Online-Journals.org

iJET International Journal of Emerging Technologies in Learning

iJET | eISSN: 1863-0383 | Vol. 18 No. 13 (2023) | OPEN ACCESS

https://doi.org/10.3991/ijet.v18i13.41915

PAPER

A New Methodology of Knowledge Point Sequence Generation and Learning Path Recommendation by Knowledge Reasoning

Jia Zhou, Dongling Liang(✉)

School of Water Conservancy & Civil Engineering, Northeast Agricultural University, Harbin, China

liangdongling@neau.edu.cn

ABSTRACT

Accurate knowledge point sequences and clear learning paths are the compass for learners to travel in the sea of massive knowledge, they can point out the way for learners and improve their learning experience. Generating knowledge point sequences and learning paths based on knowledge reasoning is conductive to optimizing the allocation of educational resources and improving the quality of higher education, and this has a profound influence on the reform of the entire educational field. In view of this, this study explored the generation of knowledge point sequences and the recommendation of learning paths based on knowledge reasoning. At first, the learning behavior of learners was subjected to collaborative analysis based on three aspects of knowledge points: learning frequency, learning duration, and pause/skip frequency, and the specific method of generating subject knowledge point sequences based on the metrics of difficulty differences was given. Then, a sequence sampling method that matches the features of Entity-Relationship (ER) diagram was proposed, which enables the system to dynamically adjust the recommended knowledge points and learning paths according to learners' learning progress with the help of biased random walks, thereby giving personalized and dynamic learning recommendations. At last, the validity of the proposed method was verified by experimental results.

KEYWORDS

knowledge reasoning, metrics of difficulty differences, knowledge point sequence, learning path recommendation

1 INTRODUCTION

Online learning has become a part of daily life for contemporary people in this information age [1–4], however, the explosion of online learning platforms and courses brings not only plentiful learning resources, but also the problem of

Zhou, J., Liang, D. (2023). A New Methodology of Knowledge Point Sequence Generation and Learning Path Recommendation by Knowledge Reasoning. *International Journal of Emerging Technologies in Learning (iJET)*, 18(13), pp. 178–192. https://doi.org/10.3991/ijet.v18i13.41915

Article submitted 2023-04-23. Resubmitted 2023-06-02. Final acceptance 2023-06-02. Final version published as submitted by the authors.

information overload, and learners always feel lost in the countless courses and knowledge points [5–9]. Besides, the conventional education mode is always rigescent and fixed, so it's impossible to adapt to the individual needs of learners and adjust the content according to their respective knowledge foundation [10–15]. Therefore, how to give effective learning path recommendations and generate suitable knowledge point sequences for learners has turned into an important research topic for field scholars.

Everyone learns and comprehends in a different way. If suitable learning paths can be formulated for each one according to their own situations, no doubt their learning efficiency and quality will be greatly improved. Accurate knowledge point sequences and clear learning paths are the compass for learners to travel in the sea of massive knowledge, they can point out the way for learners and improve their learning experience. Generating knowledge point sequences and learning paths based on knowledge reasoning is conductive to optimizing the allocation of educational resources and improving the quality of higher education, and this has a profound influence on the reform of the entire educational field [16, 17].

A learning path is the implementation of a course design, it consists of a set of learning activities that help users achieve specific learning goals. Due to differences in user limits, backgrounds, and goals, personalization of learning paths has become an important task, and researchers have made various efforts in the past ten years to use different techniques and approaches to propose all sorts of learning path personalization methods. For example, Nabizadeh et al. [18] reviewed various learning path personalization methods and explained their main parameters, and then proposed an approach to assess these methods. The authors pointed out that the most challenging thing of these methods is to improve the quality of personalization. Bian et al. [19] said that adaptive learning has attracted the attention of researchers as it can automatically recommend learning resources to learners and allow them to enjoy a personalized learning experience. Actually there are a few ways to achieve that, and one realistic approach is to recommend adaptive learning paths, namely to provide learning resources according to the needs of learners. The authors summarized existing research outcomes, and proposed their own method. At first, a learner-centred concept map was created using graph theory based on the features and concepts of learners; then the approach was adopted to generate a linear concept sequence from the concept map using a proposed traversal algorithm; after that, learning objects, which are the smallest concrete units making up a learning path, were organized based on concept sequences.

After carefully reviewing existing literatures, it's found that most available learning path recommendation systems were built based on static learning models, they generally failed to notice that learning is actually a dynamic process, and learners' knowledge mastery situation and interest point may change over time, as a result, such dynamic features can not be effectively solved by existing models. Moreover, current knowledge reasoning methods tend to express knowledge points as a simple knowledge graph, which is unable to exhibit the complexity of knowledge. For instance, knowledge graph could not capture the subtle relationships between knowledge points, or fully show the depth and breadth of knowledge, thus it is still a primary task to research more dynamic, complete, personalized knowledge expression methods, solve cold start problem, and improve the methods for generating interpretable knowledge point sequences and recommending learning paths. Out of these concerns, this study explored the generation of knowledge point sequences and the recommendation of learning paths based on knowledge reasoning.

2 GENERATION OF KNOWLEDGE POINT SEQUENCES BASED ON DIFFICULTY DIFFERENCE METRICS

Data of learning frequency, learning duration, and pause/skip frequency of knowledge points can reflect learners' learning behavior and state from multiple angles. Learning frequency reflects a learner's learning enthusiasm and learning habits; learning duration reflects learning patience and concentration; and the learning pause/skip frequency reflects learners' learning difficulty and the level of their understandings. So the collaborative analysis of learners' learning behavior based on data of these three aspects can be used for a preliminary understanding of learners' learning behavior and process.

Assuming: u represents each knowledge point contained in courses learnt by learner i, $1 \le u \le A$; A represents the collection of course knowledge points; then the frequency $FR(i,u)$ of learner i learning knowledge point u can be calculated by the following formula:

$$c_d(i,u) = \frac{FR(i,u)}{\underset{1 \le u \le A}{MA}\, FR(i,u)} \tag{1}$$

Assuming: $Y(i,u)$ represents the learning duration of learner i for knowledge point u; Y represents the original time it takes to learn knowledge point u; then a learner's learning duration for knowledge point u can be expressed as:

$$c_y(i,u) = \begin{cases} \dfrac{Y(i,u)}{3*Y}, \dfrac{Y(i,u)}{Y} \le 3 \\ U, \dfrac{Y(i,u)}{Y} > 3 \end{cases} \tag{2}$$

Assuming: $PA(i,u)$ represents the frequency of learner i to pause knowledge point u; $DR(i,u)$ represents the frequency of learner i to skip knowledge point u; then the learner's pause/skip frequency for knowledge point u can be expressed as:

$$c_{o,f}(i,u) = \frac{PA(i,u) + DR(i,u)}{\underset{1 \le u \le A}{MA}(PA(i,u) + DR(i,u))} \tag{3}$$

According to the learning behavior features of learners in these three aspects, the relative difficulty coefficient of knowledge point u for learner i can be defined as:

$$djo_{i,u} = \beta \times c_d(i,u) + \alpha \times c_y(i,u) + \varepsilon \times c_{o,f}(i,u) \tag{4}$$

Based on the coefficient values of all knowledge points attained from above formula, a relative difficulty coefficient matrix of knowledge points of learner i can be constructed.

The learning difficulty difference metrics of knowledge points can help researchers quantify the learning difficulty of different knowledge points for different learners and further reveal the differences in the learning difficulty of knowledge points, also, it offers valuable evidences for teachers to understand learners' learning state so that they could adjust the teaching strategies in a targeted manner (see Figure 1). Therefore, the relative difficulty coefficients of knowledge points for learners were taken as an important factor in this study when creating knowledge point sequences, that is, knowledge points with lower relative difficulty coefficients could be arranged first, and then gradually transitioning to those with higher relative difficulty coefficients, in this way, learners' comprehension and mastery of knowledge could be built step by step.

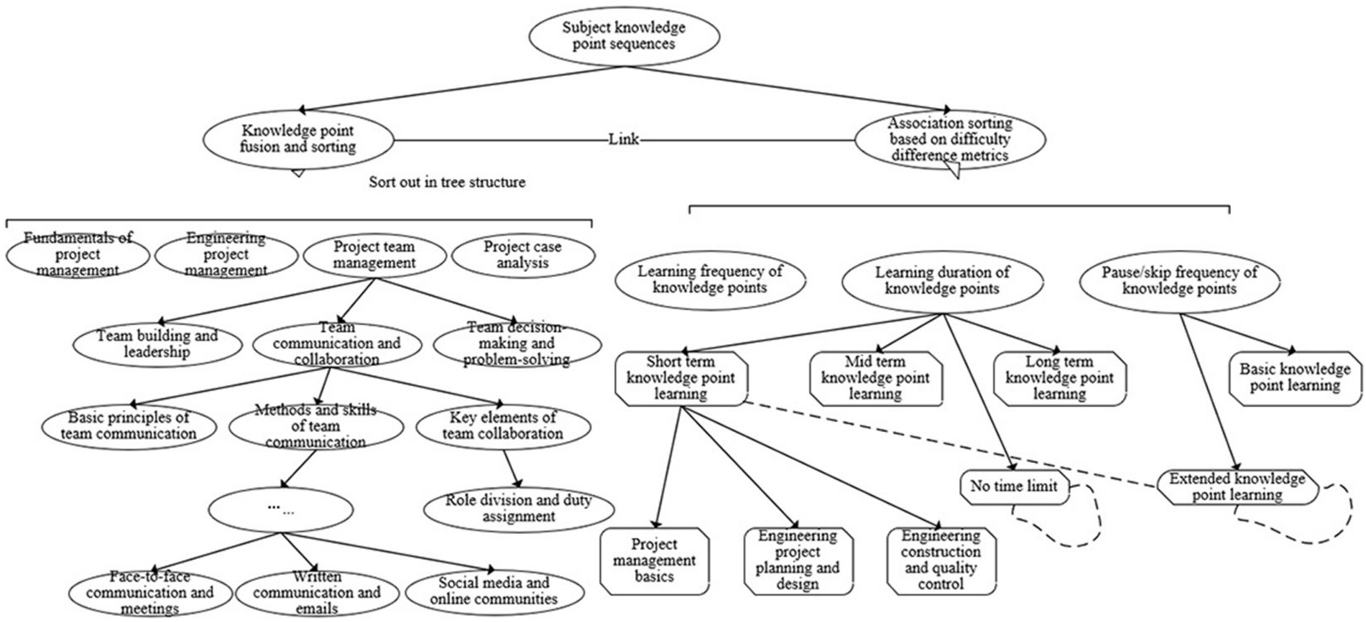

Fig. 1. Principle of creating knowledge point sequences based on difficulty difference metrics

Every learner differs in learning ability, learning method, and knowledge mastery situation, so inevitably there are differences in their learning difficulty with respect to a same knowledge point. By measuring such differences, personalized knowledge point sequences could be created and recommended to each learner based on their respective learning needs and ability levels, basic knowledge points with suitable difficulty level could be pushed to learners first, so that the learners could follow a learning path from easy to difficult, form a solid knowledge foundation, and study in a systemic way.

When measuring differences in learners' comprehension and mastery of a same knowledge point, the relative difficulty coefficients of knowledge points are an important metric. Because different learners may have different cognition in the difficulty of a same knowledge point, only using relative difficulty coefficients to represent such differences is not enough, thus it's necessary to use average relative difficulty coefficients of knowledge points to measure learners' differences based on the relative difficulty coefficients. Average relative difficulty coefficients can eliminate learners' individual differences in terms of the cognition of difficulty, making the evaluation more objective and not biased.

Assuming: $|djo_{i,u} - djo_{c,u}|$ represents the difference of different learners in a same knowledge point; $|djo^*_{i,u} - djo^*_{c,u}|$ represents the difference of different learners in the average relative difficulty coefficient of the knowledge point; $djo_{i,u}$ represents the relative difficulty coefficient of knowledge point u for learner i; $djo_{c,u}$ represents the relative difficulty coefficient of knowledge point u for target learner c; V represents knowledge points learnt by different learners; MAX and MIN represent the maximum and minimum relative difficulty coefficient of knowledge points in the knowledge point relative difficulty coefficient matrix; $djo^*_{i,u}$ and $djo^*_{c,u}$ represent the average relative difficulty coefficient of knowledge points learnt by learner i and target learner c; then the difference between learners $\rho(u,v)$ can be attained from the following formula:

$$\rho(i,c) = 1 - \frac{\sqrt{\sum_{u=1}^{V} \left|djo_{i,u} - djo_{c,u}\right| \cdot \left|djo^*_{i,u} - djo^*_{c,u}\right|}}{MAX - MIN} \tag{5}$$

Knowledge points learned by an unrelated learner may reveal some characteristics of the educational environment in which the learner is located, such as educational resources, teaching methods and evaluation systems, and these can affect the learner's comprehension and mastery of knowledge. If the learning difficulty of some knowledge points is found to be relatively high among unrelated learners, then it may indicate some problems with the education, such as defects with teaching methods, textbooks, or educational resources. Considering the learning state of unrelated learners facilitates to establishing a fairer metric, because it can eliminate the influence of educational environment on the learning difficulty of knowledge points for learners, so that the differences of learners in learning difficulty can be compared in a fairer way. Thus, considering the influence of knowledge points learned by unrelated learners on the difference measurement of learning difficulty between learners is of great significance for a more comprehensive, fairer and deeper understanding of learners' learning state, so that reasonable and useful evidences could be provided for teaching reform and improvement.

To consider the above-mentioned influence, it's assumed V represents knowledge points learnt by different learners; J_i represents knowledge points learnt by learner i; J_c represents knowledge points learnt by learner c, then the definition of ϕ is:

$$\varphi = \frac{V}{\left|J_i\right| \cdot \left|J_c\right|} \tag{6}$$

To solve the limitations of classical collaborative filtering recommendation algorithms in the difference metrics of learning difficulty of knowledge points, here ϕ and $\rho(i,c)$ were combined, the model for measuring learners' difference in the learning difficulty of knowledge points can be expressed as:

$$SIM(i,c) = \rho(i,c) \cdot \varphi \tag{7}$$

Calculations of above formula gave a nearest neighbor set that is the most similar to the target learner. Assuming: $djo'_{c,k}$ represents the predicted relative difficulty coefficient of knowledge point k for target learner c; $djo_{i,k}$ represents the relative difficulty coefficient of knowledge point k for learner i; djo^*_i and djo^*_c represent the average relative difficulty coefficient of knowledge points learnt by learner i and target learner c; b represents the set of nearest neighbors (learners); $SIM(i,c)$ represents the similarity between learner i and target learner c; then the relative difficulty coefficient of knowledge points for target learner *c* can be predicted based on the following formula:

$$djo'_{c,k} = djo_c^* + \frac{\sum_{i=1}^{b}(djo_{i,k} - djo_i^*) \times SIM(i,c)}{\sum_{i=1}^{b} SIM(i,c)} \tag{8}$$

3 SEQUENCING OF ASSOCIATED KNOWLEDGE POINTS BASED ON ER DIAGRAM

A weighted ER diagram can more accurately indicate the degree of association between knowledge points and give a thorough understanding of the knowledge structure. This method indicates not only the direct associations between knowledge points, but also their indirect associations, then the understanding of knowledge structure could be deepened. According to the similarities and differences of learners' learning difficulty, the intensity of associations between knowledge points was quantified to better recommend personalized learning paths for each learner. According to the learning difficulty of learners and the association intensity of knowledge points, knowledge learning paths that meet the individual ability levels and needs of each learner could be formulated. When recommending knowledge points or creating knowledge point sequences, the knowledge points can be sorted according to the association intensity between knowledge points, in this way, more accurate recommendations could be given.

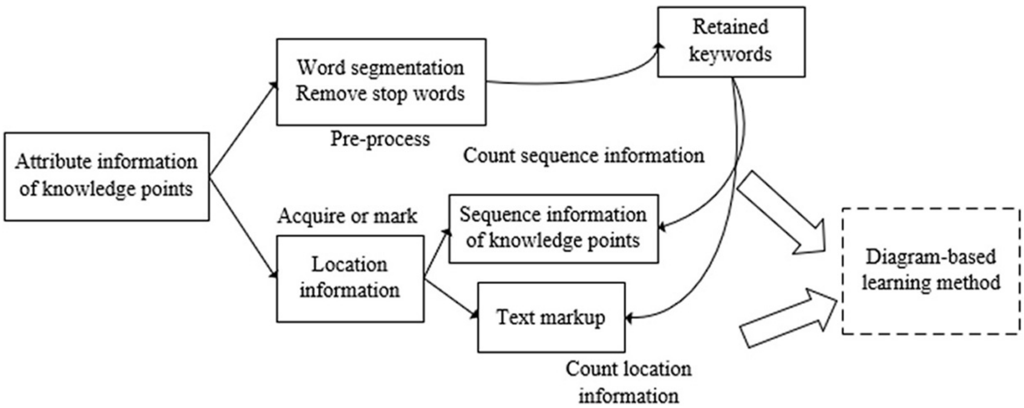

Fig. 2. Process of the diagram-based learning method

By investigating the ER diagram, the relationship between knowledge points can be fully understood (Figure 2). Not only the direct associations, also the indirection associations could be figured out clearly, thus a deeper understanding of the knowledge structure could be attained. This study proposed a sequence sampling method that matches with the features of ER diagram. The mechanism of biased random walks can make the system adjust recommended knowledge points and learning paths dynamically according to learners' learning process, so as to give dynamic and personalized learning recommendations.

Processing complex relational networks in high dimensional spaces can consume significant computational resources and time. Through diagram-based learning, the knowledge points can be mapped to a lower dimensional vector space, then the computational complexity can be reduced effectively, the processing speed could be improved, and the hidden associations between knowledge points can be captured to a certain extent via this mapping. Actually, these hidden associations may be difficult to find directly from the original ER diagram, so this study adopted the diagram-based learning method to map all knowledge points into a lower dimensional vector space, the method can reduce complexity, capture hidden associations, compare similarities flexibly, sort knowledge points, and facilitate knowledge transfer and promotion when dealing with large-scale knowledge networks.

For known $H_R = \{C,R,Q\}$, $c_u \in C$ was subjected to low-dimension vector mapping based on knowledge point sequence sampling of random walks and the *skip-gram*, assuming d represents the goal of diagram-based learning, G represents the low-dimension vector space, f represents the dimension of knowledge point vector, then the mapping process is given by the following formula:

$$d : C \to G^{|C| \times f} \tag{9}$$

In order to improve the ability of the diagram-based learning algorithm to describe associations in the ER diagram, for the features of the ER diagram, this study proposed a random walk mechanism of extended queue to modify the sampling strategy of the model.

Under the condition that $H_R = \{C,R,Q\}$ is known, for all $c_u \in C$, if there the mean of $c_u c_k \in R$, then c_k is called a neighbor knowledge point of c_u, and the set of all neighbor knowledge points of c_u is represented by JD_u; $QU_u = \{c_u(1), c_u(2), \ldots c_u(|QUy|)\}$ represents an extended queue, its length is $|QUu|$; $c_u(k)$ is the k-th entity sorted from high to low according to the difference measurement method of knowledge point learning difficulty given in the previous section, and $c_u(k) \in C - c_u$, wherein the association degree between $c_u(1)$ and c_u is the highest. Every $c_u(k) \in QU_u$ was added into BR_u. In the meantime, μ_{uk} represents the association score of c_u and c_k attained by the difference measurement method of knowledge point learning difficulty. In this way, the hypergraph H_{RA} of ER diagram with extended queues can be defined. $H_{RA} = \{C,R,Q,W\}$ is a superset of $H_R = \{C,R,Q\}$, wherein W represents the set of knowledge point extended queues, $QU_u \in W$.

For the b knowledge point sequences (length is $qm-b$) to be recommended attained from a single-time sampling, since the random walk sampling of extended queues needs to satisfy relevant properties of the Markov chain, initialization and the acquisition of original sample sequences can be completed within $I(qm/b(qm-b))$, and execution of the created extended sample sequences does not exceed $P(b \times qm^2)$, so the overall time complexity of the sampling algorithm is $P(b \times qm^2)$.

Under above mentioned sampling mechanism, the processes of sequencing associated knowledge points and creating learning paths based on the diagram-based learning method are given below.

Assuming: $c_a \in C$ is the source knowledge point; v_m represents the m-th step of the random walk, and $v_0 = c_a$; $\lambda_{lb}(v_{m-1}, v_{m+1})$ represents the hyper-parameter of the shortest path between v_{m-1} and v_{m+1}; l and b represent self-defined bias parameters for controlling the random walk; then the transition probability of the random walk from the m-th step to the m+1-th step is related to the correlation coefficient *vps*:

$$\tau(v_m, v_{m+1}) = \lambda_{ll}(v_{m-1}, v_{m+1}) \times vps(v_m, v_{m+1}) \tag{10}$$

Assuming: x represents the normalization parameter; ς represents the scaling coefficient; av represents the threshold of switching probability; further, based on above sampling mechanism, the actual transition probability j_y from c_u to c_k can be calculated by the following formula:

$$j_y = (v_m = c_k \mid v_{m-1} = c_u) = \begin{cases} (1-av)\times\dfrac{\varsigma\times\mu_{uk}\times\tau(c_u,c_k)}{x}; eo\le av \\ av\times\dfrac{\tau(c_u,c_k)}{x}; others \end{cases} \quad (11)$$

Assuming: $B_n(c_u)\subset C$ represents the neighbor network structure generated based on the above sampling mechanism, to achieve $d{:}c_u\rightarrow E^f$, $c_u\in C$ needs to be mapped to a same vector space based on the following formula:

$$\sum\nolimits_{c_u\in C} LO[o(B_n(c_u)\mid d(c_u))] \quad (12)$$

Assuming: $G^{|C|\times f}$ represents the hidden matrix space of the ER diagram attained from learning, wherein the u-th row represents the f-dimensional vector of c_u; $G_{w\rightarrow}$ represents the vector corresponding to the queried knowledge point $w\rightarrow$; G_{uk} represents the k-th value in the vector shown by the u-th row of $G^{|C|\times f}$; then the association intensity $sf(v_i,q^{\rightarrow})$ between the queried knowledge point $w\rightarrow$ and the candidate knowledge point c_u can be attained from the following formula:

$$sf(c_u,\vec{w}) = \frac{\sum\nolimits_{k=1}^{f} G_{uk}\times G_{\vec{w}}}{\sqrt{\sum\nolimits_{k=1}^{f} G_{uk}^2 \times \sum\nolimits_{k=1}^{f} G_{\vec{w}}^2}} \quad (13)$$

At this point, based on $sf(c_u,w^{\rightarrow})$, *top-k* candidate knowledge points that are most significantly associated with $w^{\rightarrow}$ can be found, forming a list of associated knowledge points $M(w^{\rightarrow})=\{sf(c_1,w^{\rightarrow}),\ldots,sf(c_j,w^{\rightarrow})\}$. Then $M(w^{\rightarrow})$ was sorted from high to low according to the association intensity, and then output as the recommended learning path.

4 EXPERIMENTAL RESULTS AND ANALYSIS

Figure 3 shows how the number of created knowledge point sequences varies with the data volume of knowledge library. As can be seen from the figure, as the data volume increases, the number of created knowledge point sequences increases as well. This is because a greater data volume can give more combinations of knowledge points, so the number of created knowledge point sequences will increase accordingly. Judging from the curve trend in the figure, the growth of the number of created knowledge point sequences is not linear, when the knowledge library only has a small data volume, the number of created knowledge point sequences grows quickly, however, as the number reaches a certain value, the growth speed begins to slow down; when the data volume reaches a certain value, the growth of number of knowledge point combinations stops, and maintains at a stable value. These observations and conclusions reveal the influence of knowledge library data volume on the created knowledge point sequences, thereby providing the possibility of recommending effective learning paths in case of large-scale knowledge libraries.

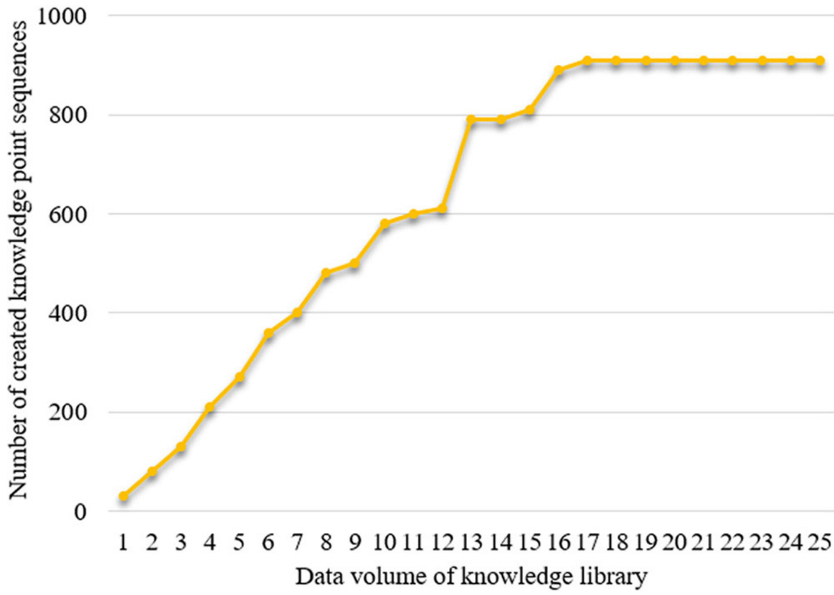

Fig. 3. Trend of the number of created knowledge point sequences

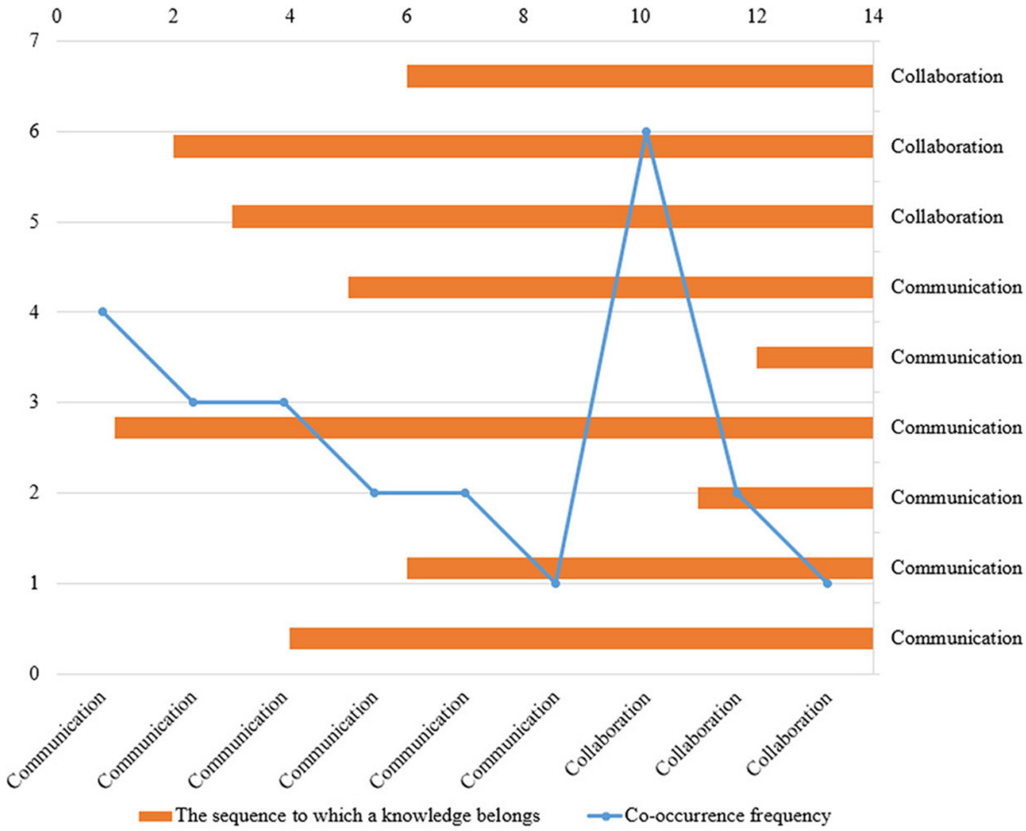

Fig. 4. Co-occurrence information of knowledge point locations

Figure 4 lists the occurrence frequency of knowledge point keywords "communication" and "collaboration" in different sequences. According to these information, it's known that the co-occurrence frequency distributes evenly in each sequence, and it's higher in sequences 4, 6, and 11. This indicates that in these knowledge point sequences, "communication" is an important topic or a concept that is frequently discussed. The frequency of "collaboration" is the highest in Sequence 3, it appeared 6 times, indicating that in Sequence 3, the knowledge point of "collaboration" is concentrated, it is a key knowledge point or a key topic. In other sequences, the co-occurrence frequency is significantly lower. Overall speaking, the co-occurrence frequency of "communication" is higher than that of "collaboration", indicating that in the whole knowledge point combinations, "communication" is more critical or complicated, and more knowledge points and content are required to explain and discuss it. The above data also facilitates the understanding and prediction of learners' learning paths and learning needs, so that more effective learning paths could be recommended. The method proposed in this study can make special recommendations of knowledge points with higher frequency in specific sequences, so as to help learners better understand and master these important topics or concepts.

Table 1. Generation information of some knowledge point sequences

Number of Created Knowledge Point Sequences	Number of Similar Knowledge Points	Total Number of Knowledge Points
1	2	61
2	5	224
3	22	251
4	8	354
5	12	388
6	31	423
7	15	568
8	8	603

Table 1 lists the number of created knowledge point sequences, the number of similar knowledge points that are related, and the total number of knowledge points. As can be seen from the table, with the increase of the number of created knowledge point sequences, the total number of knowledge points grows as well, which has reflected the complexity of knowledge point sequence generation, and this complexity comes from the complexity of the curriculum, the diversity of learners' learning needs and ability levels, and the correlation between knowledge points. The number of similar knowledge points can reflect the associations between knowledge points. If two knowledge points are generated in the same knowledge point sequence, then there must be an association between them. The increase in the number of similar knowledge points indicates an enhancement in the associations between knowledge points. The information given in the table exhibits the diversity of knowledge point sequence generation, and this diversity comes from the diversity of knowledge points, the associations between knowledge points, and the diversity of learners' learning needs and abilities.

Table 2. Comparison of experimental results of recommended learning paths

Methodology	Accuracy (*p*)	Recall Rate (*R*)	F1-value (*F*1)
Content-based recommendation method	67.45%	82.58%	73.89%
Rule-based recommendation method	75.31%	78.61%	77.64%
Matrix decomposition-based recommendation method	76.62%	81.51%	79.64%
The proposed method	78.42%	85.84%	81.64%

Next, the experimental results of learning paths recommended by different methods were comparatively analyzed. Table 2 gives the experimental results of four different learning path recommendation methods, including the content-based recommendation method, the rule-based recommendation method, the matrix decomposition-based recommendation method, and the proposed method. In terms of accuracy, the proposed method (78.42%) performed the best, indicating that under the condition of correct predictions, the accuracy of the proposed method is the highest, and the accuracy of content-based recommendation method is the lowest. In terms of recall rate, the performance of the proposed method is also the best, indicating that in all learning paths that should be recommended, the proposed method can find the most correct recommendations, while the recall rate of the rule-based recommendation method is the lowest (78.61%). When accuracy and F1-value are taken into consideration comprehensively, the proposed method still performed the best, indicating that it has found a best balance between accuracy and coverage. Figure 5 compares the experimental indicators of different recommendation methods.

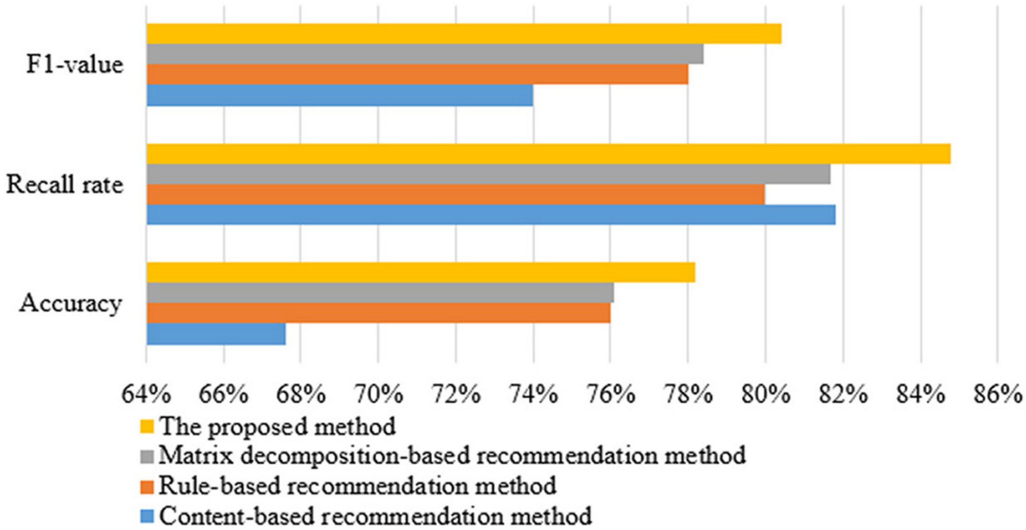

Fig. 5. Comparison of experimental indicators of different recommendation methods

Figure 5 clearly shows that the proposed method outperformed three reference recommendation methods in terms of all three evaluation indicators, and these results further verified the previous discussions, namely the advantages of the proposed method in the collaborative analysis of learners' learning behaviors (learning frequency, duration, and pause/skip frequency), the knowledge point learning difficulty metrics of learners, and the sequence sampling based on weighted ER diagram, and these have demonstrated the superiority of the proposed method in giving accurate and personalized learning path recommendations.

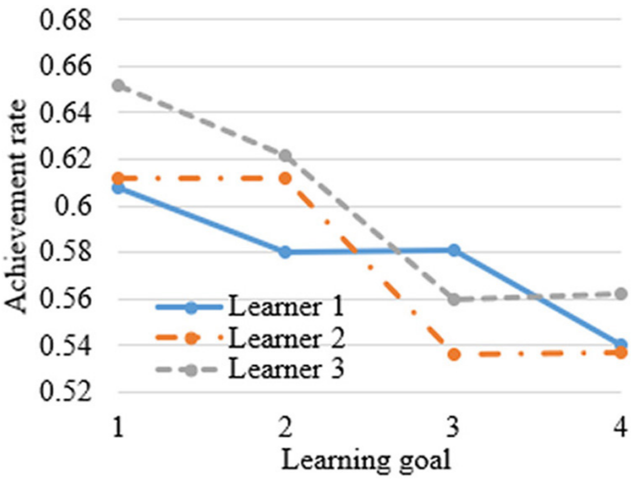

Fig. 6. Achievement rate of short-term learning goals of different learners

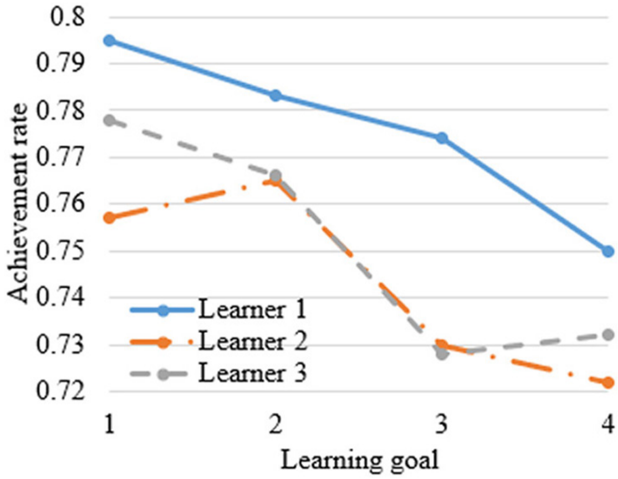

Fig. 7. Achievement rate of mid-term learning goals of different learners

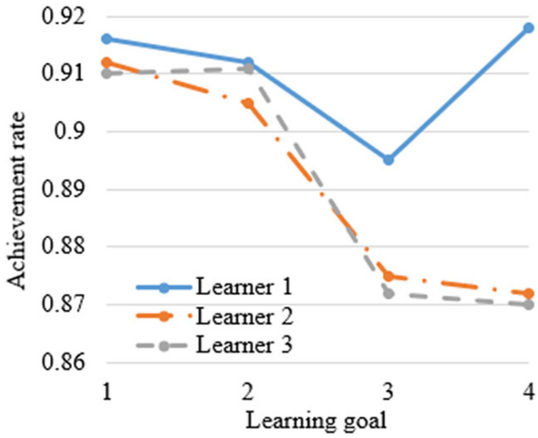

Fig. 8. Achievement rate of long-term learning goals of different learners

Figures 6, 7 and 8 respective give the short-term, mid-term, and long-term learning goal achievement rate of knowledge points of different learners after the recommended learning paths had been implemented. According to Figure 6, the achievement rate of Learner 1 in all learning goals has exceeded 0.54, and the highest achievement rate 0.608 is reached under Learning Goal 1 (the score is above 60 points), indicating that after this learner has implemented the recommended learning path, the learning effect is significant, especially under a low learning goal. In terms of Learning Goals 1 and 2 (score is above 60 and 70), the achievement rate of Learner 2 is the same, both 0.612, in terms of Learning Goals 3 and 4 (score is above 80 and 90), the achievement rate is also the same, both 0.537, indicating that the learning effect of Learner 2 in terms of all learning goals varies little. As for Learner 3, the achievement rate of all learning goals has exceeded 0.56, and the highest achievement rate 0.652 is achieved under Learning Goal 1, indicating that the learning effect of Learner 3 is the best among the three learners. Considering comprehensively, the achievement rate of short-term learning goals varies from person to person, but on the whole, learning path recommendation is helpful to improving the learning effect of learners.

As can be seen from Figure 7, under all learning goals, the achievement rate of Learner 1 has exceeded 0.75, and the highest achievement rate 0.795 is reached under Learning Goal 1, indicating that the learning effect of Learner 1 is very good, especially in the early learning stage. As for Learner 2, the achievement rate of all learning goals has exceeded 0.72, and the highest achievement rate 0.765 is reached under Learning Goal 2, indicating that Learner 2 has made steady progresses during the learning process, but has some difficulties when attempting to achieve higher learning goals. Learner 3 has reached an achievement rate higher than 0.728 under all learning goals, and the highest achievement rate 0.778 is reached under Learning Goal 1, indicating that the learner performed stably through out the entire learning process, the performance is good regardless of low or high learning goals. Overall speaking, for all learners, the achievement rate of all learning goals has exceeded 0.72, indicating a good effect of the recommended learning paths, and this is conductive to improving learners' learning effect.

According to Figure 8, for Learner 1, the achievement rate of all learning goals has exceeded 0.895, indicating that Learner 1 performed excellently in long-term learning goals, especially under Learning Goal 4, the achievement rate has reached 0.918, which is the highest of all data. Learner 2 has reached an achievement rate higher than 0.872 under all learning goals, but the achievement rate of Learning Goals 3 and 4 is slightly lower, which is 0.875 and 0.872, respectively, indicating that the learner has some difficulties in long-term learning, especially under higher learning goals. As for Learner 3, the achievement rate of all learning goals has exceeded 0.87, but the achievement rate of Learning Goals 3 and 4 is slightly lower, which is 0.872 and 0.87, respectively, this is similar to the situation of Learner 2, indicating that Learner 3 also has some difficulties in long-term learning. On the whole, after all learners have implemented the recommended learning paths, the achievement rate of long-term knowledge point learning has exceeded 0.87, indicating that the recommendation method has a significant effect on improving learners' long-term learning effect.

5 CONCLUSION

The collaborative analysis conducted in this study based on learning behavior can more accurately measure students' relative learning difficulty coefficients of knowledge points, which can facilitate accurate generation of learning paths and recommendation of learning resources, and this has important value for personalized teaching and improving students' learning effect. By creating weighted ER diagrams, the association intensity between knowledge points was better understood and measured. Further, a sequence sampling method was proposed based on the mechanism of biased random walks, with the help of this method, knowledge points that are highly related to a given knowledge point were found with high efficiency, and more effective learning paths were created. Through the learning of ER diagram, all knowledge points were mapped into a lower dimensional vector space, the association intensity between any knowledge points was effectively measured, and the final sequencing was realized. All these contribute to generating learning paths that suit students' learning needs better.

The experimental results of learning path recommendation reveal that, compared with three references methods: content-based recommendation method, rule-based recommendation method, and matrix decomposition-based recommendation method, the proposed method performed better in terms of accuracy, recall rate and F1-value, and these have further verified the effectiveness of the proposed method. According to the analysis of different learners' achievement rate in terms of short-term, mid-term, and long-term learning goals, the learning paths recommended by the proposed method can effectively improve learners' learning effect, especially under long-term learning goals. However, for some learners who have difficulties in achieving higher learning goals, more personalized support is needed.

6 REFERENCES

[1] Nguyen, T.V., Sobia, I., Alhamzah, F.A., Huma, S., Nohman, K. (2022). Online learning—Two side arguments related to mental health. International Journal of Online and Biomedical Engineering, 18(9): 131–143. https://doi.org/10.3991/ijoe.v18i09.32317

[2] Divayana, D.G.H. (2020). Utilization of DIVAYANA formula in evaluating of suitable platforms for online learning in the social distancing. International Journal of Interactive Mobile Technologies (iJIM), 14(20), pp. 50–75. https://doi.org/10.3991/ijim.v14i20.15787

[3] Gormaz-Lobos, D., Galarce-Miranda, C., Kersten, S., Hortsch, H. (2022). Attitudes and perceptions of teaching staff about the online learning during the COVID19 pandemic: A case study of engineering education. International Journal of Engineering Pedagogy, 12(3): 38–49. https://doi.org/10.3991/ijep.v12i3.29947

[4] Yang, P., Liu, X. (2022). Evaluation of comprehensive services of an online learning platform based on Artificial Intelligence. International Journal of Emerging Technologies in Learning, 17(13): 130–144. https://doi.org/10.3991/ijet.v17i13.32797

[5] Cao, F.Y., Feng, Y.T., Wei, B. (2022). Practical research of online teaching platform on reform of computer network with flipped classroom. Review of Computer Engineering Studies, 9(2): 67–70. https://doi.org/10.18280/rces.090204

[6] Li, T., Su, X., Liu, W., et al. (2022). Memory-augmented meta-learning on meta-path for fast adaptation cold-start recommendation. Connection Science, 34(1): 301–318. https://doi.org/10.1080/09540091.2021.1996537

[7] Chenoori, R.K., Kavuri, R. (2022). Online transaction fraud detection using efficient dimensionality reduction and machine learning techniques. Revue d'Intelligence Artificielle, 36(4): 621–628. https://doi.org/10.18280/ria.360415

[8] Yue, Y., Dai, H., Zhang, Y.P., Shang, X.Q., Li, Z.H. (2021). State-of-the-art survey on personalized learning path recommendation. Journal of Software, 33(12): 4590–4615. http://dx.doi.org/10.13328/j.cnki.jos.006518

[9] Kahate, S.A., Raut, A.D. (2022). Comprehensive analysis of privacy attacks in online social network: Security issues and challenges. International Journal of Safety and Security Engineering, 12(4): 507–518. https://doi.org/10.18280/ijsse.120412

[10] Yang, H., Chen, H., Li, L., Philip, S.Y., Xu, G. (2021). Hyper meta-path contrastive learning for multi-behavior recommendation. In 2021 IEEE International Conference on Data Mining (ICDM), Auckland, New Zealand, pp. 787–796. https://doi.org/10.1109/ICDM51629.2021.00090

[11] Nguyen, T.T.M., Tran, T.P.Q. (2021). A knowledge graph embedding based approach for learning path recommendation for career goals. In Computational Collective Intelligence: 13th International Conference, ICCCI 2021, Rhodes, Greece, pp. 66–78. https://doi.org/10.1007/978-3-030-88081-1_6

[12] Tang, C.L., Liao, J., Wang, H.C., Sung, C.Y., Lin, W.C. (2021). Conceptguide: Supporting online video learning with concept map-based recommendation of learning path. In Proceedings of the Web Conference 2021, Ljubljana, Slovenia, pp. 2757–2768. https://doi.org/10.1145/3442381.3449808

[13] Saito, T., Watanobe, Y. (2020). Learning path recommendation system for programming education based on neural networks. International Journal of Distance Education Technologies, 18(1): 36–64. https://doi.org/10.4018/IJDET.2020010103

[14] Liu, H., Li, X. (2020). Learning path combination recommendation based on the learning networks. Soft Computing, 24(6): 4427–4439. https://doi.org/10.1007/s00500-019-04205-x

[15] Shi, D., Wang, T., Xing, H., Xu, H. (2020). A learning path recommendation model based on a multidimensional knowledge graph framework for e-learning. Knowledge-Based Systems, 195: 105618. https://doi.org/10.1016/j.knosys.2020.105618

[16] Xie, S., Zhang, Q., Jiang, L. (2022). Current knowledge on exosome biogenesis, cargo-sorting mechanism and therapeutic implications. Membranes, 12(5): 498. https://doi.org/10.3390/membranes12050498

[17] Li, S., Zhang, Y., Liu, J., Cui, X., Zhang, Y. (2021). Recommendation model based on public neighbor sorting and sampling of knowledge graph. Journal of Electronics & Information Technology, 43(12): 3522–3529. http://dx.doi.org/10.11999/JEIT200735

[18] Nabizadeh, A.H., Leal, J.P., Rafsanjani, H.N., Shah, R.R. (2020). Learning path personalization and recommendation methods: A survey of the state-of-the-art. Expert Systems with Applications, 159: 113596. https://doi.org/10.1016/j.eswa.2020.113596

[19] Bian, C.L., Wang, D.L., Liu, S.Y., Lu, W.G., Dong, J.Y. (2019). Adaptive learning path recommendation based on graph theory and an improved immune algorithm. KSII Transactions on Internet and Information Systems, 13(5): 2277–2298. https://doi.org/10.3837/tiis.2019.05.003

7 AUTHORS

Jia Zhou, graduated from Harbin Institute of Technology with a Ph.D. in Management and now works in Northeast Agricultural University. The main research directions are engineering project management, construction engineering economics, real estate development and operation management. Email: zhoujia@neau.edu.cn

Dongling Liang, graduated from the Northeast Forestry University with a Doctor of Engineering, works in Northeast Agricultural University. The main research direction is engineering project risk management and economic evaluation. Email: liangdongling@neau.edu.cn