On Learning Path Planning Algorithm Based on Collaborative Analysis of Learning Behavior

ZHAOYU SHOU1, XIANYING LU1, ZHENGZHENG WU1, HUA YUAN1, HUIBING ZHANG2, AND JUNLI LAI1

1School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, China
2School of Computer and Information Security, Guilin University of Electronic Technology, Guilin 541004, China

Corresponding author: Zhaoyu Shou (guilinshou@guet.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61967005, Grant 61662013, and Grant U1501252, in part by the Innovation Project of GUET Graduate Education under Grant 2020YCS022, and in part by the Key Laboratory of Cognitive Radio and Information Processing Ministry of Education under Grant CRKL190107.

ABSTRACT Studying learning path planning can help find useful implicit learning behavior patterns from learners’ online learning behavior data, which is conducive to helping beginners or learners with low participation to reasonably arrange the learning sequence of online knowledge points. This paper proposes a learning path planning algorithm based on collaborative analysis of learning behavior through collaborative data analysis of online learning behaviors. The algorithm, based on the learner’s online learning behavior data set, first establishes the concept interaction degree model of knowledge points and the directed learning path network, and proposes a local structure similarity measurement method between the knowledge nodes of the directed learning path network. Second, based on the learner’s Kullback-Leibler divergence (KLD) matrix, a learning behavior similarity calculation method on the basis of eigenvector matrix similarity is proposed, which is used to perform cluster analysis on learners with similar learning behaviors and to analyze the personalized optimal learning path of each kind of learners. Finally, the clustering algorithm and the evaluation index of the directed complex network have both verified the advantages of the algorithm. This paper employs the online behavior data set and online test data set obtained from the e-learning platform to conduct an empirical analysis of the learning path planning algorithm proposed hereof. The results show that the learner’s learning effect has been improved, verifying the validity and reliability of the algorithm.

INDEX TERMS Local structural similarity on the knowledge node of the directed learning path network, similarity measure on learning behavior, concept interaction degree of knowledge points, individualized optimal learning path.

I. INTRODUCTION

In January 2018, Minister Baosheng Chen, Secretary of the Leading Party Members’ Group, of Ministry of Education, pointed out in the “Speech at the National Conference on Education” that it is high time to deepen education reform and accelerate the pace of education informatization. Big data, behavior analysis and other information technologies have played a huge role in achieving the sharing of digital educational resources, promoting comprehensive reform and development in the field of education, enhancing the learning effect of learners, and promoting the equity of modern higher education resources [1], [2]. Sidong Xiong, deputy of the National People’s Congress and president of Suzhou University, believes that there are obvious characteristics of inadequate and uneven development in the popularization of higher education. The expansion of higher education has highlighted the paradox of “quantitative” expansion and “qualitative” stability and improvement of higher education. The problems of popularization, personalization and quality control of higher education must be properly handled. Currently, educational resources have failed to meet the demand of a growing number of the to-be-educated. There are also large regional gaps as to the educational resources available to universities. In recent years, the integration and innovation of information technology and education has not only produced new forms of education such as MOOCs and micro-courses [3], [4], but also launched some sound online
learning platforms such as Sakai, ILIAS and ATutor, bringing out the “Platform + Education” service model [5]–[7]. A large number of online learning platforms and a constantly improving education management system have recorded various learning behavior data of learners. With the application of teaching management systems and the proliferation of online learning systems, various bits of students’ learning behavior data could be successfully recorded, thus constituting the educational big data.

Educational big data refers to data generated by individuals in the entire process of educational activities, through the integration of which, existing educational or teaching problems could be diagnosed from the massive, complex and diverse sources, so as to evaluate the teaching and forecast development trends, explore educational or teaching models, investigate the implicit correlation between learning behavior data and the learning effect, and achieve targeted personalized education [8]. Thereby, it plays a crucial role in improving the quality of education management [9], promoting personalized management of students [10], and assisting scientific decision-making in universities [11].

In offline learning, with educational big data, teachers are able to promptly urge students to learn, while providing timely feedback on the questions raised by them. This is also the reason why online education cannot replace offline education. In offline education, due to the large number of learners brought by the expansion of enrollment, it is difficult for teachers to teach according to the learner’s personal situation, so online learning is an important supplement to offline education [12]. To this end, online education platforms should enjoy better learning path planning algorithms to help students better understand how to learn. More and more researchers are applying personalized learning path planning technology to online learning [13]. The researchers analyze the learner’s historical log information to mine the learner’s learning habits and characteristics, so as to plan a learning path for the learner that meets their own needs. The use of learning path planning algorithms will further increase learners’ enthusiasm for learning online resources, which to a certain extent helps to improve the learning effect, while learners can freely arrange the time period and location of learning. Online education can not only make up for the shortcomings of offline education, but also alleviate the “one-to-many” dilemma in offline teacher-student interactions, and efficiently increase the availability of learning resources.

Apart from that learners can freely choose online learning resources anytime and anywhere, online learning systems can accurately record learners’ behavior data, log information and others when they learn online knowledge points. Through behavior analysis, the educators are able to file learners’ personal learning reports, summarize the learning habits in their process of learning online knowledge points, so as to plan a learning path suitable for their own learning characteristics. The study of learning path planning constitutes an important part for personalized education. It is the cornerstone of fully realizing the concept of “teaching according to the aptitude”, and provides feedback to the teaching cognition and teaching behavior process for teachers to change their teaching methods and enhance teaching effects. In view of the above, the core issue for the current study lies in how to analyze the online learning behavior data, test data and learning log data of learners in the online education system to provide learners with personalized learning paths. Complex networks have strong theoretical advantages in the method of constructing individualized learning paths. They can be used in time series prediction [14], and individualized map generation [15]. The learning behavior data of a course is the research object of this paper. In the theory of complex networks, a learning path planning algorithm based on collaborative analysis of learning behaviors is proposed to provide personalized learning paths for groups with different levels of knowledge point interaction. The main contributions of this paper are as follows:

First, based on the KLD, a local structural similarity calculation method for directed network knowledge points is proposed to better characterize the similarity measure of the in-degree and out-degree of directed network nodes.

Second, based on the learner KLD matrix, a method for characterizing learning behavior similarity by using vectors of eigenvalues is proposed.

The main content of this paper is as follows: Section 2 introduces the related work of this paper. Section 3 introduces the relevant definitions and related formulas of the algorithm. In section 4, the data set of the study is introduced for the reliability analysis of clustering algorithm and the experimental analysis of directed learning-path-planning algorithm based on the concept interaction degree of knowledge points proposed in this paper. Section 5 analyzes the proposed algorithm experimental result. Section 6 concludes the work and looks ahead.

II. RELATED WORKS

Various educational management systems and online learning systems can accurately record learners’ learning processes online. The record combined with the learner’s featured online learning behavior data can clearly characterize their online learning path and behavior feature. As we can see, on the basis of effective record and comprehensive evaluation of online learning status, the analysis, diagnosis and recommendation of online learning behavior and the targeted online learning path planning can save learners time to learn online knowledge points and improve their effectiveness. At present, the studies of learning path planning algorithms can be divided into three categories: the first refers to those that combine learning behavior data with graph theory or complex network theory to generate personalized learning paths; the second represents those that are combined with ant colony algorithm or genetic algorithm to generate personalized learning paths; and studies of the third category are combined with other methods to generate personalized learning paths.
In the study of learning path planning algorithms based on graph theory or complex network theory: Yang et al. [16] modeled learning resources, built a knowledge map based on the online knowledge points learners learned, and used a knowledge map matching algorithm to propose a personalized learning path planning algorithm based on knowledge map matching. By judging the learner’s cognitive level, the degree of correlation between various courses and other factors, the designer can push learner with different learning characteristics their “best match” of personalized learning path. In order to solve the problem of group-based recommendation lacking dependence on learning time series and knowledge, Zhu et al. [17], [18] mapped the learner’s learning log to the personal learning generation network according to the knowledge map, and proposed a metric for measuring the similarity between any two learner’s learning generation networks. According to such similarity, a clustering algorithm could be executed to obtain a learning group and the group learning network. When learners are faced with a large number of learning resources and must balance the limited time with multiple objectives, Zhu et al. [19] considering that learners may have different learning behavior characteristics and path preferences in the four different scenarios of initial learning, general review, pre-test learning and pre-test review, proposed a multi-constrained learning path planning algorithm based on knowledge map. In order to reduce the influence of recommendation quality on beginners or learners with low participation, Liu and Li [20] proposed a combined learning path recommendation method based on learning networks, which uses complex network theory to analyze the online learning relationship between the course and the learner, and constructs the course network and the learner network separately. The learners were divided into three categories in consideration of the three learning scenarios and the characteristics of the course and the learner. The researchers recommended suitable learning paths for different learners, and conducted an experimental analysis of the actual data of MOOC to verify the effectiveness and reliability of the method. Shi et al. [21] proposed a learning path recommendation algorithm based on multidimensional knowledge graph. The algorithm stores learning objects in multiple classes, builds a multi-dimensional knowledge graph framework based on the semantic relationship between the learning objects in the knowledge graph, and generates personalized learning paths that meet different learning goals.

In the study of learning path planning algorithm based on ant colony algorithm or genetic algorithm: Dharshini et al. [22] proposed an automatic learning method for online learning objects based on constraint sets using ant colony algorithm to help online learners obtain in a shorter time the appropriate learning path. Kamsa et al. [23] proposed an improved learning optimization method for online collaboration, which uses an ant colony algorithm to automatically establish an optimal path to guide each learning community. Rastegarmoghadam et al. [24] put forward an adaptive optimal learning path generation method of improved ant colony algorithm. This method can help find the path suitable for the learner’s characteristics with maximum efficiency, and take it as the main goal to find the optimal learning path based on self-organization. An adaptive online learning model proposed by Birjali et al. [25] is a learning path planning algorithm based on the combination of genetic and ant colony algorithms of MapReduce, which can provide learning goals and adaptively generate learning paths for each online learner. Considering the needs and cognitive abilities of learners and the collective experience of co-learners, Kamsa et al. [26] came up with a learning path algorithm of ant colony optimization that combines individual and group learning experience. The algorithm adaptively and dynamically finds the optimal path which is more suitable for learners in terms of learning preferences, knowledge levels and historical data on learning, and can improve learner performance and satisfaction. Vanitha et al. [27] combined ant colony optimization and genetic algorithms to propose a collaborative optimization learning path planning algorithm. The algorithm matches the learner characteristics with the learning content sequence, and gives personalized learning paths to online learners, thereby improving their online learning capabilities.

There are also studies of learning path planning algorithms based on other methods: From the perspective of context-aware computing, Intayoad et al. [28] focused on the social context of the interaction between learning objects and learners and raised a context-aware learning path planning algorithm to promote effective personalized online learning for each learner. The algorithm uses K-nearest neighbors and decision trees to classify the collected social context, uses association rules to recommend suitable learning paths and can plan suitable learning paths for different groups of online learners. Ni et al. [29] presented a learning path recommendation system based on educational big data by establishing an initial fuzzy recommendation requirement set, establishing an advanced fuzzy recommendation requirement set, designing a fuzzy recommendation algorithm, and designing a framework and functional structure for a fuzzy recommendation system of learning content in sequence. The system selects appropriate learning content based on the learner’s situation, and informs the learner of real-time information. Su et al. [30] consider that learning style influences learners’ preference for specific textbooks, learning results, and the choice of learning path. Based on the Kolb’s learning style scale, four course learning paths of different learning styles were generated for learners to better study courses. The learning path generated by the mathematical model proposed by Nigenda et al. [31] matches the learner’s learning style, thereby improving their learning ability. Tang et al. [32] put forward a reinforcement learning method to obtain a learning path based on the measurement model of the learner’s skill profile and the learning model of the learning effect. Han et al. [33] put the learning path planning in the framework of reinforcement learning to predict the familiarity of online learners with online knowledge points, so as to
better model the learner’s personalized learning and generate personalized optimal learning path. Saito and Watanobe [34] considered the learner’s past learning data, current learning ability and learning goals, used the learner’s historical learning data to construct a learning ability map, and proposed a learning path planning algorithm based on time series prediction model and recurrent neural network. Zhang et al. [35] proposed a learning path recommendation algorithm based on the feature similarity measurement of learners. The algorithm relies on clustering and machine learning technology. Firstly, it clusters a group of learners, then trains long-term short-term memory (LSTM) model to predict their learning path and performance, and finally selects personalized learning path from the results of learning path prediction.

To sum up, the current research on the individualized path of a group has yet to combine online learning behavior with knowledge points. The learning effect of the learners indicates that only the individualized path generated by a group of learners with the same concept interaction degree as to the knowledge point is of reference value to the group. Therefore, the individualized path of groups proposed in this paper features combining learners’ online learning behavior data to provide individualized learning paths for group learners with the same interaction degree of knowledge points. In this paper, the learner’s learning behavior data of online knowledge point video is modeled to obtain the learner’s concept interaction degree of online video knowledge points which is combined with the directed weighted complex network theory. Then, the clustering algorithm and the optimal path algorithm are used to generate a personalized learning path.

III. CORRELATION DEFINITION

This section describes the relevant definitions and calculation methods of the proposed algorithm, and analyzes and explains some of the definitions.

A. CONCEPT INTERACTION DEGREE OF KNOWLEDGE POINTS

Online video knowledge points refers to the orderly knowledge point videos recorded by the teacher, which are numbered by the researcher, according to the course knowledge framework of courses. When learners study online knowledge points, whether they are proficient in using each knowledge point is portrayed by virtue of the concept interaction degree of knowledge points (CKP), the learner’s mastery degree of knowledge points (MKP) and the relative difficulty coefficient of knowledge points (RKP), as demonstrated below:

\[
ckp_i = \frac{mkp_i}{rkp_i}
\]

where \(mkp_i\) is the element of \(MKP = \{mkp_i | i = 1, \cdots, m\}\), the set of learner’s mastery degree of the knowledge points through his online learning. \(MKP\) indicates the degree of mastery of knowledge points by learner \(e\) [18]. And \(rkp_i\) is the element of \(RKP = \{rkp_i | i = 1, \cdots, m\}\), the set of relative difficulty coefficient of the knowledge points through his online learning. \(RKP\) is characterized by the learning behavior characteristics of learners, such as the frequency of learning video knowledge points, the cumulative learning time, and the frequency of pause and drag [18]. And \(m\) represents the number of knowledge points of online videos that learners have learned.

B. A DIRECTED LEARNING PATH NETWORK BASED ON THE CONCEPT INTERACTION DEGREE OF KNOWLEDGE POINTS

A directed weighted learning path network based on the concept interaction degree of knowledge points is a topological network generated on the basis of the learning-behavior time series data of learners learning knowledge points online. Among them, knowledge nodes in the network represent video knowledge points that learners learn online. Knowledge node values are characterized by concept interactions of knowledge points. Directional edges between knowledge nodes are characterized by the sequential order of learners’ learning knowledge points. The weights of the edges are jointly portrayed through the concept interaction degree of knowledge points in each knowledge node. The directional weighted learning path network based on the concept interaction degree of knowledge points is shown in formula (2):

\[
DLPN = G(M, CKP, E, W)
\]

where, \(DLPN\) characterizes a directed weighted learning path network based on the concept interaction degree of knowledge points; \(M\) represents the set of knowledge nodes; \(CKP\) characterizes the concept interaction degree of knowledge points; \(E\) represents the set of edges between the knowledge nodes; and \(W\) represents the weight matrix between the knowledge nodes. The above variables are defined as shown in formula (3):

\[
\begin{align*}
M &= \{1, \cdots, m\} \\
CKP &= \{ckp_i | i = 1, \cdots, m\} \\
E &= \{e_j | i = 1, \cdots, p\} \\
W &= [w_{ij}]_{m \times m}
\end{align*}
\]

where, \(p\) represents the sum of all edges between all knowledge nodes in \(DLPN\), and the calculation method of the edge weight \(w_{ij}\) from knowledge node \(i\) to knowledge node \(j\) is shown in formula (4):

\[
w_{ij} = \frac{ckp_i}{ckp_j} \quad (i, j = 1, \cdots, m)
\]

As it demonstrates, \(ckp_i\) represents the concept interaction degree of knowledge points of the online video knowledge point \(i\), \(ckp_j\) represents the concept interaction degree of knowledge points of the online video knowledge point \(j\), and \(w_{ij}\) represents the strength of concept interaction degree of knowledge points from the online video knowledge point \(i\) to \(j\) when learners study online video knowledge points.
C. KNOWLEDGE POINT LOCAL STRUCTURE

In a learner e’s DLPN, DLPN-LSKN represents the local structure of knowledge node network of a directed weighted learning path network. Fig. 1 shows a schematic diagram of the local structure of a knowledge node in a certain learner’s DLPN.

![Diagram of Local Structure of Knowledge Nodes in DLPN](image)

The Directly Connected In-Degree (DCID) knowledge node set with the knowledge node i is defined as $DKN_ID(i)$, and the Indirectly Connected In-Degree (ICID) knowledge node set with knowledge node i is defined as $IKN_ID(i)$. The Directly Connected Out-Degree (DCOD) knowledge node set with the knowledge node i is defined as $DKN_OD(i)$, and the Indirectly Connected Out-Degree (ICOD) knowledge node set with the knowledge node i is defined as $IKN_OD(i)$. That is:

$$
\begin{align*}
DKN_ID(i) &= \{dkn_{ID}(k) | k = 1, \ldots, l_{DKN_ID}^{dkn_{ID}}(i)\} \\
IKN_ID(i) &= \{ikn_{ID}(k) | k = 1, \ldots, l_{IKN_ID}^{ikn_{ID}}(i)\} \\
DKN_OD(i) &= \{dkn_{OD}(k) | k = 1, \ldots, l_{DKN_OD}^{dkn_{OD}}(i)\} \\
IKN_OD(i) &= \{ikn_{OD}(k) | k = 1, \ldots, l_{IKN_OD}^{ikn_{OD}}(i)\}
\end{align*}
$$

(5)

where $dkn_{ID}(k)$ represents the DCID knowledge node with the knowledge node i, and $l_{DKN_ID}^{dkn_{ID}}(i)$ represents the number of elements of $DKN_ID(i)$; $ikn_{ID}(k)$ represents the ICID knowledge node with the knowledge node i, and $l_{IKN_ID}^{ikn_{ID}}(i)$ represents the number of elements of $IKN_ID(i)$; $dkn_{OD}(k)$ represents the DCOD knowledge node with the knowledge node i, and $l_{DKN_OD}^{dkn_{OD}}(i)$ represents the number of elements of $DKN_OD(i)$; $ikn_{OD}(k)$ represents the ICOD knowledge node with the knowledge node i, and $l_{IKN_OD}^{ikn_{OD}}(i)$ represents the number of elements of $IKN_OD(i)$.

The DCID corresponding to $DKN_ID(i)$ is defined as $DCI(i)$, and the ICID corresponding to $IKN_ID(i)$ is defined as $IC(i)$; the DCOD corresponding to $DKN_OD(i)$ is $DCO(i)$, and the ICOD corresponding to $IKN_OD(i)$ is $ICO(i)$.

$$
\begin{align*}
DCI(i) &= \{dni_{ID}(k) | k = 1, \ldots, l_{DCI_ID}^{dni_{ID}}(i)\} \\
IC(i) &= \{ini_{ID}(k) | k = 1, \ldots, l_{IC_ID}^{ini_{ID}}(i)\} \\
DCO(i) &= \{dni_{OD}(k) | k = 1, \ldots, l_{DCO_ID}^{dni_{OD}}(i)\} \\
ICO(i) &= \{ini_{OD}(k) | k = 1, \ldots, l_{ICO_ID}^{ini_{OD}}(i)\}
\end{align*}
$$

(6)

where $dni_{ID}(k)$ represents the degree corresponding to $dkn_{ID}(k)$, $ini_{ID}(k)$ represents the degree corresponding to $ikn_{ID}(k)$, $dni_{OD}(k)$ represents the degree corresponding to $dkn_{OD}(k)$, $ini_{OD}(k)$ represents the degree corresponding to $ikn_{OD}(k)$.

As shown in Fig. 1, the DCID knowledge nodes of knowledge node 6 are knowledge node 4 and 5, and the DCID of knowledge node 6 with knowledge node 4 and 5 are 2, 2, respectively. The ICID knowledge nodes of knowledge node 6 are knowledge node 1, 2 and 3. The ICID between the DCID knowledge node 4 and the knowledge node 1 and 2 are 1, 1 respectively. The ICID between DCID knowledge nodes 5 and knowledge node 2 and 3 are 1, 1, respectively. The DCID knowledge nodes of knowledge node 6 are knowledge node 7 and 8, and the DCID of knowledge node 6 with knowledge node 7 and 8 are 2, 2, respectively. The ICID knowledge nodes of knowledge node 6 are knowledge node 9, 10, 11 and 12. The ICOD of knowledge nodes 7 and indirectly connected knowledge nodes 9, 10, 11, 12 are 1, 1, and 1, respectively. The ICOD of directly connected knowledge node 8 and indirectly connected knowledge nodes 10, 11, 12 are 1, 1, and 1, respectively.

In the learner e’s DLPN, the in-degree knowledge node set of knowledge node i is defined as $IDKNI(i)$, namely:

$$
IDKNI(i) = DKN_ID(i) + IKN_ID(i) = \left\{dkn_{ID}(\alpha) \left(\alpha = 1, \ldots, l_{DKN_ID}^{dkn_{ID}}(i)\right)\right\} + \left\{ikn_{ID}(\beta) \left(\beta = 1, \ldots, l_{IKN_ID}^{ikn_{ID}}(i)\right)\right\}
$$

(7)

where $dkn_{ID}(i,k)$ represents the element in $IDKNI(i)$, $l_{DKN_ID}^{dkn_{ID}}(i)$ stands for the number of elements in the in-degree knowledge node set $IDKNI(i)$; $ikn_{ID}(i,k)$ represents the element in $IDKNI(i)$, $l_{IKN_ID}^{ikn_{ID}}(i)$ is represented by formula (8):

$$
l_{IDKNI_ID}^{dkn_{ID}}(i) = l_{DKN_ID}^{dkn_{ID}}(i) + l_{IKN_ID}^{ikn_{ID}}(i)
$$

(8)

The corresponding in-degree of the knowledge node set $IDKNI(i)$ is defined as $ID(i)$, namely:

$$
ID(i) = DCI(i) + IC(i) = \left\{dni_{ID}(\alpha) \left(\alpha = 1, \ldots, l_{DCI_ID}^{dni_{ID}}(i)\right)\right\} + \left\{ini_{ID}(\beta) \left(\beta = 1, \ldots, l_{IC_ID}^{ini_{ID}}(i)\right)\right\}
$$

(9)

where, $dni_{ID}(i,k)$ means the element in $ID(i)$, and the in-degree of the in-degree node $dkn_{ID}(i,k)$ is described as $idn(i,k)$.

Define the total in-degree of knowledge nodes i as $L_ID(i)$, as shown in formula (10):

$$
L_ID(i) = \left(\sum_{\alpha=1} dni_{ID}(\alpha) + \sum_{\beta=1} ini_{ID}(\beta)\right)
$$

(10)

In the learner e’s DLPN, the in-degree knowledge node set of knowledge node i is defined as $IDKNI(i)$,
namely:

\[ ODKN (i) = DKN_{OD} (i) + IKN_{OD} (i) \]

\[ = \left\{ dkn_{OD} (\alpha) (\alpha = 1, \ldots, l_{DKN}^{OD}) , \right. \]

\[ \left. ikn_{OD} (\beta) (\beta = 1, \ldots, l_{IKN}^{OD}) \right\} \]

\[ = \{ odn (i, k) | k = 1, \ldots, l_{ODKN}^{ID} \} \] (11)

where, \( odn (i, k) \) represents the element in \( ODKN (i) \), \( l_{ODKN}^{ID} \) stands for the number of elements in the in-degree knowledge node set \( ODKN (i) \); and \( l_{ODKN}^{ID} \) is represented by formula (12):

\[ l_{ODKN}^{ID} = l_{DKN}^{ID} + l_{IKN}^{ID} \] (12)

The corresponding in-degree of the knowledge node set \( ODKN (i) \) is defined as \( OD (i) \), namely:

\[ OD (i) = DCO (i) + ICO (i) \]

\[ = \left\{ dco (\alpha) (\alpha = 1, \ldots, l_{DKN}^{ID}) , \right. \]

\[ \left. ico (\beta) (\beta = 1, \ldots, l_{IKN}^{ID}) \right\} \]

\[ = \{ id (i, k) | k = 1, \ldots, l_{IDKN}^{OD} \} \] (13)

where, \( id (i, k) \) means the element in \( OD (i) \), and the in-degree of the in-degree node \( odn (i, k) \) is described as \( od (i, k) \).

Define the total in-degree of knowledge nodes \( i \) as \( L_{OD} (i) \), as shown in formula (14):

\[ L_{OD} (i) = \left( \sum_{\alpha=1} l_{DKN}^{ID} dci (\alpha) + \sum_{\beta=1} l_{IKN}^{ID} ici (\beta) \right) \] (14)

### D. KNOWLEDGE NODE LOCAL STRUCTURE SIMILARITY MEASURE

In order to calculate the local structural similarity between the knowledge nodes in the learner’s DLPN, the in-degree probability set \( P_{ID} [i] \) and the out-degree probability set \( P_{OD} [i] \) of each knowledge node in the learner’s DLPN should have the same length, and the maximum length of the in-degree and out-degree sets of all the knowledge nodes in the learner’s DLPN should be selected and defined as \( L \). When the length of the in-degree node set or the out-degree node set is less than \( L \), the remaining elements are set to 0. That is:

\[ L = \text{Max} \left( l_{IDKN}^{ID} (i), l_{ODKN}^{ID} (i) \right) \] (15)

\[ P_{ID} (i) = \{ p_{ID} (i, k) | k = 1, \ldots, L \} \] (16)

\[ P_{OD} (i) = \{ p_{OD} (i, k) | k = 1, \ldots, L \} \]

where, \( p_{ID} (i, k) \) and \( p_{OD} (i, k) \) are represented by the formula (17), which is:

\[ p_{ID} (i, k) = w_{ik} \ast \frac{id (i, k)}{L_{ID} (i)} \]

\[ p_{OD} (i, k) = w_{ik} \ast \frac{id (i, k)}{L_{OD} (i)} \] (17)

To better describe the local structural similarity between knowledge nodes, according to the literature review [35], \( P_{ID} (i) \) and \( P_{OD} (i) \) of each knowledge node need to be ordered, and the sorted in-degree probability degree set and out-degree probability set of the knowledge node \( i \) are defined as:

\[ P'_{ID} (i) = \{ p'_{ID} (i, k) | k = 1, \ldots, L \} \]

\[ P'_{OD} (i) = \{ p'_{OD} (i, k) | k = 1, \ldots, L \} \] (18)

To avoid the calculation results to infinitive, when \( (p'_{ID} (i, k) \neq 0) \), the in-degree KLD of the knowledge node \( i \) and knowledge node \( j \) calculated according to the sorted probability set should follow:

\[ H_{KL} (P'_{ID} (i) \mid P'_{ID} (j)) \]

\[ = \sum_{k=1}^{l_{ID}'} \left( p'_{ID} (i, k) \ast \ln \left( \frac{p'_{ID} (i, k)}{p'_{ID} (j, k)} \right) \right) \] (19)

where \( l_{ID} \) is represented as follows:

\[ l_{ID}' = \text{Min} \left( l_{IDKN}^{ID} (i), l_{IDKN}^{ID} (j) \right) \] (20)

The out-degree KLD of the knowledge node \( i \) and the knowledge node \( j \) is defined as follows:

\[ H_{KL} (P'_{OD} (i) \mid P'_{OD} (j)) \]

\[ = \sum_{k=1}^{l_{OD}'} \left( p'_{OD} (i, k) \ast \ln \left( \frac{p'_{OD} (i, k)}{p'_{OD} (j, k)} \right) \right) \] (21)

where \( l_{OD}' \) is expressed as:

\[ l_{OD}' = \text{Min} \left( l_{ODKN}^{ID} (i), l_{ODKN}^{ID} (j) \right) \] (22)

when \( p'_{ID} (j, k) = 0 \), \( H_{KL} (P' (i) \mid P' (j)) = 0 \) and \( H_{KL} (P' (i) \mid P' (j)) = 0 \).

Combining formula (19) and formula (21), the KLD matrix in the DLPN of learner \( e \) is obtained:

\[ KL (e) = [kl_{ij}]_{max} \] (23)

Since the KLD is not symmetrical [23], it is necessary to define each pair of knowledge nodes in the KLD matrix as follows:

\[ k_{ij} = H_{KL} (P (i) \mid P (j)) + H_{KL} (P (j) \mid P (i)) \] (24)

\[ k_{ji} = H_{KL} (P (i) \mid P (j)) + H_{KL} (P (j) \mid P (i)) \]

\[ H_{KL} (P (i) \mid P (j)) = H_{KL} (P'_{ID} (i) \mid P'_{ID} (j)) \]

\[ + H_{KL} (P'_{OD} (i) \mid P'_{OD} (j)) \] (25)

In formula (24), the values of \( H_{KL} (P (i) \mid P (j)) \) and \( H_{KL} (P (j) \mid P (i)) \) are included in \( k_{ij} \). The KLD matrix will become a symmetric matrix. The KLD matrix shows the relationship of the local structure between each pair of knowledge nodes in a directed learning path network. The greater the
correlation value based on the local structure of the knowledge node, the greater the difference in its local structure. The similarity of knowledge nodes is the other side of the difference in knowledge node structure. Therefore, the KLD value can be used to define the similarity of each pair of knowledge nodes of the directed learning path network.

**E. LEARNING BEHAVIOR SIMILARITY**

In order to improve the accuracy of clustering algorithms for clustering high-dimensional online learning behavior data, this section proposes matrix eigenvector similarity to characterize the similarity of KL divergence matrices among learners. The matrix eigenvector similarity is to obtain the KL divergence matrix similarity between learners by using the European norm similarity and cosine similarity. The matrix eigenvector similarity not only has the advantage of the traditional angle measurement method in multi-dimensional space, but also incorporates the idea of distance to make up for the defects of the traditional angle-based clustering algorithm [36]. The calculation method of matrix eigenvector similarity is as follows:

Calculate the eigenvalues of the learner $e$’s KL ($e$) matrix and the eigenvectors corresponding to the eigenvalues, sort the eigenvalues in descending order, take the top $h$ eigenvectors corresponding to the previous eigenvalues and construct the eigenvector matrix $V (e)_{hsm}$. Calculate the eigenvalues of the learner $f$’s KL ($f$) matrix and the eigenvectors corresponding to the eigenvalues, sort the eigenvalues in descending order, take the top $h$ eigenvectors corresponding to the previous eigenvalues, and construct the eigenvector matrix $V (f)_{hsm}$.

According to formula (26), calculate the Euclidean norm of $w_j$ vector of each row in the eigenvector matrix $V (e)_{hsm}$ and that of $w_j$ vector of each row in the eigenvector matrix $V (f)_{hsm}$.

$$
dis (w_i, w_j) = \left( \sum_{i,j=1}^{h} |w_i - w_j|^2 \right)^{\frac{1}{2}} \tag{26}$$

According to formula (27), calculate the Cosine similarity of $w_j$ vector of each row in the eigenvector matrix $V (e)_{hsm}$ and that of $w_j$ vector of each row in the eigenvector matrix $V (f)_{hsm}$:

$$
cos (w_i, w_j) = \frac{w_i \times w_j}{|w_i| \times |w_j|} = \frac{\sum_{i,j=1}^{h} (w_i \times w_j)}{\sqrt{\sum_{i=1}^{h} (w_i)^2} \sqrt{\sum_{j=1}^{h} (w_j)^2}} \tag{27}$$

The similarity between $w_j$ vector of each row in the eigenvector matrix $V (e)_{hsm}$ and that of $w_j$ vector of each row in the eigenvector matrix $V (f)_{hsm}$ is shown in formula (28):

$$
s_{w_jw_j} = dis (w_i, w_j) \cdot cos (w_i, w_j) \tag{28}$$

The similarity matrix $S_{ef}$ between the learner $e$’s eigenvector matrix $V (e)_{hsm}$ and the learner $f$’s eigenvector matrix $V (f)_{hsm}$ can thus be obtained:

$$
S_{ef} = [s_{w_jw_j}]_{hsm} \tag{29}$$

The value of the matrix determinant is the matrix similarity between the learner $e$’s KL ($e$) matrix and the learner $f$’s KL ($f$) matrix, i.e., the learning behavior similarity between the learner $e$ and the learner $f$. According to the classification of learners’ concept interaction degree of knowledge points, this chapter takes 3 to be $h$. According to formulas (29), the similarity matrix of KL divergence matrix of $N$ learners is defined as follows:

$$
S = [s_{ef}]_{N \times N} \tag{30}$$

**IV. LEARNING PATH PLANNING ALGORITHM BASED ON COLLABORATIVE ANALYSIS OF LEARNING BEHAVIOR**

The overall flow diagram of the algorithm in this paper is shown in Fig. 2.
Finally, the traditional DBSCAN algorithm is a similarity matrix that obtains data by distance. For the learner’s KLD matrix data, in order to better describe the similarity of learning behavior between learners and improve the clustering quality, the similarity matrix of course learning between each learner is obtained through calculating the distance between the vectors of eigenvalues of the learner’s KLD matrix and the cosine similarity. The obtained similarity matrix and the learner’s KLD matrix are used as the input of the DBSCAN clustering algorithm, and the learners who learn the course are divided into three categories of primary, intermediate and advanced levels according to the concept interaction degree of the knowledge points. The adjacency matrix of each class of learners is superimposed, and the generalized map of the learner’s directed learning path network is obtained. The path with the largest weight in the generalized map of this class of students is selected as the optimal path, and the individualized optimal learning path of the learner is obtained.

The learning path planning algorithm based on the collaborative analysis of learning behavior is shown in Algorithm 1.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this part, the DBSCAN algorithm based on the matrix eigenvector similarity (DBSCAN-EV) is compared with the SC algorithm based on the Euclidean norm similarity (SC-D) and DBSCAN algorithm based on the Euclidean norm similarity (DBSCAN-D) for clustering analysis, and the ACC and ARI are selected as the evaluation standard to judge clustering test results. Then, the clustering algorithm with better clustering effect is used to cluster the learning behavior data of the students of 2016, 2017 and 2018 (students enrolled in 2016, 2017 and 2018), after which the optimal learning path is selected. Last, the reliability of the selected optimal path is verified through the learning behavior data and achievement data of students of 2019.

A. DATA SETS

This experiment is based on the students’ learning behavior data of a school’s online learning platform. The course chosen for research is Data Structure and Algorithm. The data set is composed of 1198 learners’ learning data, including 228 students enrolled in 2016, 378 enrolled in 2017, 290 enrolled in 2018 and 302 enrolled in 2019. There are 207 video knowledge points in this course, including 207 knowledge points’ test information, learning behavior log information and students’ final exam results. The deadline for the amount of online learning behavior data and online test data obtained in this chapter is January 1, 2020. The mastery degree of the knowledge points, the relative difficulty coefficient of knowledge points and the concept interaction degree of knowledge points are shown in Fig. 3, Fig. 4 and Fig. 5.

B. CLUSTER RESULTS AND ANALYSIS

According to the literature [37], it can be known that spectral clustering and DBSCAN density clustering can have a high clustering effect on clusters of any shape. On this basis, the researcher employed the clustering method of distance measurement, and the DBSCAN algorithm which uses the similarity of eigenvalue vectors, to cluster the concept interaction degree about knowledge points of students enrolled in 2016, 2017 and 2018, respectively. The students’ concept interaction degree about knowledge points is divided into three levels, namely, primary, intermediate and advanced. According to the law of Gaussian distribution of academic performance, this study sets the following classification based on a final exam paper which carries a full mark of 100: Students with a test score between 60-75 are at the primary level of concept interaction about knowledge points; students with a test score between 76-89 are at the intermediate level, and those with test scores between 90-100 are at the

Algorithm 1 Learning Path Planning Algorithm Based on Collaborative Analysis of Learning Behavior

Input: Learner’s online learning behavior data set, $D = \{d_e | e = 1, \ldots, N\}$; MKP; RKP;
Output: Three types of learner’s optimal learning path according to the primary, intermediate and advanced concept interaction degree of knowledge points

1: According to MKP and RKP, combined with the definition 2.2, the edge weight matrix $W$ between CKP and knowledge nodes is generated.
2: for each $d_e \in D$ do
3: Construct learner $e$’s directed learning path network according to the definition 2.1
4: for $i < m$ do
5: According to the definition 2.3, construct the local structure of the knowledge node $i$; according to $DKN_{ID}(i)$, $IKN_{ID}(i)$, $DKN_{OD}(i)$, $IKN_{OD}(i)$ and $DCI(i)$, $ICI(i)$, $DCO(i)$, $ICO(i)$, calculate the $P'_{ID}(i)$ and $P'_{OD}(i)$ of the knowledge node $i$.
6: Get $P_{ID}[i]=1,\ldots,N$ and $P_{OD}[i]=1,\ldots,N$ of all the knowledge nodes in the learner $e$’s directed learning path network.
7: end for
8: Calculate the learner $e$’s KLD matrix $KL(e)$ and KLD matrix vector of eigenvalues $\nu_e$
9: Get the KLD matrix and KLD matrix vectors of eigenvalues of all learners
10: Get the learning behavior similarity matrix $S$ of learners
11: end for
12: According to the learner’s KLD matrix and the learning behavior similarity matrix $S$, the DBSCAN algorithm is used to derive the three types of learners of the knowledge point concept interaction level, primary, intermediate and advanced, and draw the optimal learning path for each type of learners.
13: return the optimal learning path of the primary, intermediate and advanced learners classified by their concept interaction degree of knowledge points
advanced level. According to the cognitive level of the learners, the learners are divided into three different groups of primary, intermediate and advanced concept interaction degree of knowledge points. According to the initial classification of learners’ concept interaction degree of knowledge points and the classification results of the three algorithms, the ACC and ARI analysis are performed, as shown in Table 1 and Table 2 respectively.

**TABLE 1. Clustering ACC analysis of three algorithms.**

| Grade | SC-D/ % | DBSCAN-D / % | DBSCAN-EV / % |
|-------|---------|--------------|--------------|
| 2016  | 66.7    | 68.1         | 75.5         |
| 2017  | 62.8    | 67.3         | 73.0         |
| 2018  | 59.4    | 61.4         | 72.8         |

**TABLE 2. Clustering ARI analysis of three algorithms.**

| Grade | SC-D/ % | DBSCAN-D / % | DBSCAN-EV / % |
|-------|---------|--------------|--------------|
| 2016  | 50.4    | 48.2         | 58.2         |
| 2017  | 49.1    | 47.5         | 54.7         |
| 2018  | 46.7    | 43.6         | 51.6         |

It can be seen from Table 1 that, compared with the SC-D and DBSCAN-D algorithms, the DBSCAN-EV algorithm increased ACC by 8.8% and 7.4% on the 2016 data set, respectively, by 10.2% and 5.3% on the 2017 data set, and by 13.4% and 11.4% on the 2018 data set.

It can be seen from Table 2, compared with the SC-D and DBSCAN-D algorithms, that the DBSCAN-EV algorithm has improved ARI by 7.8% and 10% on the 2016 data set, respectively. The accuracy rate has been improved on the 2017 data set by 5.6% and 7.2%, and the numbers for the 2018 data set are 4.9% and 8% respectively.

After clustering learners based on their learning behavior data, and comparing and classifying according to the cognitive level, the results of cluster analysis of learners’ learning behavior data show that the DBSCAN-EV algorithm is superior than SC-D algorithm and DBSCAN-D algorithm in terms of accuracy, ACC and ARI clustering indicators. Therefore, it can be said that the DBSCAN-EV algorithm can accurately cluster learners with similar learning behaviors.

Clustering learners through a clustering algorithm is to plan a learning path for learners of different levels that is more in line with their initial cognitive levels. Use DBSCAN-EV algorithm to classify the learning behavior data of learners enrolled in 2016, 2017 and 2018 by clustering primary, intermediate and advanced concept interaction degree of knowledge points to classify learners of each grade into primary, intermediate and advanced ones. The directional weighted learning path network of learners in each grade is analyzed, and the optimal learning path for learners in each grade can be obtained. Then the optimal learning paths of the three concept interaction degrees of knowledge points (primary, intermediate and advanced) obtained through the DBSCAN-EV clustering algorithm is recommended for the corresponding learners based on their cognitive levels, so as to better meet their learning demands. In the next section, the authors will use the DBSCAN-EV clustering algorithm to obtain and analyze the directed weighted learning path of the primary, intermediate and advanced learners, so as to acquire the optimal learning path for learners of each grade.
C. THE DIRECTED LEARNING PATH NETWORK RESULTS AND ANALYSIS

In this section, DBSCAN-EV algorithm is used to analyze online learning behavior data. The overlay network of original learning path for students enrolled in 2016, 2017 and 2018 is shown in Fig. 6, Fig. 7 and Fig. 8 respectively. Based on the directed learning path network algorithm of knowledge point concept interaction degree, the original learning behavior data of students enrolled in 2016, 2017 and 2018 were analyzed. These learners’ concept interaction degree of knowledge points be it primary, intermediate or advanced, is shown in the general map of directed learning path overlay network as in Fig. 9, Fig. 10 and Fig. 11 respectively.

Fig. 9(a), Fig. 10(a) and Fig. 11(a) represent the general map of the directed learning path overlay network for the students with primary concept interaction degree of knowledge point enrolled in 2016, 2017 and 2018 respectively. Fig. 9(b), Fig. 10(b) and Fig. 11(b) represent the general map of the directed learning path overlay network for the students with intermediate concept interaction degree of knowledge point enrolled in 2016, 2017 and 2018 respectively. Fig. 9(c), Fig. 10(c) and Fig. 11(c) represent the general map of the directed learning path overlay network for the students with advanced concept interaction degree of knowledge point enrolled in 2016, 2017 and 2018 respectively.

Cluster analysis is carried out for learners with similar learning behaviors, and then the learning path map of each kind of learning is superposed. The path with the greatest degree of concept interaction among knowledge nodes is taken as the optimal path. According to formula (4), the learning path with the greatest weight of concept interaction among knowledge nodes refers to the learning path with significant weight and a large number of learners in each kind of superposed map.

Fig. 12, Fig. 13 and Fig. 14 illustrates the learner’s optimal learning paths for the students of 2016, 2017 and 2018 whose knowledge point concept interaction level is intermediate, intermediate or advanced. Fig. 12(a), Fig. 12(b), and Fig. 12(c) represent the optimal learning path map for the students of 2016 with primary, intermediate and advanced concept interaction degree of knowledge points respectively. Fig. 13(a), Fig. 13(b), and Fig. 13(c) represent the optimal learning path map for the students of 2017 with primary, intermediate and advanced concept interaction degree of knowledge points respectively. Fig. 14(a), Fig. 14(b) and Fig. 14(c) represent the general map of the directed learning path overlay network for the students of 2018 with primary, intermediate and advanced concept interaction degree of knowledge points respectively.

In order to evaluate the performance of the algorithm, the topology characteristics of the real data set generation network will be analyzed. The topology features are as follows.

\[
C_D(i) = \frac{1}{N(N-1)} \cdot \sum_{j \in (DCI(i)+DCO(i))} (w_{ij} \cdot e_{ij})
\]

\[
C = \frac{1}{N} \sum_{i=1}^{N} \left( \frac{1}{s_i (k_i-1)} \sum_{k>j} \left( \frac{(w_{ij} + w_{ik})}{2} e_{ij} e_{ik} e_{jk} \right) \right)
\]

\[
L = \frac{1}{N(N-1)} \cdot \sum_{(d_{ij}, e_{ij}) \neq (i, j)} d_{ij}
\]

where \( w_{ij} \) is the weight from the knowledge node \( i \) to the knowledge node \( j \), \( e_{ij} \) represents the number of edges from the knowledge node \( i \) to the knowledge node \( j \), \( s_i = \sum_{j \neq j} (w_{ij} + w_{ik}) \) represents the weight set of the knowledge node \( i \), \( k_i \) stands for the degree of the knowledge node \( i \), and \( \sum_{k>j} (e_{ij} e_{ik} e_{jk}) \) represents the total number of triangles containing the knowledge node \( i \). \( N \) is the total number of network knowledge nodes, \( d_{ij} \) indicates the degree
of conceptual interaction of the shortest knowledge points between the knowledge node $i$ and knowledge node $j$.

In combination with formula (4), the average weighting of the knowledge node $i$ describes the average concept interaction degree of the knowledge points between the knowledge nodes directly connected to the knowledge node $i$ in the network. The average clustering coefficient $C$ stands for the average concept interaction degree between the knowledge nodes and their local nodes in the network. The average path length $L$ is defined as the average of the concept interaction degree of knowledge points between any two knowledge nodes.

According to formula (31), the average weighting, average clustering coefficient and average path length of the primary, intermediate and advanced levels of the students’ learning path network based on the concept interaction degree of the knowledge points are shown in Table 3, Table 4 and Table 5.

According to Table 3, Table 4 and Table 5, the experimental results of the average weighting, average clustering coefficient and average path length of the primary, intermediate and advanced levels of the students’ learning path network based on the concept interaction degree of the knowledge points are shown in Table 3, Table 4 and Table 5.

According to Table 3, Table 4 and Table 5, the experimental results of the average weighting, average clustering coefficient and average path length of the primary, intermediate and advanced levels of the students’ learning path network based on the concept interaction degree of the knowledge points are shown in Table 3, Table 4 and Table 5.
advanced levels of the students’ learning path network based on the concept interaction degree of the knowledge points are shown in Fig. 15, Fig. 16 and Fig. 17.

In Fig. 15, the abscissa shows the grade level, “Primary” indicates the primary learning path network based on the concept interaction degree of the knowledge points, “Intermediate” indicates the intermediate learning path network based on the concept interaction degree of the knowledge points, and “Advanced” indicates the advanced learning path network based on the concept interaction degree of the knowledge points. Fig. 15 shows the change of average weighting in the primary, intermediate, and advanced learning path networks based on the concept interaction degree of knowledge points for students of 2016, 2017 and 2018, respectively.
Fig. 16 shows the average clustering coefficient changes for the primary, intermediate, and advanced learning path network based on the concept interaction degree of the knowledge points for students of 2016, 2017 and 2018, respectively. Therefore, in Fig. 16 and Fig. 17, the difference of average clustering coefficient and the average path length among the primary, intermediate, and advanced learning path networks based on the concept interaction degree of the knowledge points gradually increase with time.

According to the above analysis results of the directed learning path network based on the concept interaction degree of knowledge points, the optimal learning path for students with primary, intermediate and advanced concept interaction degrees become available, which can be recommended to the online students of 2019 for reference according to their initial cognitive levels. Fig. 18, Fig. 19 and Fig. 20 have demonstrated the optimal learning path diagrams for the students of 2016 with primary, intermediate, and advanced concept interaction degrees of knowledge points.

It is recommended that the optimal learning paths derived from students of 2016 at primary, intermediate and advanced levels of concept interaction degree of knowledge points be used as reference for students of 2019. In order to test the effectiveness of such paths, before the students of 2019 studied this course, the 348 of them who chose this course had been divided into two groups: Group A students used the optimal learning paths of primary, intermediate and advanced levels for reference when studying this course; Group B students, on the contrary, did not use the optimal learning path of students of 2016. Secondly, according to their initial learning ability, the students in Group A and Group B were classified into three categories: primary, intermediate and advanced learners in terms of their concept interaction degree of knowledge points. Based on the final exam scores, the average scores of the two groups of students were calculated as shown in Table 6.

| Group Category | Primary | Intermediate | Advanced |
|----------------|---------|--------------|----------|
| Group A        | 70.2    | 83.41        | 92.57    |
| Group B        | 68.81   | 80.89        | 91.25    |

From Table 6, it can be seen from a horizontal perspective: The average score of group A was 82.06, and that of group B was 80.32. Group A has an average score improvement of 1.74 points compared to Group B, so compared with students who did not use the optimal learning path as a reference, the average score of those who used it would improve. From a vertical perspective, compared with those who did not use the optimal learning path as a reference for the various concept interaction degree of knowledge points, the students who used...
the optimal learning path with different concept interaction degrees of knowledge points would have improved their final average scores. Finally, we use the DBSCAN algorithm based on eigenvalue vector to cluster group A and group B, and get the change of the number of concept interaction degree of each knowledge point before and after the learning course of group A and group B, as shown in Table 7 and Table 8.

As can be seen from Table 7, Group A students took the optimal learning path when studying this course, so the number of students with primary concept interaction degree of knowledge points decreased, while the number of students with intermediate and advanced concept interaction degree of knowledge points increased. From Table 8, Group B students did not refer to the optimal learning path when studying this

| Group Category | Primary | Intermediate | Advanced |
|----------------|---------|--------------|----------|
| Before Learning| 23      | 120          | 8        |
| After Learning | 15      | 126          | 10       |

FIGURE 18. The optimal learning path for the students of 2016 with primary concept interaction degree of knowledge points.

FIGURE 19. The optimal learning path for the students of 2016 with intermediate concept interaction degree of knowledge points.

TABLE 7. Group A: the distribution of Primary, Intermediate and Advanced number of concept interaction degrees of knowledge points before and after learning the course.
focus on multi-modal education, combining the hierarchical time-series memory model and the performance indicators of deep-learning-related knowledge optimization algorithm.

In view of the fact that this paper is for group learning path planning, it does not provide personalized learning content for every individual with learning preference. In order to meet the different needs of individual learning, improve the effect of individual learning, and realize the high-level personalized education of “one person, one space”, the future work needs to study the relationship between different knowledge points, the learning traces of individuals between different knowledge points, the relationship between learning paths and learning effects, and so on, through the analysis of the behavior sequence formed by individual learners on the whole complex network. Combined with the hierarchical time series memory model and the optimization algorithm of deep learning-related knowledge, the interaction model between the structure of complex network and individual learning effect is constructed to realize the establishment and dynamic optimization of individual personalized learning path.

REFERENCES

[1] P. Prinsloo, “Big data in education. The digital future of learning, policy and practice,” Int. Stud. Sociol. Educ., vol. 29, pp. 183–186, Nov. 2020.
[2] Z. Lu, “Research on higher vocational education resources sharing based on cloud computing and big data,” in Proc. IIEASM, 2019, pp. 2231–2235.
[3] D. Andone and V. Mihaescu, “Blending MOOCs into higher education courses—A case study,” in Proc. Learns. MOOCS (LWMOOCS), Sep. 2018, pp. 134–136.
[4] J.-H. Zhang, Y.-S. Zhang, Y.-J. Jia, and Z.-K. Zhang, “The study of Internet plus continuing education pattern based on micro-learning,” in Proc. 8th Int. Conf. Inf. Technol. Med. Educ. (ITME), Dec. 2016, pp. 826–829.
[5] W. Zhang and S. Qin, “A brief analysis of the key technologies and applications of educational data mining on online learning platform,” in Proc. IEEE 3rd Int. Conf. Big Data Anal. (ICBDA), Mar. 2018, pp. 83–86.
[6] J. Liu, S. Jia, Q. Qiao, and M. Li, “Design and development of a network learning platform based on the big data concept,” in Proc. 7th Int. Conf. Educ. Innov. Through Technol. (EITFT), Dec. 2018, pp. 111–115.

[7] S. K. Ch and S. Popuri, “Impact of online education: A study on online learning platforms and edX,” in Proc. IEEE Int. Conf. MOOC, Innov. Technol. Edu.(MITE), Dec. 2013, pp. 366–370.

[8] R. Y. Toledo, Y. C. Mota, and L. Martinez, “A recommender system for programming online judges using fuzzy information modeling,” Informatics, vol. 5, no. 2, pp. 17–29, 2018.

[9] M. Attaran, J. Stark, and D. Stotler, “Opportunities and challenges for big data analytics in US higher education: A conceptual model for implementation,” Ind. Higher Educ., vol. 32, no. 3, pp. 169–182, Jun. 2018.

[10] A. Pardo, F. Han, and R. A. Ellis, “Combining university student self-regulated learning indicators and engagement with online learning events to predict academic performance,” IEEE Trans. Learn. Technol., vol. 10, no. 1, pp. 82–92, Jan. 2017.

[11] M. M. Rahman and N. A. Abdullah, “A personalized group-based recommendation approach for Web search in e-learning,” IEEE Access, vol. 6, pp. 34166–34178, 2018.

[12] S. Wang and Z. Niu, “An e-learning recommendation approach based on the self-organization of learning resource,” Knowl.-Based Syst., vol. 160, pp. 71–87, Nov. 2018.

[13] H. Chen, C. Yin, R. Li, W. Rong, Z. Xiong, and B. David, “Enhanced learning resource recommendation based on online learning style model,” Tsinghua Sci. Technol., vol. 25, no. 3, pp. 348–356, Jun. 2020.

[14] M. Wang and L. Tian, “From time series to complex networks: The phase space coarse graining,” Phys. A, Stat. Mech. Appl., vol. 461, pp. 456–468, Nov. 2016.

[15] R. S. Shaw, “The learning performance of different knowledge map construction methods and learning styles moderation for programming language learning,” J. Educ. Comput. Res., vol. 56, no. 8, pp. 1407–1429, Jan. 2019.

[16] J. Yang, J. Bai, and S. Wen, “Personalized autonomous learning path planning based on knowledge map,” in Proc. Adv. Sci. Ind. Res. Center, Int. Conf. Adv. Educ. Manage. Sci. (AEMS), 2017, pp. 230–233.

[17] H. Zhu, Y. Ni, F. Tian, P. Feng, Y. Chen, and Q. Zheng, “A group-oriented recommendation algorithm based on similarities of personal learning generative networks,” IEEE Access, vol. 6, pp. 42729–42739, 2018.

[18] H. Zhu, Y. Liu, F. Tian, Y. Ni, K. Wu, Y. Chen, and Q. Zheng, “A cross-curriculum video recommendation algorithm based on a video-associated knowledge map,” IEEE Access, vol. 6, pp. 57562–57571, 2018.

[19] H. Zhu, F. Tian, K. Wu, N. Shah, Y. Chen, Y. Ni, X. Zhang, K.-M. Chao, and Q. Zheng, “A multi-constraint learning path recommendation algorithm based on knowledge map,” Knowl.-Based Syst., vol. 143, pp. 102–114, Mar. 2018.

[20] H. Liu and X. Li, “Learning path combination recommendation based on the learning networks,” Soft Comput., vol. 24, no. 6, pp. 4427–4439, 2020.

[21] D. Shi et al., “A learning path recommendation model based on a multi-dimensional knowledge graph framework for e-learning,” Knowl.-Based Syst., vol. 18, pp. 201–215, 2020.

[22] A. P. Bharishini, S. Chandrakumarangalamp, and G. Arthi, “Ant colony optimization for competency based learning objects sequencing in e-learning,” Appl. Math. Comput., vol. 263, pp. 332–341, Jul. 2015.

[23] I. Kamsa, R. Elouahbi, F. El Khourhi, T. Karite, and H. Zouiten, “Optimizing collaborative learning path by ant’s optimization technique in e-learning system,” in Proc. 15th Int. Conf. Inform. Technol. Based Higher Educ. Training (ITHET), Sep. 2016, pp. 1–5.

[24] M. Rastegarmoghdam and K. Ziaarati, “Improved modeling of intelligent tutoring systems using ant colony optimization,” Educ. Inf. Technol., vol. 22, pp. 201–219, Mar. 2017.

[25] M. Birjali, A. Beni-Hssane, and M. Erritelli, “A novel adaptive e-learning model based on big data by using competence-based knowledge and social learner activities,” Appl. Soft Comput., vol. 69, pp. 64–92, Aug. 2018.

[26] I. Kamsa, R. Elouahbi, and F. E. Khourhi, “The combination between the individual factors and the collective experience for ultimate optimization learning path using ant colony algorithm,” Int. J. Adv. Sci., Eng. Inform. Technol., vol. 8, no. 4, pp. 1198–1208, 2018.

[27] V. Vanitha, P. Krishnan, and R. Elakkiya, “Collaborative optimization algorithm for learning path construction in e-learning,” Comput. Electr. Eng., vol. 77, pp. 325–338, Jul. 2019.

[28] W. Intayoold, T. Becker, and P. Temdee, “Social context-aware recommendation for personalized online learning,” Wireless Pers. Commun., vol. 97, no. 1, pp. 163–179, Nov. 2017.

[29] L. Ni, L. Yanshen, and L. Yi, “A recommendation system based on the education big data for e-learning,” in Proc. 7th Int. Conf. Comput. Eng. Netw. PoS (CENet), Jul. 2017, p. 91.

[30] C. Su, “Designing and developing a novel hybrid adaptive learning path recommendation system (ALPRS) for gamification mathematics geometry course,” EURASIA J. Math., Sci. Technol. Educ., vol. 13, no. 6, pp. 2275–2298, Apr. 2017.

[31] R. S. Nginda, C. M. Padron, I. Martinez-Salazar, and F. Torres-Guerrero, “Design and evaluation of planning and mathematical models for generating learning paths,” Comput. Intell., vol. 34, no. 3, pp. 821–838, Aug. 2018.

[32] X. Tang et al., “A reinforcement learning approach to personalized learning recommendation systems,” Brit. J. Math. Stat. Psychol., vol. 19, pp. 113–127, 2018.

[33] R. Han, K. Chen, and C. Tan, “Curiosity-driven recommendation strategy for adaptive learning via deep reinforcement learning,” Brit. J. Math. Stat. Psychol., vol. 10, pp. 70–81, Feb. 2020.

[34] T. Saito and Y. Watanobe, “Learning path recommendation system for programming education based on neural networks,” Int. J. Distance Educ. Technol., vol. 18, no. 1, pp. 36–64, Jan. 2020.

[35] Q. Zhang, M. Li, and Y. Deng, “Measure the structure similarity of nodes in complex networks based on relative entropy,” Phys. A, Stat. Mech. Appl., vol. 491, pp. 749–763, Feb. 2018.

[36] Z. Shou, H. Tian, S. Li, and F. Zou, “Outlier detection with enhanced angle-based outlier factor in high-dimensional data stream,” Int. J. Innov. Comput. Inf. Control, vol. 14, no. 5, pp. 1633–1651, 2018.

[37] Y. Zhang and Y. Zhou, “Review of clustering algorithms,” J. Comput. Appl., vol. 39, no. 7, pp. 1869–1882, 2019.
HUA YUAN received the B.S. degree in computer and application and the M.S. degree in electronic information engineering from the Guilin University of Electronic Technology, in 1999 and 2012, respectively.

In recent years, he presided over three department level projects, published many articles in academic journals as the first author and obtained many national invention patents. His current research interests include intelligent information processing, computer vision and image processing, machine learning, and artificial intelligence.

HUIBING ZHANG received the B.S. and M.S. degrees in computer science and technology from the Guilin University of Electronic Technology, Guilin, China, in 2000 and 2005, respectively, and the Ph.D. degree in computer science and technology from the Beijing University of Technology, Beijing, China, in 2012.

He is currently an Associate Professor with the Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology. He presided over and participated in a number of NSFC projects. His research interests include service-oriented computing, educational big data, the Internet of Things, and social computing.

JUNLI LAI received the B.S. degree in communication engineering from the Guilin University of Electronic Technology, China, in 2017, where he is currently pursuing the M.S. degree in IC engineering.

His current research interests include machine learning and big data in education.

* * *