J. COMBIN. MATH. COMBIN. COMPUT. 127b (2025) 4153--4166

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

Combinatorial Press

Identification of College Students' Online Behavior Patterns and Optimization of Online Civic Education Strategies

Xiuling Li[1,✉]

[1] *School Of Mechanical Engineering, North University of China, Taiyuan, Shanxi, 030051, China*

ABSTRACT

Focusing on the learning behavior patterns of students with network behavior, this study mainly adopts sequence cluster analysis and lag sequence analysis to convert learning behaviors into sequences, and constructs a learning behavior pattern recognition model based on network behavior sequences. Aiming at different types of classroom learning behaviors in civic education under the network behavior sequence, a targeted teaching intervention mechanism is designed to help students convert their learning behavior patterns and thus improve their learning effects. In this paper, the online behaviors are clustered into four categories of "integrated, autonomous, compliant, and deviant" according to six level 1 codes, and the correlation coefficients of the online behaviors in the four learning categories range from 0.8539 to 0.9944, which is a very strong correlation. Finally, a survey of the results of the intervention in the classroom of Civic Education found that 75.22% of the students believed that the intervention had improved the learning effect of Civic Education. 67.7% and 77.54% of the students believed that the intervention had improved the enthusiasm and motivation of Civic Education learning. 79.04% of the students were willing to continue to learn independently according to the learning behavior pattern after the intervention.

Keywords: Internet behavior, Sequential cluster analysis, Lagged sequence analysis, Pattern recognition model, Civic and political education

1. Introduction

The popularity of the Internet has gradually changed the way people think and act. Since China's access to the Internet in the late 1990s, the Web has profoundly transformed Chinese society. College students are mostly "Post-90s" and "Post-00s", who were born after China's access to the Internet, and have become "network aborigines" and the main force of Internet users [1-2].

In the Internet era, college students with distinctive personalities build various circles with the network as the main position and interest and emotion as the core bond, they are an important force for social development in the new era, they like to contact with new things, but they don't have

✉ Corresponding author.
E-mail address: leedebbie817@163.com (X. Li).

Received 10 January 2024; Revised 25 February 2024; Accepted 15 November 2024; Published Online 16 April 2025.
DOI: 10.61091/jcmcc127b-229

complete cognitive ability to all kinds of events happening in the society, and they are easily influenced by the network information [3-4]. College students have been active in cyberspace for a long time, and the effective use of ideological education to guide them to correctly use the convenience brought by the network and regulate network behavior will be conducive to the realization of ideological and political education in the Internet space of the living, and promote the formation of correct values among college students [5].

Contemporary college students grow up with the development of the Internet, is one of the groups most deeply affected by the information age, has the cognition of contacting new things, enjoys the diversified information dividends brought by the Internet, but at the same time also affects the judgment of values because of the diversity of information, how to maximize the use of the university's Civic and Political Education to regulate the behavior of college students in the network, and give full play to the advantages of the Civic and Political Education is a concern of the universities and colleges focus [6].

The popularity of the Internet has made it an indispensable part of people's life, education and work, and students, due to their age and experience, are prone to be disturbed by the negative information on the Internet, which seriously affects their normal learning and life, and even affects the development of their correct outlook on the three things, therefore, the researchers have investigated the students' Internet behavior and the influence of the Internet through the use of questionnaires, meta-analysis, and other methods, and literature [7] reveals the negative impact of online social media on Pakistani students based on a sample questionnaire and discusses in detail the relationship between oppositional and creative characteristics of online social media. Literature [8] used meta-analysis method to analyze the studies on the relationship between students' internet surfing time and mental health and synthesized the findings of several studies to show that students' internet social time is weakly associated with positive psychological indicators, negative indicators and mental health. Literature [9] systematically reviewed the research on the relationship between the degree of Internet use and academic performance, and pointed out that the time spent on Internet socialization showed a negative correlation with students' academic performance, but a positive correlation with the language test, which was more significant in the group of women and college students. Literature [10] examined the behavior of Internet addiction and Internet abuse attributed to the loss of control of Internet use through the use of tools with full validation assurance and principal component analysis of causality, and the study can provide some reference for the management of students' Internet behavior. Literature [11] decodes and analyzes the foundations of Internet ideology and describes how Internet ideology has been who home and promoted, and provides detailed illustrations through three real-life cases. Literature [12] designed a new media platform for college students' ideology education based on the background of the Internet to enrich the content of college students' ideology teaching and cultivate students' Internet ideology awareness, and the system showed good performance in the data test, which was praised by teachers and students.

This paper introduces the contents of the definition of sequence correlation of online behavior from the research object of online behavior pattern recognition and optimization of online Civic and Political Education. The correlations in students' online behaviors in the classroom of online Civics education are mined by sequence pattern and sequence clustering techniques. Then the SGT algorithm is used to extract the length of sequence features, and the correlation between the two extracted variables is explored by using the Pearson correlation coefficient. Finally, the probability of occurrence of sequence behaviors over time is assessed under the lag sequence analysis technique, and a learning behavior recognition model based on behavior sequences is constructed to investigate the learning effect after students' Civic Education interventions under network recognition.

2. Network behavior pattern recognition and performance prediction model construction

2.1. Network Behavior Pattern Mining and Clustering Techniques

2.1.1. Network Behavior Sequence Definition. Network behavior [13] is meant to be a class of action records produced by learners adopting corresponding network ways and methods throughout the network process. In the MOOC network environment, due to the high degree of freedom of online network time, place and method, learners can choose to get the resource video or courseware document they need at any time, any place and in any kind of network way. Learners in the computer web page or cell phone mobile web page in any one click that is generated by the learner online network of each network behavior.

2.1.2. Techniques related to behavioral sequence mining:
1) Sequential Pattern Mining. Association rule [14] research topics are mainly categorized into two aspects: frequent itemset mining [15] and sequential pattern mining.

In the frequent itemset mining problem, researchers mainly study how to discover the set of items hidden in large datasets that co-occur frequently. Assuming that a single item (Item) in the research subject is defined as i, the large item set (Itemset) is defined as $X=\{i_1,i_2,\ldots,i_n\}$. Let $T=\{t_1,t_2,\ldots,t_n\}$ be a transaction dataset, where each transaction consists of single or multiple items in the item set, and each transaction $t_i(i=1,2,\ldots,n)$ corresponds to a subset on the total item set X. Let $X_1 \subset X$, the support (Support) of the corresponding item subset X_i on the transaction set T be the percentage of items containing X_i that appear in the transaction set T, with the following formula:

$$Support\left(X_i\right)=\frac{\left\|\{t \in T \mid X_1 \subset t\}\right\|}{\|T\|} \tag{1}$$

The requested association rule is denoted as $X \Rightarrow Y$, which indicates some deterministic relationship between the two itemsets. A rule is supported if the percentage of transactions corresponding to the association rules of the requested itemsets X to Y exceeds a set threshold, where the set threshold is called the degree of support or the minimum degree of support.
2) Sequence clustering. Sequence clustering techniques [16] provide methods for classifying multiple sequences into clusters or groups of similar sequences. Sequence clustering algorithms are a class of hybrid algorithms that combine clustering with sequence recognition techniques for identifying classes of sequences. In sequence clustering, each cluster is associated with a probabilistic model, usually a Markov chain. If the Markov chains of all clusters are known, then each input sequence is assigned to the cluster most likely to produce such a sequence. In general, there can be more than one possible cluster, so sequences are assigned to the cluster that can produce the input sequence with a higher probability.

Since each cluster has its own Markov chain, the probability that an observed sequence belongs to a given cluster is equal to the probability that the sequence is generated by the Markov chain in which the associated cluster is located. For sequence $x=\{x_0,x_1,x_2,\ldots,x_{L-1}\}$ of length L, the formula is as follows:

$$p\left(x \mid c_x\right)=p\left(x_0;c_k\right)\cdot\prod_{i=1}^{L-1} p\left(x_i \mid x_{i-1};c_k\right) \tag{2}$$

where $p\left(x \mid c_x\right)$ is the probability of x_0 occurring as the first state in the Markov chain associated with cluster c_k; and $p\left(x_i \mid x_{i-1};c_k\right)$ is the transition probability of state x_{i-1} moving to state x_i in the same Markov chain.

2.1.3. Sequence clustering based on SGT algorithm:

1) Overview of SGT algorithm. SGT algorithm for the extraction of sequence features is mainly concerned with the relative position of each element in the sequence, the use of the relative position information of each element pair, after the corresponding calculation and transformation, to obtain a vector representation compatible with different lengths of the sequence.

Now introduce the SGT algorithm involving the relevant definitions: Suppose that there is a sequence data set denoted as S, and any sequence in the sequence database consisting of a set of letters V denoted as $s \in S, s$ may contain one or more letters in set V. If the sequence dataset $S = \{AABAAABCC, DEEDE, ABBDECCABB, \ldots\}$ consists of the corresponding elements of the set of letters $\mathrm{V} = \{A, B, C, D, E\}$. The length of any sequence s is denoted as $L^{(s)}$, and its value is equal to the number of events in the sequence. The sequence, s denotes the value of the letter taken at the corresponding position l in the sequence, where, $l = 1, \ldots, L^{(s)}$, and $s_l \in \mathrm{V}$.

The function that measures the magnitude of the effect of the relative position of two letter elements is then $\phi_\kappa(d)$. $\phi_\kappa(d)$ The geometric meaning of the function is visualized with the relative position as shown in Fig. 1 (a and b represent different paths), and in the case of $\varphi(d(l,m))$, for example, l and m are the positions of the event, and $d(l,m)$ is the distance metric between the l and m positions. In order to apply the SGT algorithm, ϕ the function needs to fulfill the following three conditions:

1) The function value needs to be strictly greater than 0, i.e. $\phi_\kappa(d) > 0$; and for any $\kappa > 0$, $d > 0$;

2) The function value is strictly decreasing as d increases, i.e., $\partial\phi_k(d)/\partial d < 0$;

3) The value of the function is strictly decreasing as κ increases, i.e., $\partial\phi_x(d)/\partial\kappa < 0$.

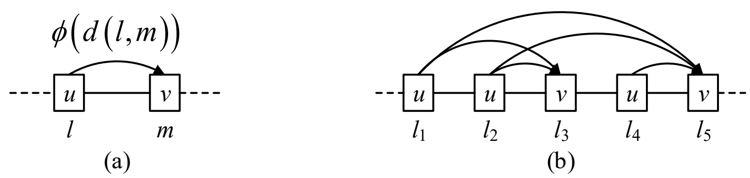


Fig. 1. $\phi_\kappa(d)$ Function geometry and relative location visualization

In the selection of ϕ function expressions, there are more functions that satisfy the above conditions to choose from, such as Gaussian function, inverse function and exponential function. Here, out of the need for interpretability of the algorithm results and function properties, the exponential function is selected as the distance metric function in the algorithm, and the expression is shown below:

$$\phi_\kappa(d(l,m)) = e^{-\kappa/m-l}, \forall \kappa > 0, d > 0 \tag{3}$$

In general, there are generally more than one pair of paths generated between two events, i.e., there is more than 1 distance function. As shown in Fig. 1 (b), there are 5 paths from event u to event v, and the 5 (u,v) pairs i.e., will correspond to the generation of 5 distance metric function values, each of which corresponds to have an impact on the link value between the two. Therefore, when the algorithm is executed, the number and position of paths between the two events to be calculated will be counted first, and then stored in a $|V| \times |\mathrm{V}|$ asymmetric matrix Λ, so at this time Λ_{uv} will contain all the relative position information of all the 5 (u,v) pairs, where the positional relationship, u is in the front, and v is always in the back. Summarized in the formula below:

$$\Lambda_{uv}(s) = \left\{(l,m) : s_l = u, s_m = v, l < m, (l,m) \in 1, \ldots, L^{(s)}\right\} \tag{4}$$

Each (u,v) pair of instance information in this matrix Λ_{uv} affects the final u,v link function ψ and in the implementation of the algorithm, the expression of the ψ function of the feedforward sequence s is obtained by performing aggregation and normalization operations on the values of the distance metric function of each pair of instances as shown below. Where $u,v \in \mathrm{V}$, the SGT feature of the sequence is denoted as $\Psi(s)=\left[\psi_{uv}(s)\right]$:

$$\psi_{uv}(s)=\begin{cases}\dfrac{\sum\limits_{\forall(l,m)\in\Lambda_{av}(s)} e^{-\kappa|m-\Lambda|}}{\left|\Lambda_{uv}(s)\right|}; \; length\ sensitive \\ \dfrac{\sum\limits_{\forall(l,m)\in\Lambda_{at}(s)} e^{-\kappa|m-l|}}{\left|\Lambda_{uv}(s)\right|/L^{(S)}}; \; length\ insensitive\end{cases} \tag{5}$$

Considering the possible influence of sequence length on the results of feature extraction, two calculation methods, sequence length-sensitive and sequence length-insensitive, are given respectively. In the sequence length-sensitive calculation, the final results take into account the influence of sequences with different lengths but similar patterns on the final feature expression, whereas in the length-insensitive calculation method, the elements of matrix Λ are normalized according to the sequence length $L^{(s)}$. The effect on the results due to sequence length issues alone is eliminated.

2.2. Data correlation analysis techniques

Data analysis refers to the process of using accurate and appropriate analytical methods and tools to analyze processed data and extract valuable information, so as to form valid conclusions and present them through visualization techniques.

Correlation analysis refers to the process of analyzing two or more correlated elements of a variable in order to measure the closeness of the correlation between the two variables. There needs to be a certain link or probability between the correlated elements for correlation analysis. Correlation is not the same as causation, nor is it simply personalization. The scope and field of correlation covers almost every aspect we see, and the definition of correlation varies greatly from discipline to discipline. In this study, in order to investigate the influence of independent learning ability on students' learning effect, it is necessary to analyze the correlation between independent learning ability and students' performance, and the Pearson correlation coefficient is mainly used to judge the correlation.

Pearson's correlation coefficient is a widely used measure of correlation between two variables. The overall correlation coefficient between two variables is defined as the quotient of the covariance over the standard deviation between the two variables, where σ_X, σ_Y represents the standard deviation of the two variables. This is shown in the following equation:

$$\rho_{X,Y}=\frac{\operatorname{cov}(X,Y)}{\sigma_X\sigma_Y}=\frac{E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right]}{\sigma_X\sigma_Y} \tag{6}$$

Based on the overall correlation coefficient, estimating the covariance over the standard deviation of the sample gives the Pearson correlation coefficient as shown in the following equation:

$$r=\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\sqrt{\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2}} \tag{7}$$

As can be seen from the formula, the absolute value of the Pearson correlation coefficient is less than or equal to 1. The closer the value is to 1, the more positive correlation is shown between the two variables, and the closer it is to -1, the more negative correlation is shown between the two variables.

2.3. Construction of Learned Behavior Recognition Model Based on Behavior Sequence

Lagged Sequence Analysis (LSA) is a method used to assess the probability of occurrence of sequential behaviors over time, which is capable of examining the significance of the resulting network behaviors, analyzing in-depth the significance of the probability of occurrence of one behavior after the generation of another, and thus uncovering the pattern of the students' network behaviors. The lagged sequence analysis process is shown in Figure 2. The lag sequence analysis method mainly utilizes the GSEQ tool for analysis. In the GSEQ analysis process, first of all, we need to strictly follow the software format requirements for behavioral coding, and then input these codes, and then the system compiled to generate the MDS format file, to check whether the code is consistent, and then conduct the analysis of behavioral sequences, to get the coefficient table of behavioral pattern identification, usually including the frequency table and the residual value table, and finally based on the size of the residual value to filter the behavioral sequences that have significant significance. Finally, we filter the behavioral sequences with significant significance according to the size of the residual values, and draw the behavioral transformation diagram.

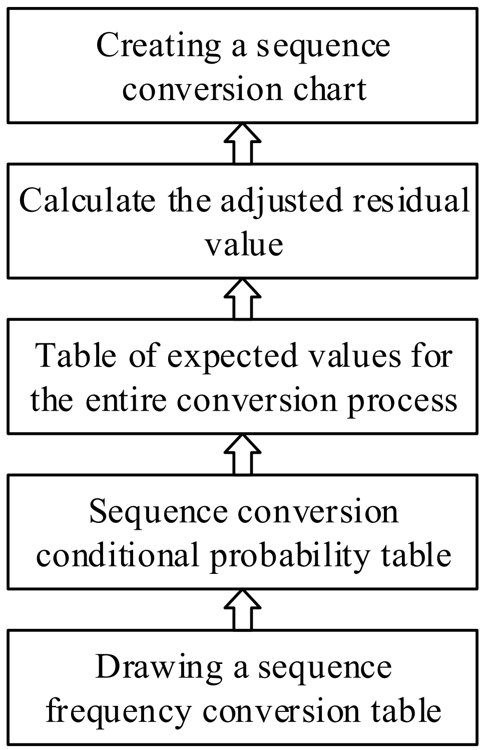


Fig. 2. Late sequence analysis process

In the process of lagged sequence analysis, the following two points need to be focused on and implemented: first: before running the lagged sequence analysis, it is necessary to encode the behavior of students interacting with the learning platform, encode the behavior of students based on the identified encoding rules, and generate the encoding sequence by recording the encoding results of each student according to the chronological order. Second: After running the lagged sequence analysis, it is necessary to pass the results of the run, namely, the standard residual value (Z-Score, Z scores), which is shown in Equation 1 below about the calculation of the standard residual value, where σ is the residual, Z scores are used to indicate the significance of the conversion between two behaviors, when the Z scores (the standard residual value) is greater than 1.96, it means that the relationship of the behavioral conversion is statistically significant , the greater the Z score, the greater the significance of the conversion relationship between the two behaviors. When the Z-score is greater than 1.96, the two behaviors show continuity:

$$z-score=\frac{\sigma-\bar{\sigma}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\sigma i-\bar{\sigma})}} \tag{8}$$

3. Results and Analysis of College Students' Internet Behavior Identification and Performance Prediction

3.1. Behavioral Sequence-Based Network Behavioral Patterns

This paper takes 58 students in a class of freshmen at University H as an example, this study selected 4 weeks of the class's Civic and Political Education class, the observation time is in units of one week, and 2 interventions on the teaching process were conducted in the 3rd and 4th weeks, respectively. The behavioral data generated by each student in the network were recorded for the recording and analysis of college students' online behavioral patterns and online Civic and Political Education strategies. In this study, the possible classroom online behaviors of the subject students were summarized into 12 types, which were A: online autonomous behaviors (including 1 looking through the textbook), B: online interactional behaviors (including 2 thinking, 3 collaborative learning, and 4 answering questions), C: online manipulative behaviors (including 5 note-taking and 6 working on exercises), D: non-internet behaviors (7 wandering off), and E: online perceptual behaviors (including 8 other, 9 listening to lectures, 10 revising exercises, and 11 looking at notes), and F: Web-based teacher-assisted behaviors (12 being asked questions). The first-level coding and explanation of online learning behaviors are shown in Table 1.

Table 1. Network learning behavior primary code and interpretation

Primary coding	Explain	Primary coding	Explain
A: Autonomous network behavior	The students read the Internet autonomously	D: Non-network behavior	Students are distracted from the classroom and do things that have nothing to do with the classroom on the Internet
B: Dating network behavior	Students communicate with their peers or teachers by network	E: Perceptual network behavior	Students listen to the teachers in the network and read their notes on the network equipment
C: Operational network behavior	Students use laptops or tablets to record notes	F: Teacher supportive network behavior	The teacher asked the students to answer the question by using the network equipment

3.1.1. Learning Behavior Pattern Clustering Based on Network Behavior Sequences. In this study, the sequence clustering method was used to cluster network behaviors. After iterative processing, the classroom online learning behaviors of 58 students in a freshman class of H University were divided into four categories based on "9 lectures→6 exercises", "9 lectures → 2 thinking", "10 modified exercises→9 lectures", "3 cooperative networks→10 modified exercises", "2 thinking→9 listening and speaking", and "average grades", as shown as Table 2. Analysis of the behavioral sequences of the four cluster groups found that:

1) Behavioral pattern recognition in cluster group 1 is the most in terms of variety and frequency, and the students in this group also have the best effect of learning in the network, which belongs to the classroom integrated learner, i.e., the students use a variety of classroom network behaviors comprehensively in the classroom, rely less on the teacher, and have a strong subjective initiative.

2) The types of behavioral pattern identification in cluster group 2 are slightly different from that of cluster group 1, and it can be inferred that this group of students is dominated by a variety of self-

study behaviors, and the students' learning behaviors in the network are more effective, and belong to the classroom-autonomous learners.

3) The number of times of behavioral pattern recognition in cluster group 3 is significantly less than that of the previous two groups, and the behavioral pattern recognition is relatively single. It is worth noting that the number of times students in this group "think→listen" is as high as 65 times, which is much larger than that of other groups, indicating that students in this group take listening to lectures in class as the main knowledge acquisition channel, and the classroom network trajectory mainly follows the teacher, which belongs to classroom-conformist learners.

4) The number of behavioral pattern identifications in cluster group 4 is basically the same as that of cluster group 3, but the frequency of distraction in this group is quite high, indicating that this group of students lacks network motivation, has little self-control, and belongs to the classroom wandering learners.

Table 2. Learning behavior clustering analysis resu1s

Clustering basis	Cluster 1 (N=8)	Cluster 2(N=12)	Cluster 3(N=31)	Cluster 4(N=7)	Total (N=58)
9→6	11	11	10	10	10.5
9→2 10→9	10	9	9	6	8.5
3→10	12	14	9	7	10.5
2→9	8	8	7	6	7.25
Average performance	97	90	78	65	82.5

3.1.2. Correlation between online behavioral pattern recognition and achievement. A combination of 12 learning behaviors generated by 58 students in a class of freshmen at the University of H produces 144 behavioral sequences. In order to explore the association between behavioral sequences and academic performance, this study obtained 14 behavioral sequences significantly related to academic performance through correlation analysis, and the behavioral sequences significantly related to online Civic Education are shown in Table 3. According to the theory of correlation analysis, if the correlation coefficient |r| is above 0.85, it means that there is a very strong correlation between the two variables, and it is evident that all the behavioral sequences that are significantly related to cyber Civic and political education are very strongly correlated with academic achievement. However, all of these behavior sequences significantly related to academic performance are scattered and do not form a unified pattern of conversion of behavioral categories, so it is necessary to analyze them with a more focused learning behavior pattern recognition in order to obtain a more accurate learning behavior pattern.

Table 3. A series of behavioral sequences that are significantly related to the Internet

Behavior classification sequence encoding	Behavior sequence coding	r	Significance
A→E	2→9	0.9944	0.0055**
B→A	4→11	0.9248	0.0499*
B→E	3→10	0.9155	0.0496*
C→A	5→1	0.8599	0.0307*
D→B	8→4	0.9184	0.0496*
D→C	7→5	-0.9933	0.0054**
D→E	7→1	0.9924	0.0046**

	8→1	0.9683	0.0493*
E→A	1→3	0.9099	0.0296*
E→C	11→5	0.9138	0.0304*
	1→5	0.9243	0.0303*
	5→5	0.8539	0.0399*
E→D	1→7	-0.9743	0.0107*
	1→8	0.9221	0.0205*

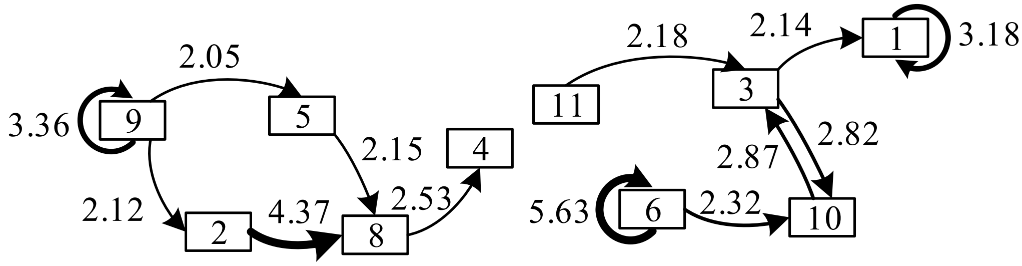
3.1.3. Sequence Transformation of Learning Behavior Based on Lagged Sequence Analysis Method. The lag sequence analysis was used to obtain the frequency table of online behavior pattern recognition for the four types of learners. The lag sequences of the four types of network behaviors are shown in Figure 3, (a, b, c, and d represent the network behavior pattern recognition results of integrated learners, autonomous learners, compliant learners, and stray learners, respectively). Where the direction of the arrow indicates the order of conversion, the thickness of the arrow and the number on the line indicate the degree of significance of the conversion (the thicker the line and the larger the number, the stronger the significance).

1) Classroom integrated learners' network behavior pattern recognition has continuity, and the length of the longest network behavior sequence produced is 6, which indicates that such learners behave well when they network ungrasped knowledge points in the classroom; at the same time, the network behavior pattern recognition is more concentrated, with strong structure, and there are no stray individual behaviors, which indicates that such learners have good classroom networking habits.

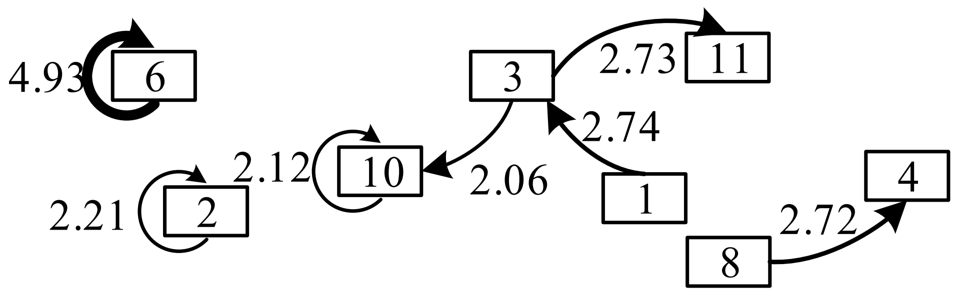
2) The length of the longest network behavior sequence produced by classroom autonomous learners is 4, indicating that such learners have good classroom practice habits. At the same time, this type of learners appeared significant for individual network behavior pattern identification, and no other significant network behaviors appeared after the two network behavior pattern identification of 3 (thinking) and 6 (doing exercises), indicating that this type of learners preferred to network autonomously in the classroom.

3) Classroom compliant learners' network behavior pattern recognition is more dispersed than the first two types of students, and a single network behavior pattern recognition occurs. 1 (listen to lecture), 6 (do exercises), the two network behavior pattern recognition does not appear after other significant network behaviors, indicating that this type of learners may just follow the teacher to carry out a shallow network.

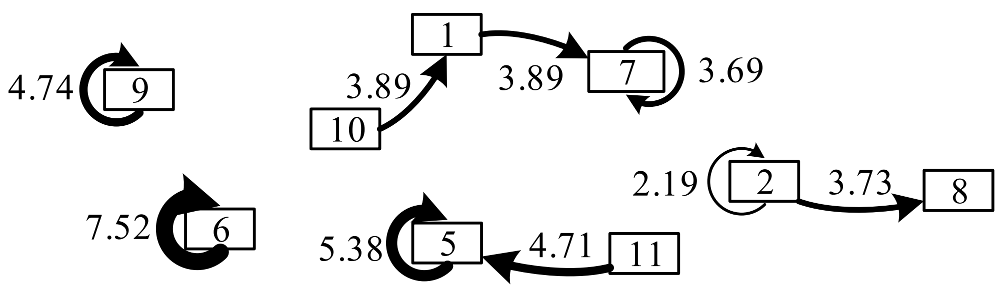
4) Classroom wandering learners had the most scattered network behavior pattern identifications among the four types of learners and the most individual network behavior pattern identifications, three of which were 1 (listening to lectures), 10 (revising exercises), and 5 (taking notes), suggesting that this type of learner is more of a passive networker in the classroom network. Although the length of the longest network behavior sequence produced by this type of learners is 5, and their pattern recognition has continuity, the good network habit of this type of learners can not be maintained without teacher's supervision because the starting behavior is being asked questions and transformed by teacher's guidance.



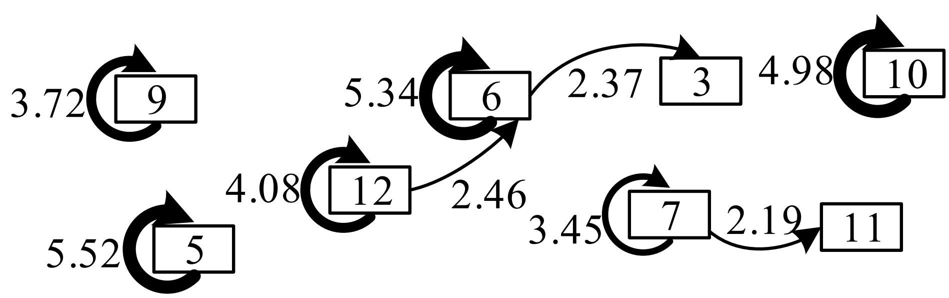
(a)Network behavior sequence conversion of synthetic learners



(b)The network behavior sequence transformation of autonomous learners



(c)The network behavior sequence transformation of the compliant learner



(d)The network behavior sequence transformation of free learners

Fig. 3. Correlation between online behavior and performance

3.2. *Correlation Analysis of Online Civic Education and Achievement*

In this study, Pearson correlation analysis was used to analyze the relationship between the frequency of learning behaviors and grades under the four learning styles.The correlation analysis between the

12 network behavior codes and the learning grades is shown in Table 4. The learning behaviors were taken as the frequency of 12 learning behaviors per lesson for each student, and the examination results were chosen from the monthly examination of Civic Education in the month when the lessons were recorded. Among them, comprehensive and autonomous students did not show 8→4 and 7→5 learning behaviors, and compliant and stray students did not show 4→11 learning behaviors.

Table 4. The correlation between 12 behavioral coding and academic performance

Behavior sequence coding	Complex type		Autonomous type		Adaptation type		Free type	
	Pearson correlation	Sig.	Pearson correlation	Sig.	Pearson correlation	Sig.	Pearson correlation	Sig.
2→9	0.1009*	0.1008	0.1026	0.1013	0.1034	0.1014	0.1009	0.0983
4→11	0.0314	0.983	0.0977	0.0979	-	-	-	-
3→10	0.0995	0.1018	0.1023	0.1011	0.1016	0.0993	0.1013	0.0997
5→1	0.0994	0.095	0.0963	0.0923	0.0944	0.0945	0.095	0.098
8→4	--	--	0.098	0.0976	0.0951	0.0962	0.0972	0.0983
7→5	-	-	0.1011	0.101	0.1002	0.1015	0.1003	0.0986
7→1	-0.0983	0.1004	0.1026	0.0983	-0.3916*	0.0314	0.1017	0.0991
8→1	-0.0952	0.1035	0.1021	0.0978	-0.3924*	0.0289	0.1107	0.0873
1→3	0.0996	0.0998	0.1006	0.0999	0.0996	0.0986	0.1	0.1019
11→5	0.1	0.1	0.1001	0.0971	0.1006	0.0998	0.0986	0.0998
1→5	0.0912	0.1005	0.0911	0.0672	0.0905	0.0983	0.0958	0.0874
5→5	0.0923	0.1003	0.0922	0.0931	0.0925	0.0910	0.0987	0.0924
1→7	0.0996	0.1013	0.1021	0.1024	0.1017	0.1002	0.1014	0.1008
1→8	0.0956	0.0997	0.0721	0.0638	0.0856	0.0938	0.0918	0.1002

3.3. Analysis of the effect of Internet Civic Education Intervention

3.3.1. Analysis of intervention performance. The results of the comparison of quiz scores before and after the intervention are shown in Figure 4. A second quiz was administered in week 4 of the course and a paired-samples t-test was performed on the pre- and post-intervention scores. The mean grade of the 58 students in the first stage quiz improved from 37.42 (standard deviation of 10.2531) to 59.33 (standard deviation of 13.9265) after the intervention, which is a significant difference ($p=0.000<0.01$), indicating that the intervention on the online self-directed learning behavioral model of the compliant learners was able to improve their academic performance. This suggests that the online self-directed learning behavioral patterns of integrative learners have a better guiding effect on compliant learners, and that the internal logic behind their behaviors can promote compliant students to process and reflect on their academic knowledge in depth, thus improving their online learning outcomes.

Overall, 58 students' grades were improved to different degrees after the intervention, but there were still a few students whose grades did not improve. This is consistent with the fact that Civic Education is a course that emphasizes the integration of theory and practice, and some students need a period of time to study and adjust before they can understand the content of the Civic Education course and find a learning pace that suits them.

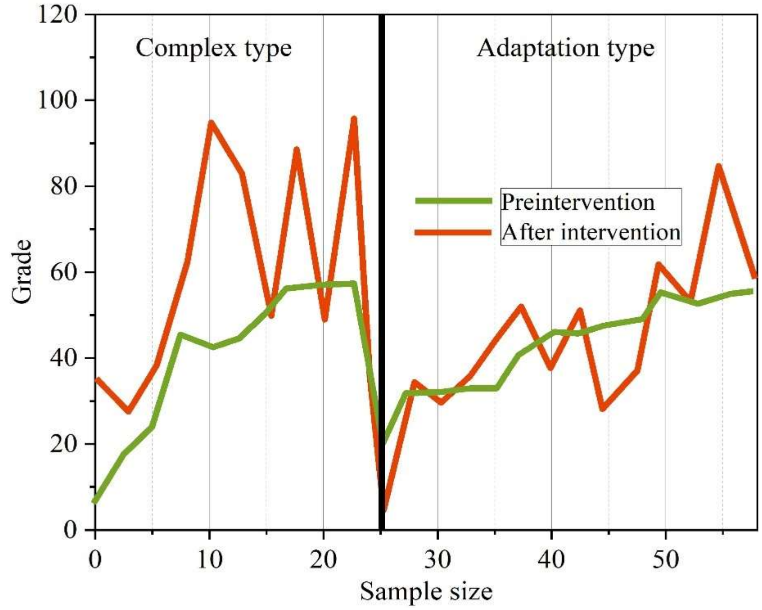


Fig. 4. Comparison of test results before and after intervention

3.3.2. Intervention effectiveness survey analysis. In this paper, the questionnaire was designed using Likert scale (strongly agree, agree, average, disagree, strongly disagree), including four aspects of online learning: learning effect, motivation, motivation, and experience (Cronbach's α = 0.927), and the results of the questionnaire and interview of the students after the intervention are shown in Table 5. From the questionnaire results, it can be seen that, similar to the empirical results, most of the students held a more positive attitude towards the learning effect after the intervention. The percentage of students who thought that the intervention on the learning behavior model improved the learning effect of Civic Education reached 75.22% (the sum of agreeing and strongly agreeing). 67.7% and 77.54% of the learners, respectively, believed that the intervention on the online learning behavioral model improved their motivation and learning motivation. Students were willing to take the initiative to follow the online learning behavioral model provided by the instructor and there was a tendency for their motivation to change from external motivation to internal motivation subconsciously. Meanwhile, 79.04% of the learners believed that the learning behavior model intervention reduced the sense of bewilderment in online learning and made the otherwise helpless learning comfortable, and were willing to continue to follow the intervention learning behavior model for independent learning. Most of the students indicated that the provision of the learning behavior model intervention could give them direction and confidence, reduce the sense of confusion and powerlessness, save a lot of fumbling time, and improve their online learning performance.

Table 5. Questionnaire survey and interview results of the intervention

Dimension	Very agree (%)	consent(%)	general(%)	Different meaning(%)	Very different(%)
Learning effect	21.74	53.48	22.56	2.22	0
Positivity	4.88	62.82	31.17	1.13	0
Learning motivation	10.12	67.42	22.46	0	0
Experience	14.92	64.12	20.96	0	0

4. Conclusion

The article adopts the lag sequence analysis method to analyze the sequence conversion of students' different online learning behaviors, explores the problems of different types of students, and designs a teaching intervention mechanism based on problematic learning behaviors to help learners change their learning behavior patterns and improve their learning effects.

In this study, a sequential clustering method was used to cluster online behaviors into four categories of “integrated, autonomous, compliant, and deviant” according to six levels of coding, and their corresponding characteristics were introduced. In addition, the results of the study show that there is a strong correlation between the 12 online behaviors in the classroom, and their corresponding correlation coefficients $|r|$ are all > 0.85. According to the significance of the sequence of the four online behaviors in the classroom, this paper constructs a sequence transformation diagram for online behavior pattern recognition, which presents the significant correlation between the behaviors in a clear way. The results of students' performance after the classroom teaching intervention indicate that the online self-directed learning behavior pattern of integrative learners has a better guiding effect on compliant learners. It can promote compliant students to process and reflect on academic knowledge in depth, thus enhancing online learning.

Funding

The 2022 project of higher education reform and innovation in Shanxi Province “The Construction of Talent Cultivation and Management System for ‘Major Enrollment and Major Diversion’ from the perspective of ‘Three-Wide Education System’ ——Taking the School Of Mechanical Engineering, North University of China "(project number: J20220580).

References

[1] Sahin, C. (2018). Social media addiction scale-student form: the reliability and validity study. Turkish Online Journal of Educational Technology-TOJET, 17(1), 169-182.

[2] Kulkarni, A. (2017). Internet meme and Political Discourse: A study on the impact of internet meme as a tool in communicating political satire. Journal of Content, Community & Communication Amity School of Communication, 6.

[3] Knoke, D., & Yang, S. (2019). Social network analysis. SAGE publications.

[4] Gao, H. W. (2023). Innovation and development of ideological and political education in colleges and universities in the network era. International Journal of Electrical Engineering & Education, 60(2_suppl), 489-499.

[5] Qiu, W. F., Ma, J. P., Xie, Z. Y., Xie, X. T., Wang, C. X., & Ye, Y. D. (2023). Online risky behavior and sleep quality among Chinese college students: The chain mediating role of rumination and anxiety. Current psychology, 42(16), 13658-13668.

[6] Ji, Y., & Zhang, T. (2024, July). Research on the Teaching Mode of Ideological and Political Courses in Colleges and Universities under the Background of “Internet plus”. In Proceedings of the 2nd International Conference on Educational Knowledge and Informatization (pp. 83-87).

[7] Abbas, J., Aman, J., Nurunnabi, M., & Bano, S. (2019). The impact of social media on learning behavior for sustainable education: Evidence of students from selected universities in Pakistan. Sustainability, 11(6), 1683.

[8] Huang, C. (2017). Time spent on social network sites and psychological well-being: A meta-analysis. Cyberpsychology, Behavior, and Social Networking, 20(6), 346-354.

[9] Liu, D., Kirschner, P. A., & Karpinski, A. C. (2017). A meta-analysis of the relationship of academic performance and Social Network Site use among adolescents and young adults. Computers in human behavior, 77, 148-157.

[10] García-Umaña, A., & Tirado-Morueta, R. (2018). Digital media behavior of school students: abusive use of the internet. Journal of New Approaches in Educational Research, 7(2), 140-147.

[11] Bory, P. (2020). The internet myth: From the internet imaginary to network ideologies (p. 169). University of Westminster Press.

[12] Meng, T. (2023). Research on the influence of new media on the ideological and political education of college students in the background of the Internet and countermeasures. Applied Mathematics and Nonlinear Sciences.

[13] Dawei Zhang. (2024). Behavior recognition algorithm based on a dual-stream residual convolutional neural network. Journal of Intelligent Systems(1),

[14] Shahad Aljehani & Youseef Alotaibi. (2025). Preserving privacy in association rule mining using multi-threshold particle swarm optimization. Information Sciences121673-121673.

[15] Deng Fan,Wang Jiabin & Lv Sheng. (2024). Optimization of frequent item set mining parallelization algorithm based on spark platform. Discover Computing(1),38-38.

[16] Ziyi Zhang,Diya Li,Zhe Zhang & Nick Duffield. (2024). Mining Spatiotemporal Mobility Patterns Using Improved Deep Time Series Clustering. ISPRS International Journal of Geo-Information(11),374-374.