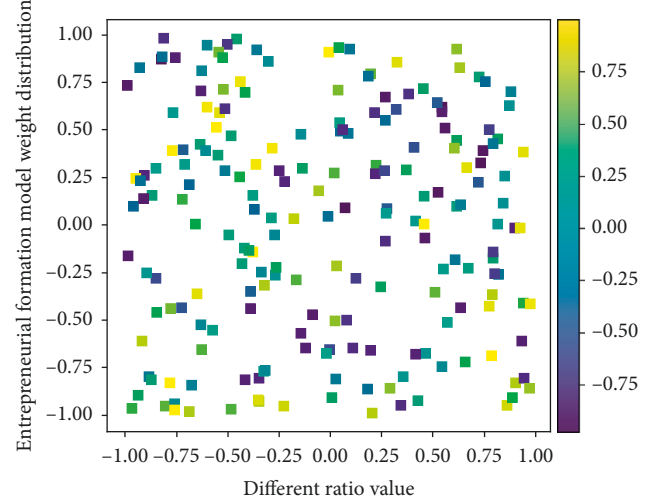


FIGURE 7: Weighted distribution of the entrepreneurial formation model.

average communication cost k when different tasks switch between different virtual machines can be quantified by the weights of the directed edges in the DAG.

If the fitness value of the optimal individual in the entrepreneurial formation model can maintain a certain algebra without significant change, the algorithm can be considered to have converged, and the iteration is terminated at this time and the corresponding optimal solution is output. In addition, in order to avoid excessive search, an upper limit is usually set for the evolutionary algebra of the algorithm. When the iteration reaches the upper limit, the algorithm will be forced to terminate. In the experiments in this section, the maximum evolutionary generation of the algorithm is set to 150 generations, and the minimum retention generation of the optimal individual is 10 generations. In terms of the influence of the training process on the training effect and the influence of training courses on quality improvement, the path coefficient is 0.286 ($P = 0.003$); in terms of the influence of training courses on social benefits, the path coefficient is 0.234 ($P = 0.021$); and in terms of the influence of supervision and feedback on quality, the path coefficient was 0.420 ($P < 0.001$). The P values of the above relationships are all less than 0.05, indicating that the corresponding path coefficients are significant, and the corresponding hypothesis relationship is established. The P value of the other path coefficient significance test is greater than 0.05, indicating that the path coefficient is not significant, and the corresponding hypothesis relationship does not hold:

$$\left[\frac{uvm(t, i, j)}{\sqrt{uvm(t, i, j)}} \sin(t, i, j) \right] = \frac{[\sqrt{uvm(t, i, j)} \sin(t, i, j)]}{[uvm(t, i, j) \sin(t, i, j)] / \cos(t, i, j)}. \quad (15)$$

The parallel search method of multi-entrepreneurial formation mode also reduces the possibility of the algorithm

converging to the local optimal solution to a certain extent in the control of system resource utilization. Since MPGA adds the constraint of load imbalance value to the evolution of individuals, it eliminates the individuals that do not meet the limited conditions in time, thus maintaining a lower and more stable load imbalance value and ensuring the correct direction of correlation. It can be seen that the time span and cost of MPGA are only about 1/2 and 3/4 of that of AGA, respectively, and its load imbalance value is only 1/30 of that of AGA, and the resource utilization rate is significantly higher than that of AGA. Therefore, both in terms of time span and cost, the optimization algorithm proposed in this article is obviously better than the task correlation algorithm based on the standard genetic algorithm. In addition, it also ideally ensures the load balancing of the correlation system. Then in the empirical process, cases first select various variables, put forward hypotheses for the model, and then make a preliminary description of the main variables.

6. Conclusion

Considering the time of the correlation algorithm, this article proposes an independent task correlation strategy based on the multi-entrepreneur formation model. In order to expand the search range of feasible correlation schemes and improve the globality of the optimal solution, it is proposed to replace the basic structure with a multi-entrepreneurial formation mode architecture. For the entrepreneurial establishment model provider, task relevance refers to the reasonable allocation of each user task to the leased virtual machine, and the establishment of an optimal mapping relationship between tasks and computing resources, so as to realize user tasks. This article first systematically summarizes the existing research results of the strategic choice of the independent entrepreneurial startup formation model, based on the industry maturity (entering a mature industry or a growing industry). The three dimensions of the target market (entering the international or domestic market) and the seven competitive strategic choices of the independent entrepreneurial new entrepreneurial establishment model based on these three dimensions are the theoretical basis. The quality of task relevance and resource allocation algorithm determines the performance of the entrepreneurial formation model application, the user's satisfaction with the service, and the income of the entrepreneurial formation model provider.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of 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 article.

Acknowledgments

This work was supported by Student Affairs Office, Xi'an Siyuan University.

References

- [1] M. Farhan, S. Jabbar, M. Aslam et al., "IoT-based students interaction framework using attention-scoring assessment in eLearning," *Future Generation Computer Systems*, vol. 79, pp. 909–919, 2018.
- [2] E. V. Kuz'mina, N. G. P'yankova, N. V. Tret'yakova, and A. V. Botsoeva, "Using data analysis methodology to foster professional competencies in business informaticians," *European Journal of Contemporary Education*, vol. 9, no. 1, pp. 54–66, 2020.
- [3] Y. Wei, H. Lv, M. Chen, M. Wang, A. A. Heidari, and H. Chen, C. Li, "Predicting entrepreneurial intention of students: an extreme learning machine with Gaussian barebone harris hawks optimizer," *IEEE Access*, vol. 8, pp. 76841–76855, 2020.
- [4] M. E. Ibrahim, Y. Yang, D. L. Ndzi, G. Yang, and M. Al-Maliki, "Ontology-based personalized course recommendation framework," *IEEE Access*, vol. 7, pp. 5180–5199, 2019.
- [5] A. Renz and R. Hilbig, "Prerequisites for artificial intelligence in further education: identification of drivers, barriers, and business models of educational technology companies," *International Journal of Educational Technology in Higher Education*, vol. 17, no. 1, pp. 18–21, 2020.
- [6] A. Qazdar, B. Er-Raha, C. Cherkaoui, and D. Mammass, "A machine learning algorithm framework for predicting students performance: a case study of baccalaureate students in Morocco," *Education and Information Technologies*, vol. 24, no. 6, pp. 3577–3589, 2019.
- [7] T. Cavanagh, B. Chen, R. A. M. Lahcen, and J. R. Paradiso, "Constructing a design framework and pedagogical approach for adaptive learning in higher education: a practitioner's perspective," *International Review of Research in Open and Distance Learning*, vol. 21, no. 1, pp. 173–197, 2020.
- [8] M. Kuziemski and G. Misuraca, "AI governance in the public sector: three tales from the frontiers of automated decision-making in democratic settings," *Telecommunications Policy*, vol. 44, no. 6, p. 101976, 2020.
- [9] C. Troussas, F. Giannakas, C. Sgouropoulou, and C. Voyiatzis, "Collaborative activities recommendation based on students' collaborative learning styles using ANN and WSM," *Interactive Learning Environments*, vol. 2020, pp. 8–14, 2020.
- [10] Y. Chen, W. Zheng, W. Li, and Y. Huang, "Large group activity security risk assessment and risk early warning based on random forest algorithm," *Pattern Recognition Letters*, vol. 144, pp. 1–5, 2021.
- [11] E. Leonidou, M. Christofi, D. Vrontis, and A. Thrassou, "An integrative framework of stakeholder engagement for innovation management and entrepreneurship development," *Journal of Business Research*, vol. 119, pp. 245–258, 2020.
- [12] Q. Jia, Y. Guo, R. Li, Y. Li, and Y. Chen, "A conceptual artificial intelligence application framework in human resource management," in *Proceedings of the international conference on electronic business*, pp. 106–114, Guilin, China, 2018.
- [13] J. I. Gimenez-Nadal, M. Lafuente, J. A. Molina, and J. Velilla, "Resampling and bootstrap algorithms to assess the relevance of variables: applications to cross section entrepreneurship data," *Empirical Economics*, vol. 56, no. 1, pp. 233–267, 2019.

- [14] A. Bogarín, R. Cerezo, and C. Romero, "A survey on educational process mining," *Review: Data Mining and Knowledge Discovery*, vol. 8, no. 1, Article ID e1230, 2018.
- [15] L. Hu, S. He, Z. Han et al., "Monitoring housing rental prices based on social media: An integrated approach of machine-learning algorithms and hedonic modeling to inform equitable housing policies," *Land Use Policy*, vol. 82, pp. 657–673, 2019.
- [16] R. M. Tawafak, A. Romli, S. I. Malik, and M. Shakir, "IT governance impact on academic performance development," *International Journal of Emerging Technologies in Learning (ijET)*, vol. 15, no. 18, pp. 73–85, 2020.
- [17] A. Tittel and O. Terzidis, "Entrepreneurial competences revisited: developing a consolidated and categorized list of entrepreneurial competences," *Entrepreneurship Education*, vol. 3, no. 1, pp. 1–35, 2020.
- [18] O. Moscoso-Zea, J. Castro, J. Paredes-Gualtor, and S. Lujan-Mora, "A hybrid infrastructure of enterprise architecture and business intelligence & analytics for knowledge management in education," *IEEE Access*, vol. 7, pp. 38778–38788, 2019.
- [19] J. Finocchiaro, R. Maio, F. Monachou et al., "Bridging machine learning and mechanism design towards algorithmic fairness," *Fairness, Accountability, and Transparency*, vol. 2021, pp. 489–503, 2021.
- [20] E. Abad-Segura and M.-D. González-Zamar, "Sustainable economic development in higher education institutions: a global analysis within the SDGs framework," *Journal of Cleaner Production*, vol. 294, p. 126133, 2021.
- [21] J. Saltz, M. Skirpan, C. Fiesler et al., "Integrating ethics within machine learning courses," *ACM Transactions on Computing Education*, vol. 19, no. 4, pp. 1–26, 2019.
- [22] G. Secundo, P. Rippa, and R. Cerchione, "Digital Academic Entrepreneurship: a structured literature review and avenue for a research agenda," *Technological Forecasting and Social Change*, vol. 157, p. 120118, 2020.
- [23] T. D. Ullmann, "Automated analysis of reflection in writing: validating machine learning approaches," *International Journal of Artificial Intelligence in Education*, vol. 29, no. 2, pp. 217–257, 2019.
- [24] F. Corradini, A. Ferrari, F. Fornari et al., "A Guidelines framework for understandable BPMN models," *Data & Knowledge Engineering*, vol. 113, pp. 129–154, 2018.
- [25] M. Injadat, A. Moubayed, A. B. Nassif, and A. Shami, "Multi-split optimized bagging ensemble model selection for multi-class educational data mining," *Applied Intelligence*, vol. 50, no. 12, pp. 4506–4528, 2020.