J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 3583--3602

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Research on Employment-Oriented Talent Cultivation Model for the Pathway from Vocational Education to Higher Education Based on Computational Optimization

Ruiqian Su[1], Yanfang Wang[2,✉]

[1] School of Foreign Languages, Xiamen Institute of Technology, Xiamen, Fujian, 361021, China
[2] College of Humanities, Xiamen Huaxia University, Xiamen, Fujian, 361024, China

ABSTRACT
In recent years, socio-economic development and the process of massification of vocational education have been accelerating. The article surveys the current situation of the articulation between vocational education and undergraduate education through questionnaires. On this basis, in order to better realize the cultivation of employment-oriented talents, it designs a teaching resource acquisition method based on computational optimization, constructs a crawler search method by fusing genetic algorithm and ant colony algorithm, and realizes automatic clustering by using a clustering algorithm based on the combination of K-mean and particle swarm algorithm in random search direction. The results show that only 23.3% of the students think that there is no duplication of content between vocational and undergraduate education, 89.6% of the students want to set the teaching content according to different needs, and the current talent cultivation for the articulation of vocational and undergraduate education suffers from poor wholeness and monotonous tendency. The proposed crawler search method and automatic clustering method show superior performance and can accurately extract teaching resources and process structured information. Finally, the employment-oriented talent cultivation model is proposed to actively explore the path of integrating vocational and undergraduate education and promote the development of vocational education.

Keywords: genetic algorithm, ant colony algorithm, particle swarm algorithm, K-mean algorithm, employment-oriented talent cultivation

1. Introduction

With the rapid development of social economy and the adjustment of industrial structure, vocational education has become a key link to improve the national scientific and technological innovation

✉ Corresponding author.
E-mail address: wyflcf@163.com (Y. Wang).
Received 20 July 2024; Revised 10 October 2024; Accepted 05 February 2025; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-200

ability and economic competitiveness. As the foundation of vocational talent cultivation, vocational education shoulders the dual mission of employment and further education [1-4]. With the construction of China's modern vocational education system and the arrival of the intelligent era, in the modern manufacturing industry, strategic emerging industries and other fields, there is a higher demand for the quality of technical and skilled personnel, and vocational education plays an important role in providing various types of highly skilled personnel [5-8]. However, there are still certain difficulties and challenges in the articulation between vocational education and undergraduate education. In order to better promote the cultivation of talents in the articulation between vocational education and undergraduate education, China has formed a relatively mature education articulation model, i.e., “3+4” talent cultivation model, in the course of long-term development [9-12].

The “3+4” talent cultivation model is the implementation of vocational vocational and undergraduate segmented cooperative cultivation, “seven-year segmented consistent undergraduate (Bachelor's degree) + preparatory technician (industry vocational qualification)” cultivation system, that is, fresh junior high school graduates in vocational vocational colleges and universities to complete the three years of learning and After passing the process assessment and transfer examination, fresh junior vocational school graduates will complete 4 years of undergraduate study in the corresponding majors of undergraduate colleges and meet the requirements for undergraduate degree conferment, and obtain the certificate of bachelor's degree [13-16]. The effective articulation pathways of “3+4” talent cultivation model include career orientation, integration of teaching resources, practical teaching and internship opportunities, cooperation with training institutions, employment services and school-enterprise cooperation [17-20]. Through the effective articulation of these pathways, the development of vocational and tertiary education can be better promoted, and the vocational competitiveness and employability of students can be improved [21-23].

The study conducts a questionnaire survey on schools of vocational education and undergraduate education in a city to explore the current situation of the articulation between vocational education and undergraduate education from the four dimensions of course objectives, content, implementation and evaluation. Subsequently, a crawler search method based on genetic and ant colony algorithms is designed and combined with K-mean and PSO algorithms to achieve automatic clustering of teaching resources on specific topics, a structured teaching resources collection and organization model based on focused crawler technology is proposed, the relative returns of different crawler search methods are evaluated to analyze the search performance of GA-ACO, and different clustering is performed on three different dimensional datasets. Comparison experiments of different clustering methods on three different dimensional datasets to explore the automatic clustering effect of the RVPSO-K algorithm in this paper. Finally, the employment-oriented talent cultivation model for the articulation of vocational education and undergraduate education is designed in terms of the catalog of professional specifications, the articulation of talent cultivation objectives, the articulation of curriculum standards, the articulation of professional teachers, and the articulation of teaching resources.

2. Survey and analysis of the interface between vocational and undergraduate education

Vocational and undergraduate education is an organic whole composed of various elements, including the basic content of curriculum objectives, curriculum content, curriculum implementation, curriculum evaluation and related constraints---knowledge, students, social environment (including examination and enrollment system). Through the questionnaire survey and data analysis, we survey the current situation of the articulation between vocational education and undergraduate education, and try to find out the current problematic features and problematic issues of the articulation between vocational education and undergraduate education, so as to provide practical experience and theoretical basis for the employment-oriented talent cultivation mode of the articulation between vocational education and undergraduate education.

2.1. Questionnaires

Based on the study of literature related to the articulation between vocational education and undergraduate education, a set of Questionnaire on the Status of Research on the Articulation of Undergraduate Education in Higher Vocational Institutions and a Questionnaire on the Status of Research on the Articulation of vocational education in Vocational Schools are compiled. The structure of the questionnaire consists of two parts. The first part of the questionnaire is the basic information of the teacher or student being surveyed, including eight aspects of basic information, such as name, gender, date of birth, the school in which they are studying, the type of study, the grade in which they are studying, the specialty in which they are studying, and their family situation. The second part of the questionnaire, i.e., the core content part of the survey questionnaire, includes four major survey dimensions, such as curriculum objectives, curriculum content, curriculum implementation, and curriculum evaluation, etc., for the interface between vocational and undergraduate education, and many specific and detailed smaller survey dimensions are designed under each of the four major survey dimensions.

The questionnaires were distributed in a number of vocational education and undergraduate education schools in a city respectively. The questionnaires were sent out 498 times, and 485 test questionnaires were retrieved, of which, 463 were valid test questionnaires, with a questionnaire validity rate of 95.46%. Among them, in terms of educational level, there were 240 undergraduate students, accounting for 51.84% of the total number of students, and 223 vocational students, accounting for 48.16% of the total number of students.

2.2. Statistical analysis of survey results

In this study, the EXCEL analysis tool was used to describe the results of the survey using a combination of frequency distribution and analysis of variance to describe the results, with the frequency distribution describing the overall picture and the analysis of variance describing the changes between the covariates.

2.2.1. Articulation analysis of curriculum objectives. The curriculum objective articulation between vocational education and undergraduate education is analyzed in six dimensions, namely, the curriculum objective adapting to the talent demand of the industry X1, the teaching and learning implementation of different student sources X2, the significant level of the job X3, the comprehensive ability enhancement X4, the professional skill enhancement X5, and the teaching and learning settings X6. The results of the analysis of course objective articulation are shown in Fig. 1, A~E indicate completely non-conformity, comparatively non-conformity, general conformity, comparative conformity, and complete conformity, respectively, the same below.

In the survey of course objectives adapting to industry talent demand, more than 88.9% of the intermediate and undergraduate students thought that the course objectives conformed to the demand for talents in the industry of socio-economic development. Only 26.9% of the vocational and undergraduate students thought that the teaching program varied according to the student population. The teaching program does not take into account the source of students and the actual situation, and the adoption of the same teaching program for different sources of students will inevitably cause disappointment in the learning effect for students with different learning bases or habits, thus affecting the overall quality of teaching.

In the item of “undergraduate graduates and vocational school graduates of the same specialty have obvious level distinction in employment positions”, 18.7% think that it is completely in line with the situation, 43.4% think that it is relatively in line with the situation, 20.8% think that it is generally in line with the situation, 10.9% think that it is relatively not in line with the situation, and 6.2% think that it is not in line with the situation at all. Compared with vocational schools, 82.2% of the students thought that there was a significant improvement in the professional knowledge and skills learned in undergraduate education, and 83.3% of the students thought that there was a significant improvement in comprehensive ability. Undergraduate institutions had a substantial improvement and significant results in students' professional knowledge and skills as well as comprehensive ability. 89.6% of the students agreed that different teaching contents should be set for students with different needs for further studies and employment. Curriculum objectives should meet the diversified development directions of students, and teaching contents should meet the diversified needs of students according to curriculum objectives.

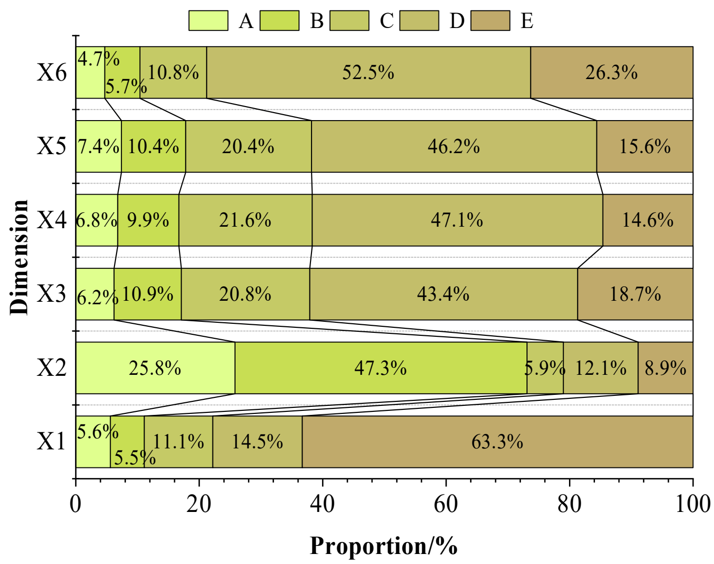


Fig. 1. The course goal link analysis results

2.2.2. Analysis of course content articulation. Course content articulation is analyzed in terms of duplication of content in specialized courses X7, strengthening of theories in specialized courses X8, strengthening of basic courses X9, rationalization of disciplinary settings X10, satisfaction with teaching materials X11, and progression of course levels X12. The results of course content articulation analysis are shown in Figure 2. In the item of “Compared with the intermediate courses, there is no duplication in the teaching content of undergraduate specialized courses”, 23.3% of the students think that it is completely consistent with the item. More than 70% of the students want to strengthen the theory and basic courses of the professional program. The percentage of students who were very satisfied and relatively satisfied with the teaching was 61.7%, and as many as 70.2% of the students thought that the professional disciplines were reasonably set up, and 83.6% of the students thought that the courses were progressive in level.

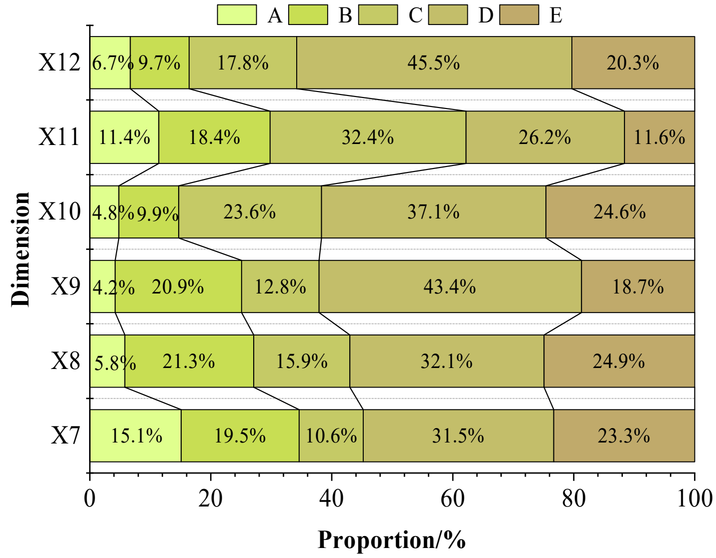
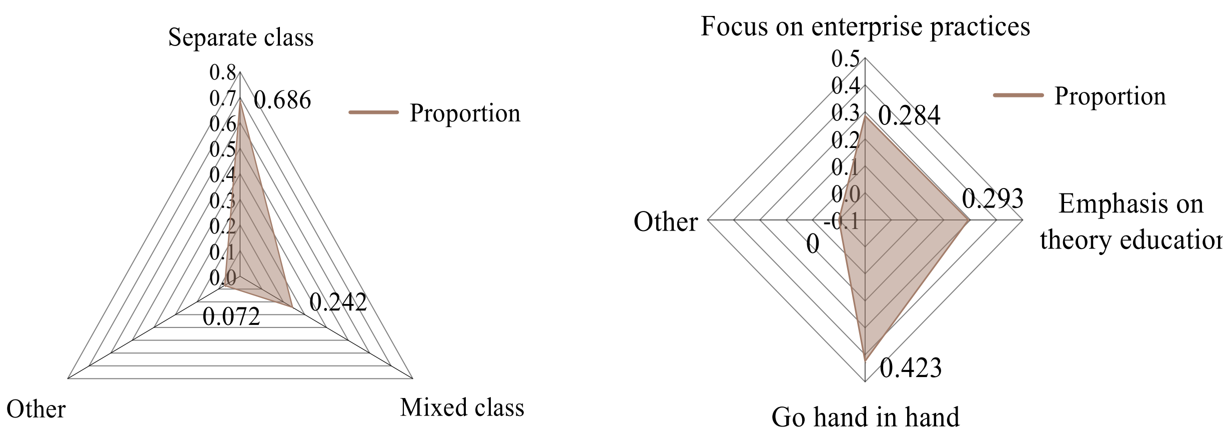


Fig. 2. The course content link analysis results

2.2.3. Articulation analysis of curriculum implementation. The results of the survey on student source management methods and school-enterprise cooperation training methods are shown in Figure 3(a) and Figure 3(b). In terms of student source management methods, separate classes amounted to 68.6%, mixed classes amounted to 24.2%, and others accounted for 7.2%. In terms of school-enterprise cooperation cultivation methods, 28.4% thought that the focus was on enterprise practice, 29.3% thought that the focus was on theoretical education, and 42.3% thought that it went hand in hand.



(a)Students' source management

(b)School and enterprise cooperation

Fig. 3. The results of the students' source management and the school and enterprise cooperation

The survey on the expected teaching focus of vocational education and undergraduate education is shown in Figure 4. 57.9% chose to impart knowledge, 86.4% chose to cultivate professional skills, 83.8% chose to cultivate students' vocational qualities, 63.7% chose to cultivate students' learning methods, 73.9% chose to cultivate students' outlook on life and values, 59.4% chose to cultivate students' interests and hobbies, and 71.3% chose to cultivate students' ability to adapt to society.

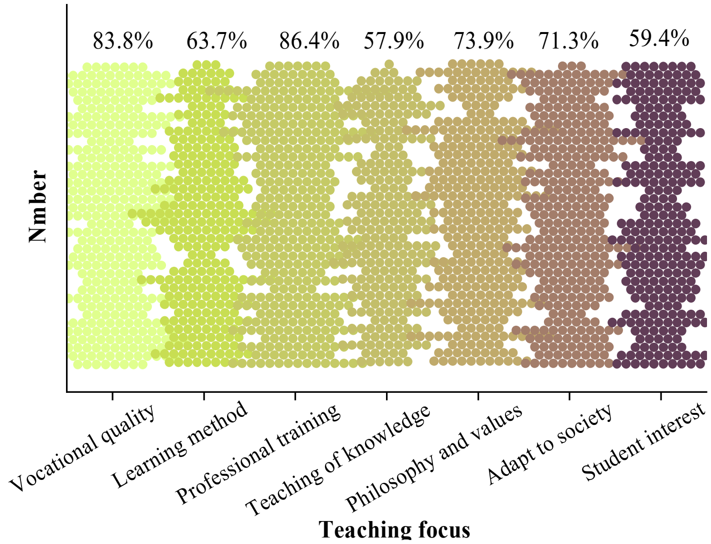


Fig. 4. The research result of the expectant teaching focus

2.2.4. Articulation analysis of course evaluations. The survey on the factors that have a greater influence on employment is shown in Figure 5. Test scores, certificates, professional operation skills, personal comprehensive ability, internship experience, school popularity and others accounted for 54.0%, 83.2%, 84.2%, 85.1%, 73.4%, 35.6% and 7.9% respectively.

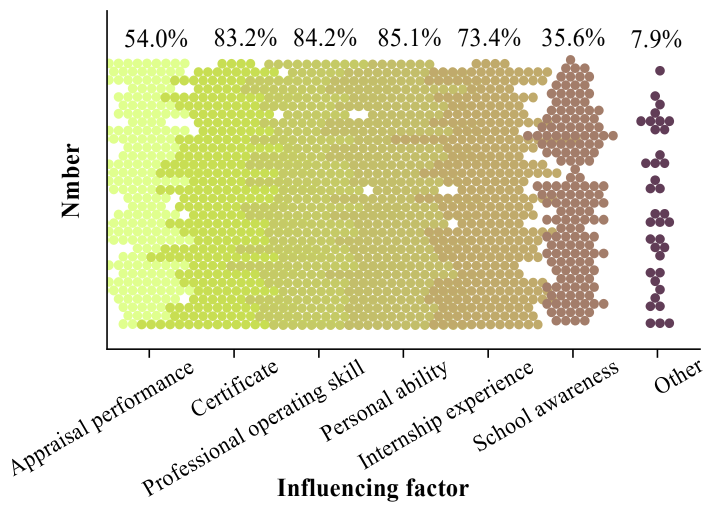


Fig. 5. The survey of factors that affect employment

The survey of course evaluation methods was conducted in three dimensions: assessment method, assessment subject and value orientation, and the results of the survey on course evaluation methods are shown in Figure 6. In the three dimensions of assessment method (written test-based), assessment subject (teacher assessment) and value orientation (encouraging evaluation), 86.6%, 77.9% and 88.2% of the students thought that they were in line with the three evaluation methods, respectively, and the percentage of the three evaluation methods was high.

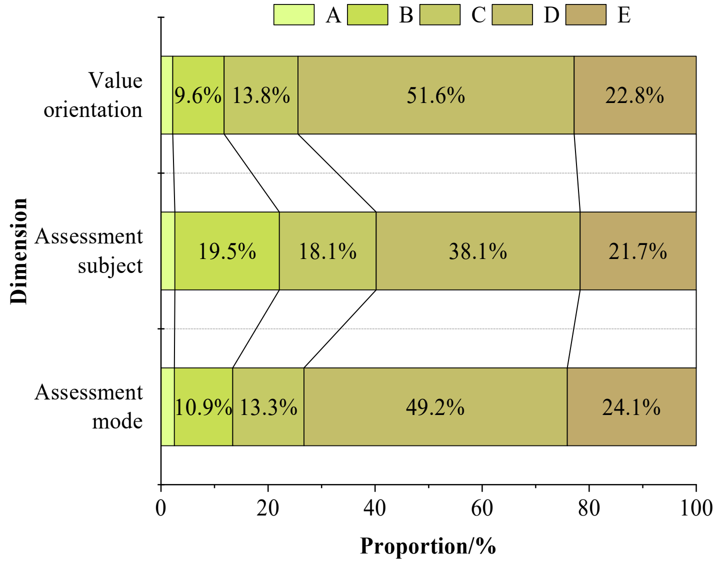


Fig. 6. The results of the course evaluation method

Comprehensive investigation and analysis of the current articulation of vocational education and undergraduate education, there are problems such as (1) poor stability of curriculum goal orientation and the tendency of disciplining the goal determination, (2) poor integration of curriculum content orientation and the tendency of liberalization of curriculum management, (3) poor openness of curriculum implementation orientation and the tendency of classroom-based curriculum teaching, and (4) poor formative orientation of curriculum evaluation and the tendency of monotonous evaluation mode. Therefore, it is necessary to build the articulation mode of vocational education and undergraduate education in line with the cultivation of employment-oriented talents, to build a curriculum system with significant coherence at different levels, to realize the common construction and sharing of teaching resources between vocational education and undergraduate education, and to promote the cultivation of employment-oriented talents.

3. Computationally optimized access to teaching resources

At present, national and local governments are paying more and more attention to the cultivation of employment-oriented talents, and projects such as professional teaching resource base of colleges and universities are emerging one after another, and the investment in teaching resources is increasing, and the conditions for running schools have been greatly improved. However, due to the fact that vocational education and undergraduate education belong to different management systems, the informatization teaching resources are independent of each other and cannot realize the common construction and sharing. In order to better realize the sharing of teaching resources between vocational education and undergraduate education, this chapter researches on the automatic collection and organization of teaching resources, and proposes a structured teaching resources collection and organization model based on focused crawler technology.

3.1. GA-ACO based focused crawler

3.1.1. Theme crawlers. Effective URL analysis models provide a better picture of the performance of topic crawlers. Distinguishing between topic crawlers differs mainly by how URL queues are determined. Based on the search strategy of web link structure, Hyperlink Induced Topic Search algorithm (HITS) is one of the algorithms for analyzing web links.

S_1, S_2, S_3 is a collection of URLs of web pages that satisfy the topic relevance requirement. These collections include collections of web pages that S_i can be linked to, and also collections of web pages that can be linked to S_1. V_c, V_m is the set of web pages that come from the S_1 crossover operation as well as the S_2 mutation operation. The formula is as follows:

$$V_n = V_m \cup V_c \tag{1}$$

3.1.2. GA-ACO algorithm. A genetic algorithm is a search heuristic used to solve optimization. The algorithm is a randomized search method evolved by drawing on the laws of evolution in biology, including the survival of the fittest and the genetic mechanism of survival of the fittest. Ant colony algorithm is a bionic algorithm from nature, which is an odds-based algorithm used to find an optimal path by simulating the foraging behavior of ant colonies.

Ant-Cycle model is a better model for global optimization in ACO algorithm. In Eq. (2), τ_{ij} denotes the pheromone intensity on path (I, j) at the moment t, and $\Delta\tau_{ij}^k$ denotes the amount of pheromone of the kth ant on path (I, j). The pheromone retention factor is $\rho(0<\rho<1)$ and the pheromone on each path is updated according to Eq. (2):

$$\tau_{ij}(t+1) = \rho\tau_{ij}(t) + \Delta\tau_{ij}, \Delta\tau_{ij} = \sum_{k=1}^{m} \Delta\tau_{ij}^k(t) \tag{2}$$

Let Z_k be the length of the path traveled by the knd ant in the model, which can be obtained as $\Delta\tau_{ij}^k(t) = Q / Z_k$, where Q is a constant. η_{ij} is the desired degree of path (I, j), which is usually $1/d_{ij}$, where d_{ij} is the length of path (I, j), β is a parameter indicating the importance of pheromone, and α denotes the degree to which the pheromone accumulated by the ant during its travel plays a role in its choice of path. The probability of the ant k shifting from position i to j at the moment of t is represented by equation (3):

$$p_{ij}^k(t) = \begin{cases} \dfrac{\left[\sigma_{ij}(t)\right]^{\alpha}\left[\eta_{ij}\right]^{\beta}}{\sum\limits_{j\in U}\left[\tau_{il}(t)\right]^{\alpha}\left[\eta_{il}\right]^{\beta}}, j \in U \\ 0 \end{cases} \tag{3}$$

The combination of genetic and ant colony algorithms is often referred to as genetic-ant colony algorithm and the basic idea is to utilize the strengths of both algorithms to overcome their weaknesses. The aim of this algorithm is for better accuracy as well as lower time loss. The main process is to make full use of the advantages of fast, stochastic and global optimization of the genetic algorithm and the parallelism, positive feedback characteristics of the ACO algorithm to achieve a highly efficient and low-loss system. The process of combining genetic algorithm and ant colony algorithm is shown in Fig. 7. The fitness function is used to solve the combinatorial problem due to the use of real number metric computation and is determined by the projection function. The real numbers generated by the stochastic function are used as the encoded population. Sequence crossover algorithm is used here.

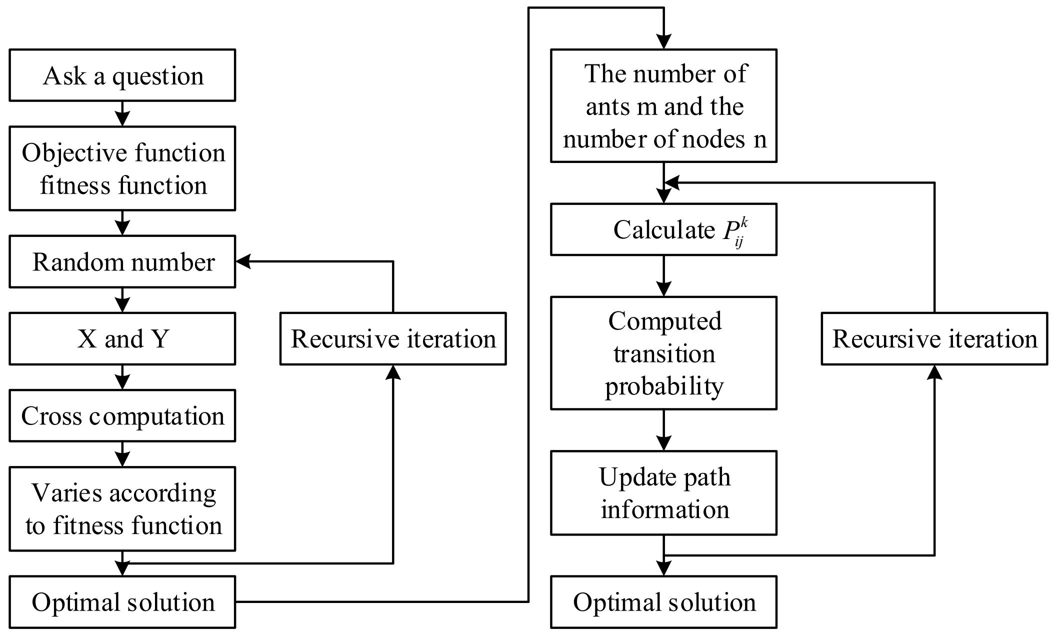


Fig. 7. The process of combining genetic algorithms and ant colony algorithms

3.1.3. Page Preprocessing. Before analyzing the relevance of a page document, the web page needs to be preprocessed to a certain extent, and the classical vector space model is chosen for the preprocessing of the web page. After the basic thesaurus is built, feature words are extracted from the content of the web page to build the web page information vector, so that the relevance operation can be carried out according to the vector space model. The TF-IDF algorithm is used to extract the topics from the information of the web page, TF denotes the number of words that appear in the text for word t, and the IDF operation is based on the following formula:

$$DF(t) = \frac{N}{M} \tag{4}$$

where N is the total number of all text collections and M is the number of texts that contain the word t in the document collection.

The concept of vector space modeling is used to evaluate the relevance of web page content and topic, and the given evaluation value determines the degree of relevance of web page content and topic. The traditional TF-IDF positional weighting algorithm for a page A has a keyword weighting vector in a n-dimensional vector space: $\overline{A} = (a_1, a_2, \cdots, a_n)$ where a_i represents different weighting coefficients. A set of topic vectors $\overline{P} = (p_1, p_2, \cdots, p_n)$, p_i in the vector space represent different weighting coefficients, and the current expression of the degree of relevance of the page content and topic in the vector space is:

$$\sin(\overline{A}, P) = \frac{\sum_{i=1}^{n} \alpha_i p_i}{\sqrt{\sum_{i=1}^{n} a_i^2 \sum_{i=0}^{n} \sum_{i=1}^{n} p_i^2}} \tag{5}$$

The number resulting from this expression should be between [-1, 1], and if positive, the larger the value, the more relevant the content is to the topic. If the number is negative, the smaller the value is, the more contrary the content is to the topic. If the result is equal to 0, it means that the content is completely irrelevant to the topic, and the page can be abandoned.

3.1.4. Web site analysis model. The contents of web pages are usually written by editors and then uploaded, which leads to web pages with many contents that are not related to the topic.In this paper, we design a new URL analysis model with the following main steps:

1) Manually create description vectors.

2) Preprocess the web pages.

3) Apply the fitness algorithm and calculate the crossover probability p_{ij}^{k}.

4) Calculate the fitness value.

5) Use selection, expansion, crossover, and mutation algorithms.

6) If the fitness is less than $\tau_{ij}(t)$, or if there is still no match after iterations have been completed, return to step 5, otherwise end.

7) The result of the fitness function value will be determined by the positive feedback system of the ACO algorithm and the combined genetic-ant colony algorithm, and then the vector can be modified by the following equation:

$$V_{new} = \alpha * V_{old} + \beta * \frac{\sum D_{rel}}{N_{rel}} - \gamma * \frac{\sum D_{nrel}}{N_{nrel}} \tag{6}$$

where, V_{new} and V_{old} are the old and new topic vectors respectively. α, β, γ is the weight value, D_{rel} and D_{nrel} are the relevant as well as irrelevant documents, and N_{rel} and N_{nrel} are the total number of documents in these two categories.

8) Ranking of URL queue based on relevance.

3.2. Automatic clustering based on RVPSO-K

Using clustering module can automatically organize the teaching resources on specific topics collected from the Web by focus crawlers, reduce manual labor and improve the efficiency of organization. The K-means algorithm and particle swarm algorithm are combined to propose an improved RVPSO-K clustering algorithm based on search direction.

3.2.1. K-means algorithm. The K-means method first randomly selects K objects, each of which serves as an initial cluster center, then assigns the remaining objects to the most similar clusters based on the distance of each cluster center for that object, and next updates the mean value for each cluster. This cycle is repeated until all objects are selected. And so that there is a high degree of similarity within each cluster and a high degree of dissimilarity between clusters.

1) Randomly select k objects as the centroids of the initial clusters.

2) Calculate the distance between the unclassified objects and the center point of each cluster, and find the distance between them according to the following formula:

$$d_{ij} = \left\| x_i - c_j \right\| \tag{7}$$

Classify that x_i point into the closest cluster.

3) Calculate the new clustering centers, i.e., update the mean of the clusters and calculate the average value of the objects in each cluster:

$$\tilde{c}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i \tag{8}$$

n_j is the number of samples in the input samples that belong to the j nd cluster at the time of determining the center of the network, here x_i is the sample that belongs to the j th cluster.

4) Calculate $\varepsilon = \left\| \tilde{c}_j - c_j \right\|$, if E is less than the given value, the cluster center is stable, otherwise $c_j \leftarrow \tilde{c}_j$ go to step 2.

3.2.2. Particle Swarm Algorithm. Particle Swarm Optimization (PSO) algorithm is an evolutionary computation technique based on swarm intelligence methods, which first originated from the study of the foraging behavior of bird flocks.

It is described mathematically by assuming that in a d-dimensional search space, there are m candidate solutions of particles representing potential problem solutions, and the degree of superiority of the solutions is determined by a fitness function. The particles search for new solutions by continuously adjusting their positions Z_i. The position of each particle represents a solution to the problem, and each particle remembers the best solution it has searched for, denoted $P_i = (p_{i1}, p_{i2}, \dots, p_{id}), i = 1, 2, \dots, m$, as well as the best position experienced by the entire swarm of particles, denoted $P_g = (p_{gl}, p_{g2}, \dots, p_{gd})$, i.e., the optimal solution currently searched for, where, $g = 1, 2, \dots, m$. Each particle has a velocity variable, denoted $V_i = (v_{i1}, v_{i2}, \dots, v_{id}), i = 1, 2, \dots, m$. Then:

$$V_i' = \omega V_i + c_1 rand()(P_i - Z_i) + c_2 rand()(P_g - Z_i) \tag{9}$$

where ω is the inertia weight, a larger ω facilitates jumping out of the local maxima, a smaller ω facilitates convergence of the algorithm, c_1, c_2 is a non-negative constant called the learning factor, which is a parameter that regulates the relative importance of P_i and P_g, and $rand()$ is the random number generating function. This gives the next position at which the particle moves:

$$Z_i' = Z_i + V_i \tag{10}$$

From equations (9) and (10), it can be seen that the direction of movement of the particle is determined by three parts, the first part is the original velocity of the particle V_i illustrates the current state of the particle, the second part is the cognitive part, the distance between the particle and its own optimal experience $(P_i - Z_i)$, which indicates the particle's own thinking, and the third part is the social part, the distance between the particle and its optimal experience with the group $(P_g - Z_i)$, which guides the particle to converge to the region of the current optimal solution. The three parts together determine the spatial search ability of the particle.

3.2.3. RVPSO-K algorithm. Set the nst sample x_n to be a vector of dimension d, denoted $x_n = (x_{n1}, x_{n2}, \dots, x_{nd}), n = 1, 2, \dots, N$, and let $x_\alpha = (x_{\alpha 1}, x_{\alpha 2}, \dots, x_{\alpha d})$ and $x_\beta = (x_{\beta 1}, x_{\beta 2}, \dots, x_{\beta d})$ be any two sample vectors for which the ith particle belongs to the kth class in the clustering process. There are random numbers $\overline{r}_1, \overline{r}_2, \dots, \overline{r}_d \in [0,1)$ forming matrix $R = [\overline{r}_1, \overline{r}_2, \dots, \overline{r}_d]$, $V^k_{i\_random}$ denoting the random velocity variable of class k for the ith particle, $V^k_{i\_max}$ being the velocity threshold of class k for the ith particle, and V_{i_random} consisting of K $V^k_{i\_random}$. There are the following relational expressions:

$$\left| V^k_{i\_\max} \right| = \left\{ \min \left| x^k_{\alpha 1} - x^k_{\beta 1} \right|, \min \left| x^k_{\alpha 2} - x^k_{\beta 2} \right|, \dots, \min \left| x^k_{\alpha d} - x^k_{\beta d} \right| \right\} \tag{11}$$

$$V^k_{i\_random} = R.* V^k_{i\_\max} \quad l = 1, 2, \dots, d \tag{12}$$

That is, $V^k_{i\_random}$ in the RVSPO-K algorithm is given by the following equation:

$$V^k_{i\_random} = \left\{ \overline{x}_1 \min \left| x^k_{\alpha 1} - x^k_{\beta 1} \right|, \overline{r}_2 \min \left| x^k_{\alpha 2} - x^k_{\beta 2} \right|, \dots, \overline{r}_d \min \left| x^k_{\alpha d} - x^k_{\beta d} \right| \right\} \tag{13}$$

The improved expression for the particle search velocity is as follows, i.e., Eq. (9) is redefined as:

$$V_i = \omega V_{i\_random} + c_1 rand()(P_i - X_i) + c_2 rand()(P_g - X_i) \tag{14}$$

If it is desired to have different particles searching in other possible search directions, then the first part of Eq. (10) can be changed to a random variable V_{i_random} instead of inheriting the speed from the last iteration, see Eq. (14).

The process is described as follows:

1) Initialize the particle swarm, i.e., randomly set the initial position Z_i and initial velocity V_i of each particle, for each particle Z_i, according to the nearest neighbor law, to determine the clustering division for that particle, calculate the initial fitness value $F(x_i)$ of each particle according to the fitness function, and initialize P_i, P_g.

2) Adjust the velocity and position of the particles according to equations (14) and (10).

3) For the new generation of particles, optimize according to the following K-mean algorithm:

(1) According to the clustering center of the particle, follow the nearest neighbor law to determine the cluster division for that particle.

(2) Calculate the new clustering center according to the clustering division.

4) Calculate the fitness value $F(z_i)$ for each particle based on the fitness function.

5) For each particle, compare its fitness value with the fitness value of the best position P_i it has experienced, and if better, update P_i.

6) For each particle, compare its fitness value with the fitness value of the best position P_g experienced by the population, and if better, update P_g.

7) If the end condition is reached (sufficiently good position or maximum number of iterations), end and get i.e. the optimal solution searched so far P_g. Otherwise go to step 2.

3.3. *Data extraction*

The data extraction module is to extract and organize the subject-related web pages collected by the crawler according to format and category. The data extraction module formulates corresponding extraction rules and patterns based on the characteristics of each subject domain and the learning of the sample web pages labeled by experts, and generates the corresponding wrapper libraries. The data extracted by the wrappers are stored as entity files according to the pre-set format, and XML is used to represent them. It is then processed by the structured repository engine module and stored in the structured repository. With the filling of new structured files, retraining is performed periodically or quantitatively to add the new feature items to the feature item library of the classification submodule.

3.4. *Structured Teaching Resource Library*

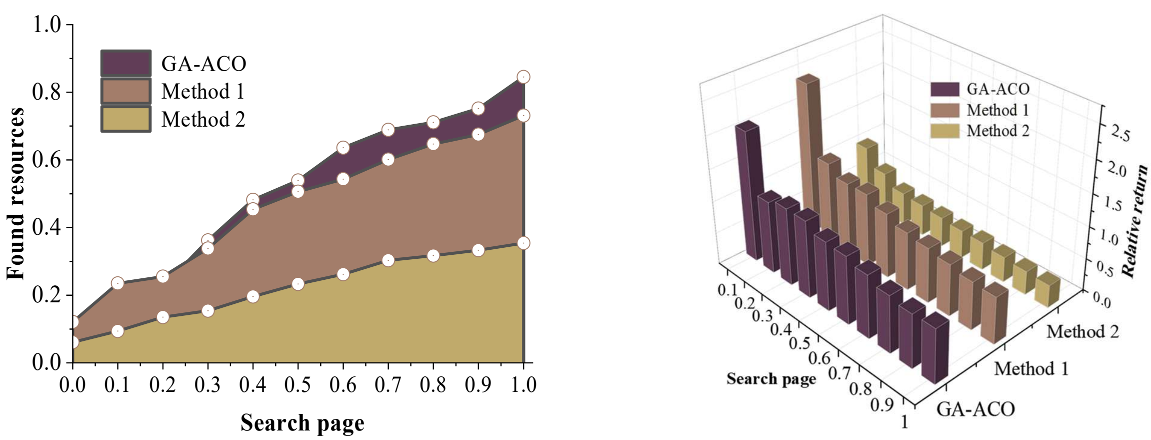
Structured teaching resource library is based on a certain classification system structure. The classification system structure includes concepts, ontologies, knowledge points and so on. At present, most structured teaching resource libraries are built on the commonly used classification system structure of specific disciplines, usually a tree structure. Corresponding functions are provided at the nodes of each level of the catalog, such as the functions of adding, deleting, modifying, copying, moving up and down, exporting, etc. of the nodes, the function of restricting the total amount of resources, and the function of restricting the type of application of the resources.

3.5. *Experimental results and analysis*

3.5.1. Analysis of GA-ACO experiments. In order to verify that the GA-ACO based topic crawler search strategy has better global search capability and adaptivity than the traditional breadth-first algorithm and best-first based search strategy, an educational topic crawler is built on top of the Heritrix crawler.The Heritrix crawler is extensible and customizable. The Heritrix crawler is extensible and customizable. By defining an ACOCrawler plugin to crawl web pages of a specific topic. The theme is "English teaching resources", and three educational websites are selected as seed sets.

The system runs for 24 hours, and a total of 3,280,000 web pages and resources are crawled. In order to facilitate the comparison, the search results based on the breadth-first algorithm (Method 2) and the best-first search algorithm (Method 1) are tested separately, and the number of web pages and resources about physics teaching resources searched by the crawlers implemented by the three search algorithms are counted, and the relative return rate is used to evaluate the performance of the crawlers. The performance of the crawler is measured by the ratio of the percentage of pages with relevant topics found to the percentage of pages searched at a certain point in time.

The performance comparison of the three search algorithms is shown in Fig. 8, with (a) and (b) indicating the changes of the proportion of discovered resources and relative return with the proportion of searched pages, respectively. Among the three search strategies, the breadth-first algorithm has a lower performance than the other two "heuristic" algorithms. The best-first algorithm found more than 54.3% of relevant ELT resources after visiting 60% of the pages, with a relative return of 90.5%. The GA-ACO-based search strategy found more than 63.6% of the relevant English teaching resources after visiting 60% of the pages, with a relative return of 106%, which shows the superiority of the crawler search strategy based on the genetic algorithm and ACO algorithm. The performance of the GA-ACO based search strategy is relatively significant, except that its discovery ability is slightly lower than that of the search algorithm based on the best-first strategy at the beginning of the search, in the subsequent searches, the performance of the algorithm in this paper is significantly higher than that of the search algorithm based on the best-first strategy.



(a) Found resources (b) Relative return

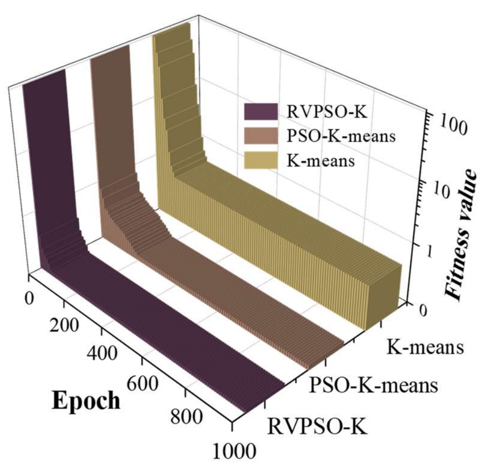
Fig. 8. Comparison of three search algorithm performance

3.5.2. RVPSO-K experimental analysis. The experiments are conducted to evaluate the clustering performance experiments from three data sets on low dimensional data, high dimensional data and multidimensional samples high dimensional respectively. The experimental results of the algorithm RVPSO-K are compared with those of PSO-K-means and K-means.

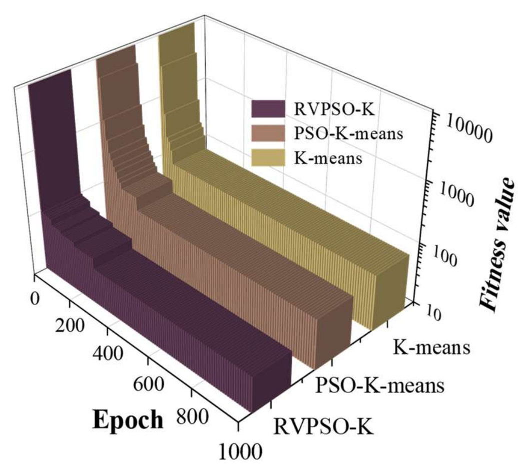
The experimental data use three types of datasets commonly used in similar studies: the Iris plant sample dataset for the low-dimensional feature dataset, the Wine dataset for the high-dimensional feature dataset, and the Wisconsin Breast Cancer dataset for the multi-sample high-dimensional data. In the experiment, the value of the adaptation value function $f()$ is the value of the function corresponding to the current position.

Comparison of the data clustering performance of the three algorithms is shown in Fig. 9, (a)~(c) correspond to the experimental results of the Iris, Wine and Cancer datasets.The K-means algorithm has the fastest convergence speed, but it is easy to fall into the local optimum. And the PSO-K-means clustering algorithm makes the algorithm's global optimal solution search ability improved by combining it with the traditional particle swarm algorithm. However, as the number of iterations increases and the role of K-means gradually decreases, the particle swarm focuses and gradually converges, and then falls into the local optimum. Different from the above two methods, in this paper, a clustering algorithm based on the combination of K-mean and PSO algorithm with random search direction jumps out of the local optimum and approaches the global optimum.

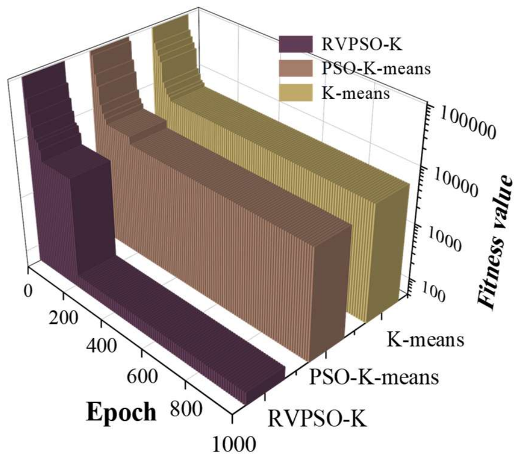
The comparison results of the three algorithms for clustering stability of data are shown in Figure 10. As can be seen from the variance values and means of the various algorithms, the K-means algorithm, due to the large influence of the initial solution, has mean values of 27.03, 121.63 and 689.71 on the Iris, Wine and Cancer datasets, respectively, which are larger than those of the RVPSO-K algorithm and the PSO-K-means algorithm, which is less stable and prone to fall into the local optimum. The PSO-K-means hybrid clustering algorithm has better stability because it introduces the K-means operation through each iteration, which makes the algorithm have a better ability to search for the global solution in the initial stage, and thus the algorithm itself reduces the dependence on the distribution of the initial solution, and the algorithm has better stability. The RVPSO-K algorithm in this paper has the smallest mean on Iris, Wine and Cancer datasets, which are 12.84, 59.18 and 203.52 respectively.The algorithm proposed in this paper has obvious advantages for high dimensional and diversified high-dimensional feature data processing, especially for the complex diversified high-dimensional feature data, and its performance has a great improvement compared with other algorithms.



(a)Iris dataset



(b)Wine dataset



(c)Cancer dataset

Fig. 9. Clustering performance comparison of three algorithm data

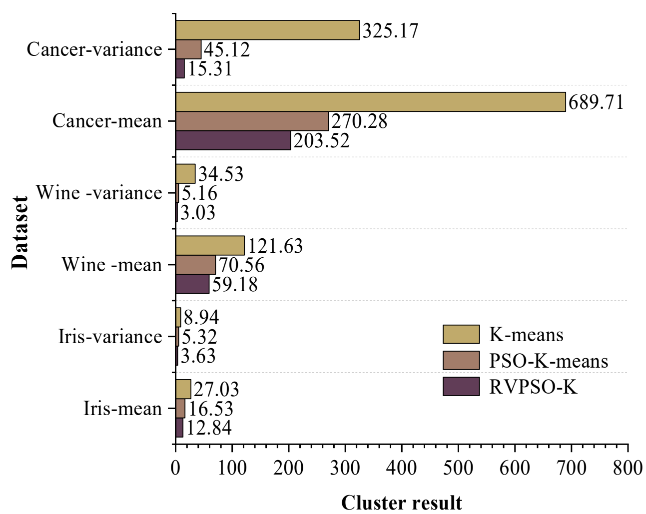


Fig. 10. The comparison results of the three algorithms

4. Employment-oriented talent training model

Based on the previous investigation and analysis of the articulation of vocational and undergraduate education and the method of acquiring teaching resources, this chapter discusses the employment-oriented talent cultivation model for the articulation of vocational and undergraduate education in terms of the professional catalog specifications, the articulation of talent cultivation objectives, the articulation of curriculum standards, the articulation of professional teachers, and the articulation of teaching resources. The employment-oriented talent cultivation model is shown in Figure 11.

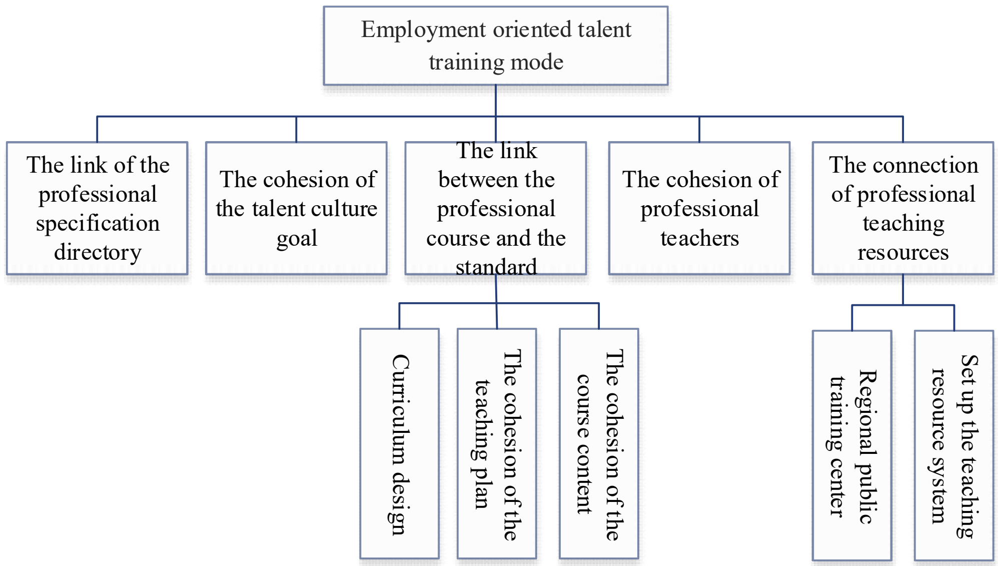


Fig. 11. Employment oriented talent training mode

4.1. *Articulation of the catalog of professional specifications*

Improving the employment-oriented talent training system and establishing a path for professional degree advancement in vocational and undergraduate education. Follow the "three-three-three" operation mechanism for revising the catalogs of vocational and undergraduate majors, and innovate the rules for revising the catalogs of articulated majors. Firstly, the three bases for revising the catalog articulation are solidified: the current professional catalog is the basis, the Occupational Classification Dictionary is the basis, and the National Occupational Qualification Certificate is the basis. Second, abide by the three principles of professional articulation catalog compilation: the principle of policy orientation, the principle of adapting to industrial development, and the principle of giving equal importance to science and application. Thirdly, it follows the three phases of compiling the professional articulation catalog: the research and demonstration phase, the determination and implementation phase, and the feedback and optimization phase.

4.2. *Articulation of talent training objectives*

The objective of talent training in vocational education is mainly to provide basic knowledge, technology and skills education for junior and senior high school graduates, and to provide a reserve source of students with a professional foundation for undergraduate education. Undergraduate education is to cultivate high-level technical and skill-oriented talents for enterprises and institutions or self-employed entrepreneurs. At the stage of convergence between vocational and undergraduate education, students' cultural and professional skills are mapped out, and targeted talent training programs are formulated to improve students' professional theory and technical skills. Undergraduate education should adhere to the intermediate level as the basis, according to the corresponding level of students' cultural and professional skills, improve students' professional theory, professional technology and practical skills level, and truly realize the docking of intermediate and undergraduate education in the goal of cultivating employment-oriented talents.

4.3. *Articulation of specialized course standards*

4.3.1. Curriculum design. According to the cognitive characteristics of students, the development of vocational abilities is carried out by utilizing the "layer-by-layer advancement technique", and the curriculum is divided into vocational and undergraduate education courses on the basis of the layer-by-layer advancement technique according to the vocational abilities required by the cultivation goals of vocational and undergraduate education, so that curriculum differentiation is carried out progressively from a single to a comprehensive, from simple to complex, and the cumulative curriculum is developed in a cumulative ability-based curriculum. Teaching model for curriculum development.

4.3.2. Articulation of teaching programs. In order to make the smooth connection between vocational education and undergraduate education, a spiral progression of teaching links is established. The integration of science and practice, single practical training, comprehensive practical training, top internship and other teaching links are scientifically and rationally carried out throughout the whole process of talent cultivation in accordance with the requirements of skill enhancement and quality development.

4.3.3. Curriculum content articulation. The curriculum system based on the analysis of typical work tasks and the progression of vocational job competence is the main line, making a distinction in the content and degree of course learning. From single to comprehensive, from experience to strategy, from popularization of technology to high-tech, from operation to management and other hierarchical characteristics. Sending samples can avoid the waste of teaching resources in vocational education due to the duplication of course content, and at the same time, it can enable students to further improve their vocational literacy and vocational ability.

4.4. Articulation of specialized faculties

First, the optimization of the structure of teacher competence. Teachers at the intermediate level have higher requirements in terms of vocational quality and basic engineering practice ability, focusing on teaching grass-roots positions at the front line of production, management and service. Teachers at the undergraduate level have higher requirements for comprehensive practical ability and applied technology research and development ability to meet the teaching requirements for students' developmental positions. Secondly, establish an interactive exchange mechanism for teachers. Establish a platform for mutual recruitment and exchange of teachers, and implement a system of exchange, posting, mutual recruitment and mutual visits of middle and high school teachers in the region. To realize “vertical integration” in the teacher management system and “horizontal integration” in the mutual exchange of teachers.

4.5. Articulation of professional teaching resources

4.5.1. Regional public practical training centers. Public practical training centers can be led by the relevant government departments, integrating the stock of resources of various institutions, setting up several sub-centers in accordance with the major categories of specialties, establishing a management body to exercise the duty of coordination and monitoring, and the sub-centers and institutions to exercise the duty of operation and management, and accepting vocational and undergraduate colleges and universities as well as various kinds of social training in the region.

4.5.2. Establishment of a teaching resource system. Cooperating institutions jointly build a digital professional teaching resource library and an online learning platform, and the vocational and undergraduate institutions respectively determine the administrator's rights to be open to teachers and students. The professional resource library establishes modules such as professional information, cultivation program, curriculum standard, teaching design, case video, resource expansion, etc. It collects and organizes teaching resources by using teaching resource acquisition methods optimized based on intelligent algorithms, and builds a digital teaching resource information management service system. It is jointly constructed by intermediate teachers and undergraduate teachers, and students can study through the assigned permissions. Enterprise practice project cases are introduced to promote the sharing and updating of digital teaching resources. This open and co-built shared teaching resource platform can provide support for teacher-student exchanges, students' independent choice of learning content, and the promotion of individuality and characteristic development.

5. Conclusion

The articulation between vocational and undergraduate education has become an important issue affecting the healthy development of vocational education endeavors. This paper investigates the articulation between vocational education and undergraduate education and finds the problems that affect and restrict the articulation between vocational and undergraduate education. Based on this, it constructs a computationally optimized teaching resource acquisition method and discusses the employment-oriented talent cultivation model for the articulation between vocational and undergraduate education in a comprehensive analysis of the article.

The interface between vocational education and undergraduate education has the problems of poor stability of curriculum objectives, poor integration of curriculum content, poor openness of curriculum implementation and poor formative evaluation of curriculum. 89.6% of the students think that different teaching contents should be set according to the different needs of further education and employment. More than 70% of the students hoped to strengthen the theory and basic courses of professional courses.

The method of collecting and automatically organizing teaching resources based on focused crawler technology is proposed for the construction of teaching resources. Relative to other search algorithms, this paper's crawler search algorithm based on GA-ACO has the highest relative return rate, which is as high as 106% when the proportion of searched pages is 60%. Meanwhile, the RVPSO-K based automatic clustering algorithm has superior clustering stability, and the fitness values on different datasets are lower than those of the comparison methods.

The proposed employment-oriented talent cultivation model includes professional specification catalog articulation, talent cultivation goal articulation, curriculum standard articulation, professional faculty articulation and teaching resources articulation. Through the design of employment-oriented talent cultivation model, it points out the direction and the way for institutions to carry out the articulation of vocational and undergraduate specialties in modern vocational education.

Funding

This paper is supported by the Management Culture Research Institute at Xiamen Institute of Technology. The project is titled “One-Stop Digital and Intelligent Community Employment Support for Students Under the leadership of the Communist Party of China (CPC): a Perspective Based on Field Theory” (Project Number: GYY2024009).

This paper is a research outcome of the 2024 Xiamen Vocational Education Research Project titled “Study on the Learning Status of the Pathway from Vocational Vocational Education to Higher Education: A Case Study of the First Pilot Class in Xiamen” (Project Number: ZJKT202410).

References

[1] Sargent, T., & Xiao, Y. (2018). Teaching reform in Chinese undergraduate education. Chinese Education & Society, 51(4), 245-247.

[2] Yin, H., & Ke, Z. (2017). Students' course experience and engagement: An attempt to bridge two lines of research on the quality of undergraduate education. Assessment & Evaluation in Higher Education, 42(7), 1145-1158.

[3] Wang, G., & Wang, Z. (2023). Vocational education: a poor second choice? A comparison of the labour market outcomes of academic and vocational graduates in China. Oxford review of education, 49(3), 408-427.

[4] Guo, D., & Wang, A. (2020). Is vocational education a good alternative to low-performing students in China. International Journal of Educational Development, 75, 102187.

[5] Wang, A., & Guo, D. (2019). Technical and vocational education in China: enrolment and socioeconomic status. Journal of Vocational Education & Training, 71(4), 538-555.

[6] Wang, G., Zhang, X., & Xu, R. (2023). Does Vocational Education Matter in Rural China? A Comparison of the Effects of Upper-Secondary Vocational and Academic Education: Evidence from CLDS Survey. Education Sciences, 13(3), 258.

[7] Xue, E., & Li, J. (2022). Exploring the type-based vocational education system: Insights from China. Educational Philosophy and Theory, 54(10), 1670-1680.

[8] Liu, S., & Hardy, I. (2023). Understanding Chinese national vocational education reform: a critical policy analysis. Journal of Vocational Education & Training, 75(5), 1055-1077.

[9] Marginson, S. (2016). High participation systems of higher education. The Journal of Higher Education, 87(2), 243-271.

[10] Li, L., & Guanghua, W. (2021, August). The Connection between Secondary Vocational Education and Applied Undergraduate Education Research on Students' University Learning Investment. In 2021 International Conference on Big Data Engineering and Education (BDEE) (pp. 120-125). IEEE.

[11] Oviawe, J. I., Uwameiye, R., & Uddin, P. S. (2017). Bridging skill gap to meet technical, vocational education and training school-workplace collaboration in the 21st century. International Journal of vocational education and training research, 3(1), 7-14.

[12] Wei, X., Hou, Z., & Li, W. (2016, December). Constructing the Curriculum System of Cultivating High-end Skilled Talents through Secondary Vocational Education and Undergraduate Teaching. In 2016 6th International Conference on Mechatronics, Computer and Education Informationization (MCEI 2016) (pp. 1453-1457). Atlantis Press.

[13] Korchak, E. S. (2017). Implementation of Supplementary Educational Programs into Secondary Vocational Education System. Contemporary Problems of Social Work, 3(4), 108-115.

[14] Feng, W., Li, Z., & Pan, Y. (2022). Study on the "Secondary and Undergraduate" Comprehensive Training Curriculum System of Tourism Management Vocational Education. Tourism Management and Technology Economy, 5(1), 12-15.

[15] Khairullina, E. R., Valeyev, A. S., Valeyeva, G. K., Valeyeva, N. S., Leifa, A. V., Burdukovskaya, E. A., & Shaidullina, A. R. (2015). Features of the programs applied bachelor degree in secondary and higher vocational education. Asian Social Science, 11(4), 213.

[16] Shuwu, L., Jiajue, F., Libao, L., Xiaoli, W., Yanhai, W., Hongmei, L., ... & Yanmei, D. (2020, September). Research on Cultivating Personnel for the Cohesiveness between Secondary Vocational School and Applied Undergraduate College-take A college as an example. In 2020 International Conference on Modern Education and Information Management (ICMEIM) (pp. 142-145). IEEE.

[17] Field, S., & Guez, A. (2018). Pathways of progression: Linking technical and vocational education and training with post-secondary education. UNESCO Publishing.

[18] Jiang, X. (2017, November). The Problem Analysis of the Education Connection about" 3+ 4" Electrical Engineering and Automation Profession of Secondary Vocational-Undergraduate. In International Conference on Education Innovation and Social Science (ICEISS 2017) (pp. 354-357). Atlantis Press.

[19] Li, L., & Han, S. (2022). Research on "Connecting with Secondary Vocational and Undergraduate Transition Examination" of Preschool Education Specialty Based on Ability Orientation. Mobile Information Systems, 2022(1), 6998394.

[20] Yuen, M., Yau, F. S., Tsui, J. Y., Shao, S. S., Tsang, J. C., & Lee, B. S. (2019). Career education and vocational training in Hong Kong: Implications for school-based career counselling. International Journal for the Advancement of Counselling, 41, 449-467.

[21] Zhao, Z., & Liu, Y. (2019). Vocational education and training in economic transformation in China. The Wiley Handbook of Vocational Education and Training, 495-511.

[22] Szőköl, I., Pupala, B., & Kováč, O. (2023). Key competences of a secondary vocational school graduate. In EDULEARN23 Proceedings (pp. 7-15). IATED.

[23] Cheng, H. (2023). Research on Vocational Education Personnel Training Under the Strategic Background of Guangdong-Hong Kong-Macao Greater Bay Area--Taking Zhaoqing City as an Example. Economics & Management Information, 1-10.