International Journal for Housing Science and Its Applications
Publish August 4, 2025. Volume 46, Issue 3 Pages 2360-2370

https://doi.org/10.70517/ijhsa463197

Analysis of ESP Demand and Vocational Skills Cultivation Path of Chinese+Vocational Skills Courses Based on Big Data Algorithms--A Case Study of a Guangdong University

Lu Meng[1,*]
[1] University International College, Macau University of Science and Technology, Macau, 999078, China
Corresponding authors: (e-mail: trudymong1413@gmail.com).

Abstract With the continuous development of vocational education, accurately grasping students' ESP needs has become a necessary path for optimizing students' vocational skills training path. Based on the Apriori algorithm, this paper mines the association rules between the vocational skills learning data of students in a university in Guangdong, reveals the intrinsic connection between students' vocational skills and students' ESP needs for Chinese+vocational skills courses, and designs the students' vocational skills cultivation path based on the TPACK framework. Then, social network analysis was used to visualize the relationship data in the vocational skills cultivation path network, analyze the differences of different individual centrality indexes in different student nodes, and explain the key roles played by different student nodes in the vocational skills cultivation path. The study shows that based on Apriori algorithm, we can effectively identify students' Chinese+vocational skills related features, and the improvement between “public liability” and “litigation” is as high as 1474.23%. 1474.23%, and there is a high degree of intrinsic connection between the two. Based on the network architecture analysis of vocational skills training path based on the social network analysis method, it is accurately identified that among the different students, the point degree centrality of the students who are in the core position in the class is higher, and the leadership role of this kind of students can be given full play to, and the optimization of the vocational skills training path can be provided to the students in a targeted way.

Index Terms Apriori algorithm, association rule mining, TPACK framework, social network analysis, vocational skills training path network

I. Introduction

As we all know, English learning still occupies an extremely important position in China's basic education and higher education stages. Facing such a large group of learners and foreign language teachers, the foreign language education sector has been exploring and experimenting with different educational concepts and teaching modes [1]. At the same time, as China's national power continues to grow and its international status continues to improve, more and more international students choose to come to China to learn Chinese under the background of “cultural confidence” [2], [3]. These groups of Chinese learners are of different ages, different genders, different levels of Chinese language proficiency, and have different purposes and motivations for learning [4]. How to teach Chinese learners according to their needs has become a topic that the international Chinese language teaching community has been exploring and trying to solve [5].

With the rise and development of the new Chinese language learning mode of “Chinese + Vocational”, and in the face of the diversification of the learning needs of language learners around the world, it is of practical significance to explore the teaching of specialized languages and the mechanism of their emergence and development [6]-[8]. For other language education, the relevant courses should also explore new curriculum models according to the characteristics of their own language-technology integration and the background of the current era with the high-quality development of “Chinese + vocational skills” [9]-[11]. In this context, the proposed reform of English for Specialized Purposes (ESP) education in terms of language education concepts and vocational skills cultivation paths can provide certain teaching ideas and teaching practice references for the majority of foreign language teachers in China as well as Chinese language teachers in the international community [12]-[15].

In order to solve the problem of inefficiency of students' vocational skills development caused by the complexity of the intrinsic association of Chinese + vocational skills courses. In this paper, Apriori algorithm and social network analysis are introduced to reveal the potential association rules of Chinese + vocational skills courses and optimize the vocational skills cultivation path. The Chinese+vocational skills course data of the overall first-year students of a university in Guangdong were collected, and the Apriori algorithm was used to deeply mine the students' professional course information under different minimum confidence thresholds and minimum support thresholds to obtain the correlation between vocational skills. Then social network analysis is used to visualize the correlations in the vocational skills development pathway network, measure the reliability of the network structure using multiple evaluation indexes, and analyze the variability of different student nodes.

II. Association analysis of Chinese+vocational skills based on association rules

II. A. ESP Demand Analysis Grounded Theory

II. A. 1) Association rules

Association rule mining [16] is an important analytical method in data mining techniques. Data mining is the process of xtracting potentially valuable information and knowledge from a large amount of incomplete data. Commonly used mining methods mainly include association rule mining, correlation analysis, temporal analysis, classification, clustering and so on.

The association rule was originally motivated by merchants' desire to understand consumers' shopping habits, based on “basket analysis”. An association is a connection that occurs when something happens to something else. The association rule is described as follows: let $I=\{i_1,i_2,\cdots,i_n\}$ be the set of terms, D be the set of transactions, and each transaction T be a set of certain terms, i.e., $T\subseteq I$, and the association rule be the implication in the shape of $A\Rightarrow B$, where $A\subset I,B\subset I$ and $A\cap B=\varnothing$. Indicates some kind of correlation between transaction A and transaction B , determined by support and confidence.

Definition 1 Support: rule $A\Rightarrow B$ holds in transaction D with support support. where support is the percentage of transactions in D that contain AUB, i.e., the probability of $A\cup B$ occurring $P(A\cup B)$ as shown in equation (1).

$$support(A\Rightarrow B)=P(A\cup B) \tag{1}$$

Definition 2 Confidence: rule $A\Rightarrow B$ has confidence CONFIDENCE in the set of transactions D , where CONFIDENCE is the percentage of transactions in D that contain A that also contain B , i.e., the conditional probability $P(B\mid A)$. as shown in Equation (2).

$$confidence(A\Rightarrow B)=P(B\mid A) \tag{2}$$

There are two phases of association rule mining, the first phase is to collect the data and find out all the K-item frequent sets from the database, a frequent set is a set of items whose support meets a predefined minimum support threshold; the second phase is to generate association rules from the frequent sets, using the former K-item frequent set group to generate association rules, if the rule meets the minimum confidence level then the rule is said to be an association rule.

II. A. 2) Classical Apriori Algorithm

Apriori algorithm [17] is the earliest classical association rule algorithm based on frequent itemsets, including finding frequent itemsets and finding strong rules in two parts, the core is to find frequent itemsets, which contains two-step operations of connecting and pruning.

The basic idea of Apriori algorithm is to scan the database multiple times D to find out all the frequent itemsets, starting from 1-item frequent set, recursively generating 2-item frequent set, 3-item frequent set, and so on until all the frequent itemsets are generated.The important property of Apriori algorithm is that all the non-empty subsets of the frequent itemsets are also frequent. The algorithm implementation flow is shown in Fig. 1.

The main realization steps are as follows:

(1) Firstly, find out all frequent 1-item sets, denoted as L1;

(2) Connect and prune: find out the potential frequent 2-item sets denoted as C2, combine the algorithmic properties to determine the item sets in C2, and then mine the set of frequent 2-item sets L2;

(3) Repeat the previous step until no more frequent k-term sets can be found.

Finally, the frequent itemsets are used to construct rules that satisfy the minimum confidence level. The classical Apriori algorithm is easy to understand and implement.

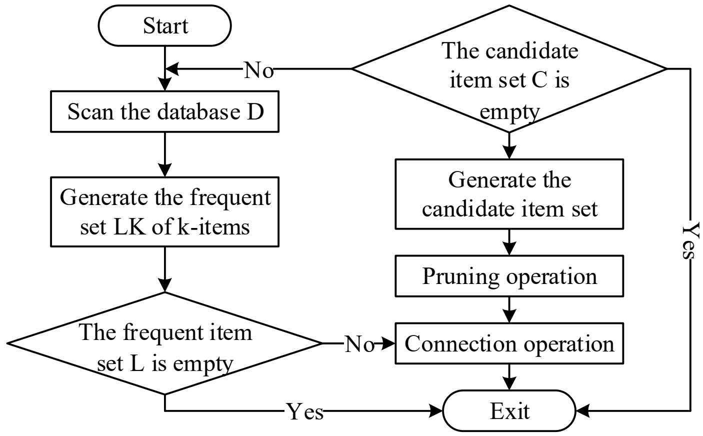

Figure 1: Flowchart of the Apriori algorithm

II. B. *Chinese+Vocational Skills Linkage Analysis*

II. B. 1) Chinese+Occupational Skills Association Analysis Model

By correlation, we mean the intrinsic connectedness between two or more things in a collection of things. If there are more and strong correlations in a collection of things, then it is possible to predict some of the things by some of them. The laws that exist between transactions are known as association rules. Association analysis, or association mining, is the search for frequent patterns, associations, correlations, or causal structures that exist between sets of items or sets of objects. The purpose of association analysis is to find the hidden correlations between data items in a given set of data records and to characterize the closeness between data items.

The data mined by association rules is called a dataset, denoted as $D=\{T_l,T_2,...,T_k,...,T_n\}$, $T_k(k=1,2,...,n)$ is called a record; each record T_k is an item list, $T_k=\{i_l,i_2,...,i_m\}$.

In this paper, the dataset D is specified as 40681 occupational records from the experimental data, T_k is one of the occupational records, and the item list $\{i_l,i_2,...,i_m\}$ in T_k is the set of occupational skills normalized to that occupational record. In this way, an occupational skills association analysis model is built.

In association rule mining, the metrics of association rules used are confidence, support, expectation and lift:

Confidence, a measure of the accuracy of an association rule, measures the strength of the association rule. The confidence of rule $X \rightarrow Y$ in dataset D is the frequency of occurrence of Y given all occurrences of X, denoted as: $confidence(X \rightarrow Y)=P(Y\mid X)=|\{T:X\cup Y\subseteq T,T\in D\}||\{T:X\subseteq T,T\in D\}|\times 100\%$.

Support, a measure of the importance of an association rule, reflects the prevalence of the association and indicates how representative a rule is. The support of rule $X \rightarrow Y$ in dataset D is the frequency of occurrence of both X and Y in all records, denoted as: $support(X \rightarrow Y)=P(X\cup Y)=|\{T:X\cup Y\subseteq T,T\in D\}|/|D|\times 100\%$, where $|D|$ is the number of records in dataset D.

Expectation, which refers to how well the successor is supported in the current dataset. In rule $X \rightarrow Y$, describes how often Y occurs in all transactions for association rule $X \rightarrow Y$ when there is no conditional influence, denoted as: $expectation(X \rightarrow Y)=P(Y)=|\{T:Y\subseteq T,T\in D\}|/|D|\times 100\%$

Boost, which describes how much the occurrence of Y is affected when X occurs, is calculated as confidence divided by expected confidence, denoted as:

$lift(X \rightarrow Y)=P(Y\mid X)/P(Y)=(|\{T:X\cup Y\subseteq T,T\in D\}|*|D|)\ /(|\{T:X\subseteq T,T\in D\}|*|\{T:Y\subseteq T,T\in D\}|)\times 100\%$

For the vocational skills association analysis model, association analysis methods can be used to mine the association rules between vocational skills items and quantitatively analyze the confidence, support, expectation and improvement of the relevant association rules.

II. B. 2) Characteristics of Chinese + vocational skills association

In association rule mining, the minimum confidence threshold and minimum support threshold are usually set. If the confidence of an association rule is greater than the minimum confidence threshold and the support is greater than the minimum support threshold, then this association rule is called a strong association rule, which is the association rule needed in this paper. In this paper, the number of strong association rules that can be mined by using a typical association rule mining algorithm with different minimum confidence thresholds and minimum support thresholds is selected from the Chinese+vocational skills course data provided by the overall first-year students of a university in Guangdong, some of which are shown in Table 1.

Under different minimum confidence thresholds and minimum support thresholds, a certain number of vocational skill association rules can be generated. Therefore, under different application scenarios, a corresponding set of strong association rules can be obtained by setting different minimum confidence thresholds and minimum support thresholds according to the specific requirements for the strength of the association rules. For example, in the case where the minimum support threshold is set to 1% and the minimum confidence threshold is set to 30%, there are 3,200 vocational skill association rules for a single antecedent and a single consequent that can be mined.

Table 1: The number of rules produced under different minimum threshers

Minimum confidence threshold	Minimum support threshold									
	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
10%	2203904	23992	1666	388	131	50	19	17	11	6
20%	2144257	22230	1424	340	132	49	18	11	3	5
30%	2005056	20854	1189	267	114	39	17	24	3	5
40%	1744479	18986	1014	211	93	32	16	18	3	5
50%	1399117	16498	847	165	60	33	15	5	2	4
60%	1023202	13590	647	129	45	30	14	4	2	4
70%	686038	9867	439	77	19	13	8	4	1	1
80%	406617	5826	267	49	9	7	5	4	1	0
90%	155245	2522	99	14	5	3	4	3	0	0

Sorting by confidence level is shown in Table 2. There are indeed a large number of strong correlations between vocational skills, for example, the correlation "public liability→litigation" has a high confidence level of 95.08%, a high support level of 1.09%, and a high enhancement level of 1474.23%. This indicates that there is a high degree of intrinsic connection between "public liability" and "litigation" in students' vocational skills behavior. Through the correlation analysis of the experimental data, it can be found that there are a large number of correlations between the students' vocational skills, and some of the correlation rules have very high confidence, enhancement and support, which suggests that there are some inherent links between the students' vocational skills.

Table 2: Examples of association rules under corresponding thresholds

N	Front piece	Afterpiece	Confidence/%	Degree of ascension/%	Support/%	Expectancy/%
1	Executive development	Executive coaching	96.45	3271.56	1.05	2.85
2	Casualty insurance	Insurance	96.31	805.25	2.66	11.56
3	Wrongful death	Litigation	95.56	1475.56	1.23	6.54
4	Inland marine	Insurance	95.45	806.51	1.05	11.56
5	Torts	Litigation	95.31	1485.56	2.13	6.54
6	Personal injury litigation	Litigation	95.24	1478.51	1.56	6.54
7	Executive development	Leadership development	95.11	1644.23	1.03	5.45
8	Public liability	Litigation	95.08	1474.23	1.09	6.54

II. C.ESP Needs Analysis and Vocational Skills Development Pathway

II. C. 1) Analysis of students' ESP needs

Demand analysis is the most critical part of student teaching, and effective demand analysis is the core of guaranteeing the effectiveness of student training. At present, the theoretical construction and empirical research on the demand analysis of various student groups from different perspectives are constantly deepening, but they also face the problems of a single source of analytical data, in-depth exploration of demand, and lack of analysis of differences in the demand of the group, which is due to the fact that the information obtained based on objective questionnaire statistics is susceptible to bias and superficial, and it is difficult to deeply explore the students' hidden needs. According to the theory of verbal behavior, the content of verbal information can reflect individual will and psychological behavior, therefore, the content of the text is an important basis for in-depth needs analysis, and the combination of scientific and technological methods can effectively promote the in-depth development of needs analysis. In the field of international Chinese language education, student training is a long-term strategic task, but the effect of student training is still unsatisfactory, and the learning needs of “Chinese + vocational skills” students need to be explored. The learning needs of “Chinese + Vocational Skills” students are written based on their own teaching situations and pain points and difficulties, which are the verbal information naturally outputted by students in the form of words before the training, and carry rich information about students' learning needs, so that the conclusions of the needs analysis based on them can be more accurate and objective.

II. C. 2) TPACK framework for students

The construction of the discipline of international Chinese language education has entered a new era of comprehensive transition to modernized teaching and research, and this new situation has put forward higher requirements for the information technology literacy of international Chinese language students.TPACK is a framework used to describe students' professional skill literacy, which is shown in Fig. 2, and it includes technological knowledge (TK), pedagogical knowledge (PK), disciplinary content knowledge (CK), pedagogical knowledge of integrating technology (TPK), subject content knowledge for integrating technology (TCK), subject pedagogical knowledge (PCK), and subject instructional knowledge for integrating technology (TPACK). Among them, technical knowledge, pedagogical knowledge, and subject content knowledge are the bottom core elements, pedagogical knowledge of integrating technology, subject content knowledge of integrating technology, and subject pedagogical knowledge of integrating technology are the middle composite elements, and subject pedagogical knowledge of integrating technology is the top composite element.The TPACK framework is highly applicable, highly recognized, and has full coverage, and is able to systematically characterize what students' vocational skills literacy should look like in the digital era.

As a new theory, research on the connotation of the TPACK framework, the expansion of its content, and the effectiveness of its application has been continuously improved, and the TPACK framework has been used in student education research in many disciplines, scientifically guiding the development of student training. In the field of international Chinese language education, the TPACK framework has profound theoretical value and practical significance, but the concept of international Chinese language student training based on the framework has not yet attracted enough attention, and its potential for application should be further explored. Based on the TPACK framework, this study aims to explore the learning needs and characteristics of different “Chinese + vocational skills” students in the field of international Chinese language education.

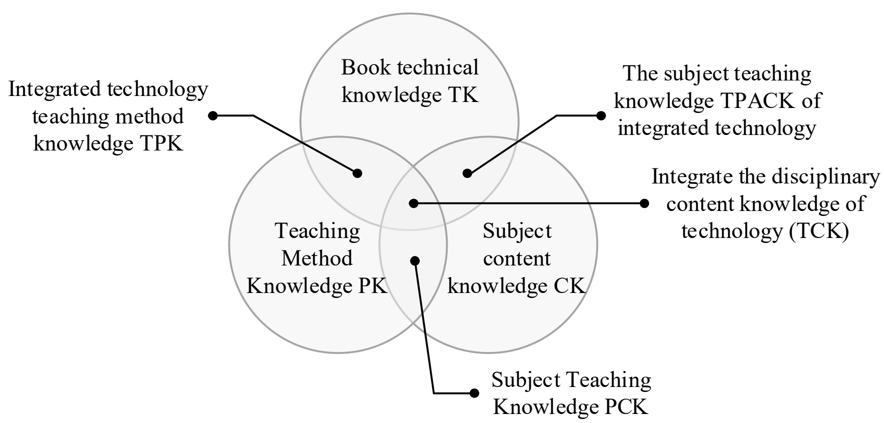

Figure 2: TPACK framework model

III. Cultivation path analysis based on social network analysis

III. A. Social network analysis methods

III. A. 1) Relevant overview of social network analysis methods

Social network analysis method [18] is a sociological research paradigm, as the extension and development of social network theory, which emphasizes that the society is not composed of individuals but by the network, the network contains nodes and the relationship between nodes, the social network analysis method can be through the network of relational data, visual expression and clear data analysis, and then to explore the structure of the network and the characteristics of the attributes. In this paper, we analyze the village public space network and villagers' behavioral network through the social network analysis method, and combine the quantitative results of various types of data to provide scientific basis for the optimization strategy of rural public space network layout and enhancement.

III. A. 2) Data representation of social network analysis methods

Community diagrams and data matrices are two forms of representation that together constitute a data representation of social network relationships. One of them is the community diagram, which is a chart reflecting the overall situation of a network, first used by Moreno and widely used in social networks, this charting method can clearly show the connections and interactions between various parts of a network. Another notation is data matrix, which is a matrix model to study two relationships or a combination of relationships, which helps researchers to systematically analyze social network data.

(1) Community Graph

Community graphs, with their unique network expressions, are used to describe the interrelationships between social actors. Each dot in the community graph symbolizes an actor, while the connection represents the connection and interaction between actors. At the same time, according to different classification criteria, different types of community graphs can be obtained, if the direction of the relationship is taken as the scale, the community graph can be divided into "directed graph" and "undirected graph", the former highlights the directivity of the relationship between actors, and the latter pays more attention to the two-way and interactive nature of the relationship. Furthermore, according to the closeness of the relationship, the community graph can be divided into "binary value graph" and "multi-value graph", the former simplifies the relationship to yes or no, and the latter assigns different levels and depth to the relationship through the depth of the value.

(2) Data matrix

A matrix is a combination of related elements arranged in the form of a rectangle, the size of which is determined by the number of rows and columns. The rows and columns in the matrix represent the "actors", i.e., the nodes in the social network diagram, and the elements at the intersection of the rows and columns represent the "relationships" between these actors. In the data matrix, the position of each element is the identification of its identity, and if the rows and columns of the matrix represent the "social actors" in the set of actors, then the elements of the matrix reflect the "relationship" between these actors, and such a network is a 1-mode network. Conversely, if the rows and columns of a matrix represent two different sets of "social actors", then the elements of the matrix reflect the "relationship" between the individuals in the two groups of actors, and this network is classified as a two-mode network.

III. A. 3) Parametric indicators of social network analysis methods

As an important tool for studying social structure and relationships, social network analysis commonly uses network structure indicators such as network density and cut points, which can reveal the basic characteristics of the network structure after in-depth analysis. The following are the specific meanings of these network structure indicators:

(1) Network density

Network density is the ratio of the number of nodes actually existing in the overall network to the number of potentially connectable nodes, which reflects the strength of the network's internal correlation. When the cluster contains n public space, its maximum possible connection value is $n(n-1)/2$. If the actual number of connections is closer to this maximum value, it indicates that the network is more connected. The specific calculation method is as follows:

$$P = L \times \frac{n(n-1)}{2} \tag{3}$$

where, P is the network density, L is the number of connections actually present in the network and n is the number of spatial nodes in the network.

(2) Centripetal potential

This indicator measures the degree of construction of the network around the core node, reflecting the network's wholeness and centripetal force. The increase of network central potential means the enhancement of network stability, and its value ranges from 0 to 1. The specific calculation is as follows:

$$C = \sum_{i=1}^{n}(C_{\max} - C_i) / \max(\sum_{i=1}^{n} C_{\max} - C_i) \tag{4}$$

where, $C_{\max}$ is the maximum node degree of each node in the network, C_i is the node degree of node i, and n is the number of nodes in the network.

(3) Factions

A faction, also known as a clique or cohesive subgroup, is a subset of a social network whose members have reciprocal relationships and are closely connected to each other, and the connections between members within a faction are significantly closer than those with other groups. From the perspective of network graphs, factions are composed of at least three nodes, forming a fully connected subgraph, i.e., there is a direct line between any two nodes, which makes factions play an important role in social networks, and the close interactions and reciprocal relationships among their internal members have a far-reaching impact on the structure and function of the network.

(4) Point centrality

Point centrality index is an important criterion to measure actors' control over resources, which can reveal how many pairs of connection paths between nodes a node plays a key role, if a node has n connection path, then the node's point centrality value is n. The increase of its value means the number of connections is increasing, which also symbolizes the core position of spatial nodes in the network is gradually highlighting. Analyzing the point centrality of network nodes is not only a measure of network structure, but also a key to resource allocation and power distribution.

(5) Cutpoints

Cutpoints are in the key position of the internal structure of the network, with its unique position connecting the major clusters in the network, like a bridge to connect all parts of the tightly, once the cutpoints are damaged, the whole network will be turned into several isolated network segments. Because of this, cutpoints are often regarded as an important indicator for evaluating the robustness of a network, and the higher the number of cutpoints in a network node, the higher the risk of vulnerability the network faces.

III. B. Network analysis of vocational skills development pathways

III. B. 1) Individual student centrality analysis

Centrality describes what position an individual is in the network, such as whether it is at the center or the edge of the network, and it can reflect the importance of the individual in the whole network (this study is the vocational skills training path network).

Accordingly, this study selected the whole of a freshman class in a university in Guangdong, which has a total of 30 students, as the research object, and used social network analysis to study the students' vocational skills cultivation pathway network, explore the structure of students' vocational skills cultivation pathway network, and understand the vocational skills competence of individual students.

In this study, the obtained student data were imported into pajek, and then their point centrality, mediated centrality, and proximity centrality were calculated, and finally plotted and exported.

The degree of entry of individuals in the vocational skills development pathway network of students in this class is shown in Fig. 3, the color of the circle is red, orange, yellow, green, cyan, blue, purple, which represents the entry value 0, 1, 2, 3, 4, 5, 6, respectively, and the number is the number of the student. student No. 25 has the largest degree of entry, and the value of the degree of entry is 6. The students with a lower degree of entry are 1, 10, 16, 20, 21, 28, respectively, and the degree of the above students' degree of entry is 0. The data will be imported into SPSS130 to do Spearman rank correlation with teacher's subjective evaluation of students. data into SPSS 13.0 and the teacher's subjective evaluation of the students to do Spearman rank correlation, get the correlation coefficient r = -0.425, P < 0.05 (the reason why the correlation coefficient is negative is because the teacher's subjective evaluation and the student's degree of entry of the scores in the opposite direction). The results indicate that the higher the entry degree, the lower the subjective rating, i.e., the higher the students' vocational skills competence, and the results indirectly indicate the reliability of the online data.

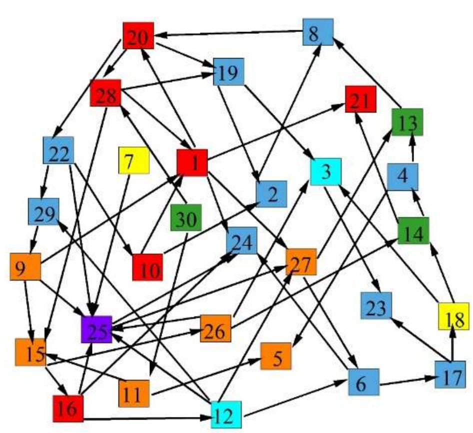

Figure 3: Enrollment of individual students

The frequency histogram of the students' vocational skills entry degree is shown in Figure 4, which shows that: the entry degree of the students in the class does not obey the normal distribution, and the values of the entry degree are concentrated in the range of 0 to 6, among which the most students have the entry degree value of 5.

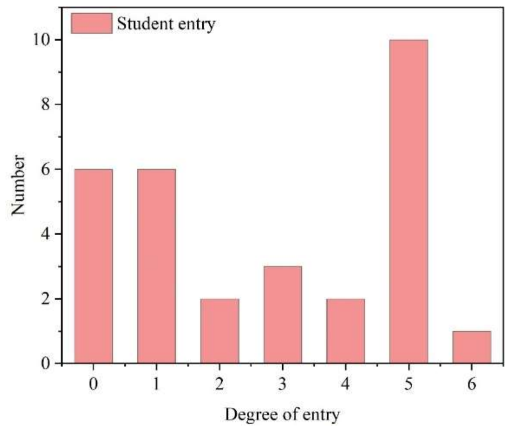

Figure 4: Student entry frequency histogram

Individual mediational centrality is shown in Figure 5.The mediational centrality of student No.25 is the largest with a value of 0.31.The mediational neutrality of students No.8 and No.19 is also relatively high with a value of 0.25.The data were imported into SPSS13.0 to do a Spearman rank correlation with the subjective ratings of the teachers and a correlation coefficient of r=-0.406 was obtained, P<0.05.The results indicate that the higher the mediational centrality , the lower the subjective rating, i.e., the higher the vocational skill competence of the students.

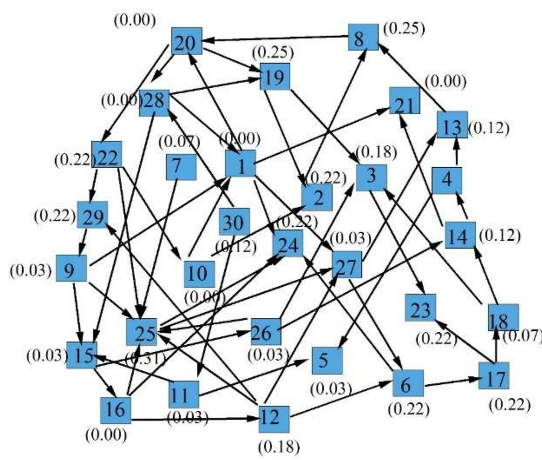

Figure 5: The intermediary center of the individual

Individual centrality of proximity is shown in figure 6. students number 8 and 21 have the largest centrality of proximity with a value of 0.44, student number 27 has a value of 0.43. student number 20 has a value of 0.41. some students even have a centrality of proximity of 0.00, such as students number 10 and 26. These data were imported into SPSS 13.0 and correlated with teachers' subjective ratings on a Spearman scale, and the correlation coefficient was obtained as r=-0.592, P<0.05. The results indicate that the higher the centrality of proximity, the lower the subjective ratings, i.e., the higher the vocational skills competence of the students.

The centrality of representative individuals is shown in Table 3, from the three centrality indexes, it can be seen that although all of them are significantly negatively correlated with subjective vocational skills competence, the same nodes are different in different individual centrality indexes, the point degree centrality and mediator centrality of student No. 5 and student No. 9 are consistent, but their proximity centrality are different, respectively, 0.03 and 0.23. Then, as in the case of student No. 7 and Student #18, their dot centrality and proximity centrality are the same, but their mediated centrality is 0.07 and 0.00, respectively.

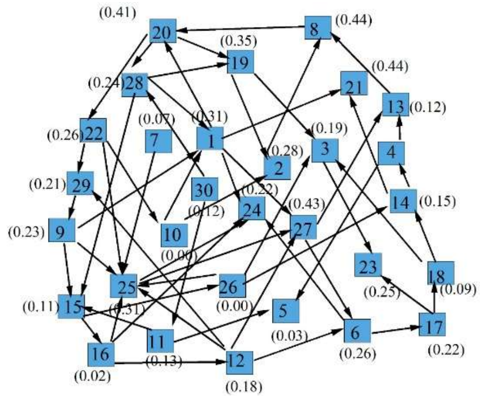

Figure 6: The proximity of the individual

Table 3: The center of the representative individual

Node number	Point center degree	Intermediate center	Proximity center
5	1	0.03	0.03
9	1	0.03	0.23
7	2	0.07	0.07
18	2	0.07	0.00

III. B. 2) Analysis of cohesive subgroups of vocational skill development pathways

Cohesive subgroups are in the middle of the pack in terms of the frequency of use of the studies covered by social network analysis, with roughly 45% of the studies using the method. Social network analysis uses relational data to delineate cohesive subgroups, or cliques, based on graph theory, with the main assumption being that inter-individual relationships within cliques are stronger than inter-individual relationships between cliques. The main method of delineating cliques is to identify the subgraph components of the relational structure graph that are relatively closely related, and to determine the members of the cliques accordingly. The level of cohesion of the whole network can be measured by the average point degree of all vertices, but since each student basically nominates multiple other students, this metric is not very meaningful. In addition, k kernel refers to a sub-network in which all vertices in the group have an incidence value of not less than k, and the same members of the network with higher k kernel values are prone to form more closely related groups. In this study, the obtained data of students' friendship nominations were imported into Pajek, the k values were calculated, plotted and exported.

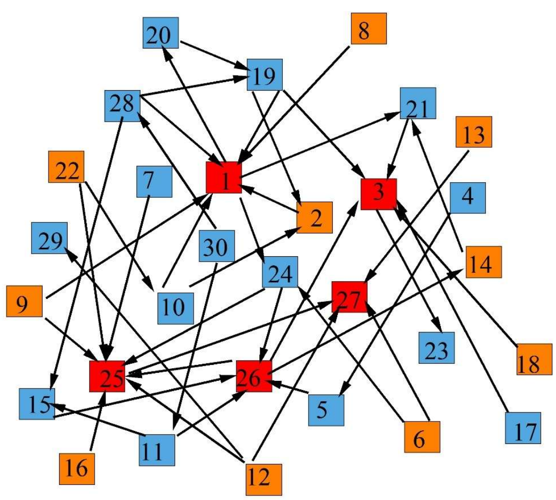

Figure 7: K network map

Firstly, the average point degree of all the vertices in the network was calculated to be 5, and the density was 0.065. Indicating that the cohesion of the network was good. Pajek was used to draw the k kernel network graph of friendship and companionship, k kernel network graph is shown in Figure 7, the different colors of the rectangle represent the students with different k kernels, the red color represents the students with the k value of 2, while the orange and the blue color represent the students with the k value of 1 and 0 respectively. The graph shows that there are several students in the class who have a k-core value of 2, such as #1, #2, #25, #26, and #27. The rest of the students have k kernel values of 1 and 0. It can be seen that in the cooperative interactions of the class a larger group was formed with students #1, #2, #25, #26, and #27 as members.

IV. Conclusion

Through the correlation analysis of students' data in the Chinese+vocational skills course, it is found that the correlation relationship “public liability→litigation” has a high confidence and support level, with a high degree of enhancement of 1474.23%. This indicates that there are a large number of inherent correlations between students' vocational skills, which lays the foundation for the construction of the correlation network of students' vocational skills cultivation path.

The analysis of the individual centrality of the students' vocational skills cultivation path network shows that the students' individual centrality indexes are consistent with the teachers' subjective vocational skills competence evaluation of the students, indicating the reliability of these indexes, and the vocational skills cultivation path network constructed in this paper can be utilized to guide the students' learning and competence enhancement.

Students who are in the core position in the class have higher point degree of centrality and stronger leadership, which can drive other students to improve their vocational skills level together, so that more learners can join in the vocational skills improvement activities. In addition, in the development of students' vocational skills can make full use of the class intermediary center of the higher degree of students, this kind of students have a wider range of human relations, can contribute to the smooth implementation of students' vocational skills improvement activities. Students with higher proximity centrality are the closest to other students in the class, so they can be trained to take the role of leaders and set an example for other students to have a higher awareness of vocational skills enhancement. Using the incomplete consistency of the three indicators, analyzing the differences in the indicators of the students from different perspectives, and taking advantage of the differences that exist among these students in the vocational skills cultivation pathway network, targeting different vocational skills cultivation strategies will be conducive to the further optimization of the vocational skills level of the students.

References

[1] Nasihin, A. (2022). The Impact of Using English Curriculum Design Based on Industry Needs in English Teaching on Vocational School to Improve Students English Skill for Industry Standard Working Communication. Education Quarterly Reviews, 5(1), 138-144.

[2] Liu, W., & Lin, X. (2016). Meeting the needs of Chinese international students: is there anything we can learn from their home system?. Journal of Studies in International Education, 20(4), 357-370.

[3] Gong, Y., Lai, C., & Gao, X. (2020). The teaching and learning of Chinese as a second or foreign language: The current situation and future directions. Frontiers of Education in China, 15(1), 1-13.

[4] Chen, C., & Du, X. (2022). Teaching and learning Chinese as a foreign language through intercultural online collaborative projects. The Asia-Pacific Education Researcher, 31(2), 123-135.

[5] Mohamed, O., & Wei, Z. (2017, December). Motivation and satisfaction of international student studying Chinese language with technology of education. In 2017 International Conference of Educational Innovation through Technology (EITT) (pp. 272-277). IEEE.

[6] Jiang, B. (2022). Research on the application of Chinese traditional culture teaching in higher vocational education. Kuram ve Uygulamada Egitim Bilimleri, 22(2), 1-13.

[7] Wang, D., Moloney, R., & Li, Z. (2013). Towards internationalising the curriculum: A case study of Chinese language teacher education programs in China and Australia. Australian Journal of Teacher Education (Online), 38(9), 116-135.

[8] Dai, D. W. (2023). What do second language speakers really need for real-world interaction? A needs analysis of L2 Chinese interactional competence. Language Teaching Research, 13621688221144836.

[9] Fadlia, F., Zulida, E., Asra, S., Rahmiati, R., & Bania, A. S. (2020). English subject on English for Specific Purposes (Esp) in vocational schools. Language Literacy: Journal of Linguistics, Literature, and Language Teaching, 4(2), 358-367.

[10] Widodo, H. P. (2016). Teaching English for specific purposes (ESP): English for vocational purposes (EVP). English language teaching today: Linking theory and practice, 277-291.

[11] Ali, A., Firissa, B., & Legesse, A. (2017). Investigating the relevance of ESP (English for Specific Purpose) for TVET (Technical and Vocational Education and Training) Colleges in Ethiopia. Int. J. of Multidisciplinary and Current research, 5.

[12] Hyland, K. (2022). English for specific purposes: What is it and where is it taking us?. ESP Today-Journal of English for Specific Purposes at Tertiary Level, 10(2), 202-220.

[13] Maulani, D. S., Friatin, L. Y., & Rachmawati, E. (2022). Voicing the Students' Need of English Vocabulary English for Specific Purposes in Vocational High School. JELITA: Journal of Education, Language Innovation, and Applied Linguistics, 1(1), 56-66.

[14] Fitria, T. N. (2020). Teaching English for specific purposes (ESP) to the Students in English language teaching (ELT). Journal of English Teaching Adi Buana, 5(01).

[15] Ayuningtyas, P., Mauludin, L. A., & Prasetyo, G. (2022). Investigating the anxiety factors among English for Specific Purposes students in a vocational education setting. Language Related Research, 13(3), 31-54.

[16] Chuang Yang & Huanxin Chen. (2025). Analysis of power consumption influencing factors in different modes of semiconductor factory refrigeration station based on association rule mining. International Journal of Refrigeration,175,187-195.

[17] Mingyang Liu,Jiake Li,Yafang Li,Weijie Gao & Jingkun Lu. (2025). Data-driven identification of pollution sources and water quality prediction using Apriori and LSTM models: A case study in the Hanjiang River basin.. Journal of contaminant hydrology,272,104570.

[18] Xi Kang,Saiyong Li & Kanchaya Chaivirutnukul. (2025). A study on the impact of entrepreneurial bricolage on enterprise performance management using the BPNN-DEMATEL method and social network analysis. Technology in Society,82,102883-102883.