Journal of Network Intelligence ©2025 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **10**, Number **3**, August **2025**

Knowledge Point Recommendation Algorithm Based on Enhanced Correction Factors and Weighted Sequential Pattern Mining

Yu-Li Li[1,*], Kun Yao[2], Ying Zhang[3]

[1]School of Software Engineering, Institute of Guangzhou, Guangzhou 510900, P. R. China
13826171564@163.com

[2]Education Information and Equipment Service Centre of Guangzhou Tianhe District, Guangzhou 510651, P. R. China
yao_200307@163.com

[3]Faculty of Engineering and Information Technology, INTI International University, 71950 Jurai Newtown, Negeri Sembilan, Malaysia
eh7451@163.com

[*]Corresponding author: Yu-Li Li

Received July 8, 2024, revised November 23, 2024, accepted February 18, 2025.

Abstract. *The application of knowledge point recommendation systems in personalised learning in modern education has important research value and necessity. However, traditional recommendation methods often fail to fully consider factors such as differences in user ratings, item popularity, and changes in user interest over time, resulting in limited accuracy and personalisation of the recommendation results. To solve these problems, this study proposes a knowledge point recommendation algorithm based on Enhanced Correction Factors (ECF) and Weighted Sequential Pattern Mining (WSPM). Firstly, the ECF is used to accurately estimate the similarity between users by combining the rating value difference correction factor, item popularity difference correction factor and time decay factor, which improves the accuracy of similarity calculation. Secondly, the WSPM algorithm is used to mine high-weighted sequence patterns from the user's learning behaviour sequences, identify important patterns reflecting the user's learning behaviour, and generate a personalized list of knowledge point recommendations by combining the user's similarity. Experimental results on three public data sets, namely, ASSIST2009, ASSIST2015 and EdNet KT1, show that compared with traditional collaborative filtering methods, the ECF-WSPM-based recommendation algorithm has significant improvement in recommendation accuracy. The proposed method provides an effective tool for educators and students to recommend knowledge points based on students' personalised needs, which can enhance the learning effect and learning experience.*
Keywords: Knowledge point recommendation; Collaborative filtering; Correction factor; Personalised learning; Weighted sequence model

1. **Introduction.** In the field of modern education, the research of knowledge point recommendation system is of great value and necessity. With the wide application of online education platforms and digital learning resources, students can access various learning resources at any time and any place [1, 2, 3]. However, the massiveness and complexity of learning resources also bring the problem of information overload. When facing numerous learning resources, students often find it difficult to choose the knowledge that best suits their current learning stage and needs. Therefore, it is of great practical significance to develop a system that can accurately recommend relevant knowledge points according to students' individual needs and learning behaviours [4, 5]. An effective knowledge

recommendation system can not only help students learn in a more targeted way and improve their learning efficiency, but also stimulate their interest in learning and enhance their learning experience, thus promoting the improvement of the quality of education in general.

Collaborative filtering algorithms [6, 7] are widely used in various recommender systems as they are able to better capture students' common learning behaviours and preferences by analysing the similarities between users to make recommendations. However, traditional collaborative filtering algorithms have some limitations in dealing with rating sparsity and cold-start problems. Sequential pattern mining techniques [8, 9, 10], on the other hand, are able to discover frequently occurring patterns from the sequence of students' learning behaviours and reflect the importance of the knowledge points by assigning different weights to them, thus providing more personalized and accurate recommendations. By combining these two techniques, higher recommendation accuracy and personalised services can be achieved in knowledge point recommendation systems. Therefore, collaborative filtering recommendation algorithms and sequential pattern mining techniques have a broad application potential in knowledge point recommendation systems.

The aim of this paper is to propose a new knowledge point recommendation algorithm, by combining user rating data and behavioural sequence data, calculating the similarity between users and mining their learning behavioural patterns, in order to achieve personalised knowledge point recommendation, thus contributing new technical means to the development of online education.

1.1. **Related work.** In the field of education, research on knowledge point recommendation systems has made significant progress, but there are still many challenges and room for improvement. Traditional knowledge recommendation methods mainly include content-based recommendation and collaborative filtering-based recommendation.

Content-based recommendation methods rely on the features of the knowledge points and students' history to make recommendations by similarity calculation. Ghauth and Abdullah [11] proposed a content-based recommendation algorithm that makes recommendations by analysing textual features of knowledge points and students' learning preferences. However, this method has high computational complexity when dealing with large-scale data and it is difficult to capture the dynamic changes in students' interests. Another widely used method is Collaborative Filtering (CF). CF algorithms make recommendations by analysing similarities between users or similarities between items. Aghdam et al. [12] improved the CF algorithm by matrix decomposition technique to improve the accuracy and scalability of recommendations. However, the traditional CF algorithm performs poorly in dealing with rating sparsity and cold- start problems. In addition, Muruganandam and Srininvasan [13] proposed a recommendation method based on sequence patterns to mine useful learning paths by analysing the sequence of students' learning behaviours. However, this method fails to fully consider the weight and time factors of learning behaviours, resulting in the personalisation and accuracy of recommendation results to be improved.

Current research focuses on solving the problems of traditional methods and exploring new recommendation algorithms to improve the accuracy and personalisation of recommendations. Firstly, rating sparsity and cold-start problems are the main challenges faced by collaborative filtering algorithms. Li and Han [14] proposed a hybrid model-based approach that combines collaborative filtering with content-based recommendation to alleviate the rating sparsity problem. Secondly, the importance of the time factor in recommender systems is increasing. Liu et al. [15] improved the matrix decomposition algorithm by introducing the time factor, which enabled the recommender system to better capture the dynamic changes of user interests.

In addition, the application of sequence pattern mining techniques in recommender systems has received widespread attention. The research conducted by Nitu et al. [?] has presented an innovative approach to generating recommendations through the analysis of weighted sequential data, which can more accurately reflect students' learning interests and preferences by assigning different weights to learning behaviours. However, the method has high computational complexity when dealing with large-scale data. In order to further improve the recommendation effect, Liu et al. [16] proposed a recommendation algorithm combining deep learning and sequence pattern mining, which extracts features through a deep learning model and combines them with sequence patterns to make recommendations, and achieved better results. Nevertheless, these techniques exhibit certain constraints when addressing issues related to the fusion of different data sources and the mining of complex patterns.

1.2. **Motivation and contribution.** Through the above analysis, we can find that the problems of related research mainly include: (1) the fluctuating nature of user preferences, which are subject to temporal evolution, but it is often difficult for existing methods to capture this change; (2) the problem of data sparsity, with the increase in the number of knowledge points, the user rating matrix becomes more and more sparse, affecting the recommendation accuracy; (3) the demand for real-time recommendation, with the development of online learning platforms, the user's demand for real-time recommendation is growing, but the existing methods do not perform well in real-time, which affects the accuracy of recommendation. In order to solve these problems, this paper proposes a knowledge point recommendation algorithm based on Enhanced Correction Factors (ECF) and Weighted Sequential Pattern Mining (WSPM). The main innovations and contributions of this work include.

(1) Aiming at the problems of rating sparsity and cold start in traditional collaborative filtering methods, the ECF algorithm proposed in this paper can more accurately calculate the similarity between users by introducing the rating value difference correction factor, the item popularity difference correction factor, and the time decay factor, which can effectively alleviate the problems of rating sparsity and cold start, and improve the accuracy of recommendation.

(2) Aiming at the problem that existing methods fail to fully consider the sequence of users' learning behaviours, this paper introduces the WSPM algorithm, which can more accurately reflect the students' learning interests and preferences by assigning different weights to each item in the users' learning behaviours. The WSPM algorithm is able to mine the frequently occurring high-weight sequence patterns from the user's learning behaviour sequences, which provides stronger support for personalized recommendation.

(3) Aiming at the importance of the time factor in recommender systems, this paper captures the dynamic changes of user interests by introducing a time decay factor, which enables the recommender system to respond to user needs in a more timely and accurate manner. In addition, by combining ECF and WSPM algorithms, the recommender system proposed in this paper is able to provide more personalised and accurate knowledge point recommendations in different learning stages and environments, thus enhancing students' learning effectiveness and satisfaction.

2. **ECF based CF recommendation.**

2.1. **Traditional similarity calculation methods.** CF recommendation algorithm is one of the most widely used methods in modern recommender systems [17, 18], the core of which is to make recommendations by calculating the similarity between users. The calculation of similarity is the key of collaborative filtering algorithm, which determines the precision and efficacy of the generated recommendation outcomes. The traditional similarity calculation methods are mainly as follows:

2.1.1. *Cosine similarity.* It is one of the most commonly used similarity measures, and the basic idea is to measure the similarity of two vectors by calculating the cosine of the angle between them. Assuming that the ratings of a set of items by a user u and a user v are represented as vectors R_u and R_v, respectively, the cosine similarity is computed as follow:

$$\mathrm{sim}_{\cos}(u,v) = \frac{\sum_{i\in I_{uv}} R_{u,i}\cdot R_{v,i}}{\sqrt{\sum_{i\in I_{uv}} R_{u,i}^2}\sqrt{\sum_{i\in I_{uv}} R_{v,i}^2}} \tag{1}$$

where I_{uv} denotes the set of items jointly rated by user u and user v, and $R_{u,i}$ and $R_{v,i}$ denote the ratings of item i by user u and user v, respectively.

2.1.2. *Pearson Correlation Coefficient.* It is used to measure the linear correlation between two variables [19], and its value ranges from $[-1,1]$, with larger absolute values indicating stronger correlation. The Pearson correlation coefficient between user u and user v is calculated as follows:

$$\mathrm{sim}_{\mathrm{pearson}}(u,v) = \frac{\sum_{i\in I_{uv}}(R_{u,i}-\bar{R}_u)(R_{v,i}-\bar{R}_v)}{\sqrt{\sum_{i\in I_{uv}}(R_{u,i}-\bar{R}_u)^2}\sqrt{\sum_{i\in I_{uv}}(R_{v,i}-\bar{R}_v)^2}} \tag{2}$$

where $\bar{R}_u$ and $\bar{R}_v$ denote the average ratings of user u and user v on all rating items, respectively.

2.1.3. *Adjusted cosine similarity.* It takes into account the scoring bias when calculating similarity, making the similarity measure more reasonable, and is calculated as follows:

$$\mathrm{sim}_{\mathrm{adjcos}}(u,v) = \frac{\sum_{i\in I_{uv}}(R_{u,i}-\bar{R}_i)(R_{v,i}-\bar{R}_i)}{\sqrt{\sum_{i\in I_{uv}}(R_{u,i}-\bar{R}_i)^2}\sqrt{\sum_{i\in I_{uv}}(R_{v,i}-\bar{R}_i)^2}} \tag{3}$$

where $\bar{R}_i$ denotes the average rating of item i.

The above traditional similarity calculation methods have some problems in practical applications [20], such as failing to fully consider the differences in user ratings and the differences in item popularity. These problems will affect the accuracy of the recommendation results, and thus need to be improved on this basis to enhance the performance of the tweaking system. In the next section, we introduce improved algorithms that introduce correction factors to address these issues and improve the accuracy of similarity computation.

2.2. **Improved algorithm introducing correction factor.** Traditional similarity calculation methods are widely used in collaborative filtering recommender systems, but they have some shortcomings in practical applications, such as failing to take into account the differences in the values of user ratings, the differences in the popularity of items, and the changes in user interests over time. In order to improve the accuracy and performance of recommender systems, this section proposes an improved algorithm that introduces correction factors, i.e., the ECF algorithm. The algorithm improves the traditional similarity calculation method by introducing multiple correction factors to enhance the recommendation effect.

2.2.1. *Correction factor for differences in rating values.* In traditional similarity calculation methods, the difference in the numerical values of users' ratings for the same items is not taken into account. To solve this problem, we introduce a rating value difference correction factor to adjust the impact of user ratings. The rating value difference correction factor is calculated as:

$$\mathrm{fac}_1(u,v) = \exp\left(-\frac{1}{n}\sum_{i\in I_{uv}} |R_{u,i} - R_{v,i}|\right) \tag{4}$$

where $\mathrm{fac}_1(u,v)$ denotes the correction factor for the difference in the values of specific ratings between user u and user v; $R_{u,i}$ denotes the ratings of user u for item i, and $R_{v,i}$ is the ratings of user v for item i; I_{uv} is the set of jointly rated items for user u and user v; and n is the the number of jointly rated items of user u and user v. When the correction factor is smaller, the difference in rating values between two users is larger.

2.2.2. *Item popularity difference correction factor.* The traditional similarity calculation method has the same weight value for each item, and this method cannot fully reflect the similarity between users. For this reason, we propose a correction factor for item popularity differences, which is used to adjust the influence weights of items. The correction factor for item popularity difference is calculated as follows:

$$\mathrm{fac}_2(i) = \exp\left(1 - \frac{N(i)}{N_{\mathrm{all}}}\right) \tag{5}$$

where $\mathrm{fac}_2(i)$ is a correction factor measuring the difference in popularity of item i; $N(i)$ is the number of all users who have a rating record for item i; and N_{all} is the number of all users in the user-item rating matrix.

2.2.3. *Time decay factor.* Users' interest preferences change over time factors, so the time factor needs to be taken into account when calculating similarity. We introduce a time decay factor to capture the dynamic changes in user interest preferences. The time decay factor is calculated as follow:

$$f(t) = \exp(-\lambda t) \tag{6}$$

where λ is the decay rate, which represents the rate of decay of the user's interest preference; t is the user's rating time of the item. The change of user's interest preference can be realised by adjusting the size of λ, and if the user's interest preference changes faster, it can be adapted by increasing λ.

2.2.4. *Improved similarity calculation method.* By introducing the above correction factors into the traditional similarity calculation formula, an improved similarity calculation method can be obtained. Considering the score value difference correction factor, item popularity difference correction factor and time decay factor, the improved cosine similarity calculation method is shown as follow:

$$\mathrm{sim}_{\text{ecf-cos}}(u,v) = \frac{\sum_{i\in I_{uv}} R_{u,i}R_{v,i}\cdot \mathrm{fac}_2(i)\cdot \mathrm{fac}_1(u,v)\cdot f(t_{u,v})}{\sqrt{\sum_{i\in I_{uv}} R_{u,i}^2}\sqrt{\sum_{i\in I_{uv}} R_{v,i}^2}} \tag{7}$$

where $t_{u,v}$ denotes the time difference between the ratings of user u and user v on item i.

The improved modified cosine similarity is calculated as follows.

$$\mathrm{sim}_{\text{ecf-adjcos}}(u,v) = \frac{\sum_{i\in I_{uv}} (R_{u,i}-\bar{R}_i)(R_{v,i}-\bar{R}_i)\cdot \mathrm{fac}_2(i)\cdot \mathrm{fac}_1(u,v)\cdot f(t_{u,v})}{\sqrt{\sum_{i\in I_{uv}} (R_{u,i}-\bar{R}_i)^2}\sqrt{\sum_{i\in I_{uv}} (R_{v,i}-\bar{R}_i)^2}} \tag{8}$$

By introducing the above ECF, we improve the traditional similarity calculation method, which can more accurately reflect the similarity between users, thus improving the performance of the recommender system and the accuracy of the recommendation results.

3. Weighted Sequential Pattern Mining (WSPM) based recommendation.

3.1. Weighted sequential pattern.

In recommender systems, a user's behavioural sequence can reflect the user's interests and preferences [21, 22]. Therefore, pattern mining based on user behaviour sequences is important to improve the accuracy of recommendation results. Weighted Sequential Pattern (WSP) is a method that considers the weights of each item in the sequence when mining user behaviour sequence patterns [23], which can more accurately reflect the user's level of interest in different items.

Given a sequence of user behaviours $\mathcal{E}_u = (s_1, s_2, \ldots, s_n)$, where each s_i denotes a record of a user's visit or rating of item i at time t_i. A sequence pattern can be defined as a sequence of items that appear consecutively in a sequence of user behaviours. To take into account the importance of different items, we assign a weight w_i to each item, indicating the importance of item i in the sequence pattern.

The weighted sequence pattern $\mathcal{WEP}_u$ can be expressed as

$$\mathcal{WEP}_u = ((s_1, w_1), (s_2, w_2), \ldots, (s_n, w_n)) \tag{9}$$

where w_i is the weight of item i, which indicates the user's interest level or rating of item i.

Different from the traditional sequence pattern mining, WSP considers the user's preference weights for different items in the mining process, which makes the mining results reflect the actual interests of users more accurately [24, 25]. By assigning different weights to each item in the sequence, WSP can flexibly adapt to the personalised needs of different users and improve the accuracy and personalisation of the recommendation results. In the process of WSP mining, the weight information can be used to filter the unimportant items, thus reducing the computational complexity and improving the mining efficiency [26].

WSP can be widely used in various recommendation scenarios, e.g., by mining the sequence patterns of users' ratings of movies and considering the weights of the ratings, the movies that users are interested in can be recommended more accurately. In e-commerce platforms, by analysing the sequential patterns of users' browsing and purchasing behaviours and considering the weights of each product, the accuracy of product recommendation can be improved. In this study, we try to apply WSP to online education platforms, by mining the sequence patterns of students' learning behaviours and considering the weights of different courses, we can personalize the recommendation of suitable learning contents for students. By weighting the user behaviour sequence, the user's interest changes and preferences can be better captured, thus improving the accuracy and personalization of the recommendation results.

3.2. The WSPM algorithm.

Definition 1: Given a database of sequences $SDB = \{S_1, S_2, \ldots, S_n\}$ and let the set of sequence transactions containing item b in SDB be $S_b = \{S \mid b \in S, S \in SDB\}$ then define the weights of b in the sequence α as follow:

$$w_\alpha(b) = \frac{T_\alpha(b)}{L(\alpha)} \tag{10}$$

where $T_\alpha(b)$ represents the number of times item b occurs in α; $L(\alpha)$ is the length of α; $|SDB|$ and S_b represent the number of sequence transactions in the sequence database SDB and the number of sequence transactions in the set S_b, respectively.

From Definition 1, if the frequency of item b in sequence α is higher, and the number of sequences containing item b in the whole sequence database is less, then the sequence differentiation ability of item b is considered to be stronger, and it is more suitable for personalised recommendation.

Definition 2: Define the weights of the sequence S_i as follow:

$$w(S_i) = \sum_{b \in S_i} w_{S_i}(b) \tag{11}$$

From Definition 2, the weight of sequence S_i is the sum of the weights of all items contained in S_i.

Definition 3: Given a minimum weight ratio δ $(0 \leq \delta \leq 1)$, the minimum weight threshold is defined as follow:

$$w_{\min} = \delta \times \sum_{S_i \in SDB} w(S_i) \tag{12}$$

Sequences with weights not under the minimal weight limit are called high-weight sequences. Since high weight sequences no longer satisfy the anti-monotonicity commonly used in frequent pattern mining, it increases the complexity of traversing the search space. To improve the mining efficiency, we define the estimated weights of the sequences.

Definition 4: Define the estimated weights of the sequence α as follow:

$$w(\alpha) = \sum_{S_i \in g(\alpha)} w(S_i) \tag{13}$$

From Definition 4, it is clear that the estimated weight of the sequence α is the sum of the weights of all sequence transactions in the support set of α.

Theorem 1: Given sequences α and β, if $\alpha \subseteq \beta$, then $w(\alpha) \geq w(\beta)$.
Since $\alpha \subseteq \beta$, therefore $g(\alpha) \supseteq g(\beta)$. It follows from Definition 4:

$$w(\alpha) = \sum_{S_i \in g(\alpha)} w(S_i) \geq \sum_{S_j \in g(\beta)} w(S_j) = w(\beta) \tag{14}$$

Theorem 1 shows that the proposed WSPM algorithm satisfies anti-monotonicity, i.e., the subsequence of the estimated high-weight sequence is also the estimated high-weight sequence, which can speed up the mining process of the estimated high-weight sequence.

3.3. **Matching of sequences.** The prerequisite for accurate recommendation is to discover high-weighted sequences that have a high degree of match with the target user's access behaviour. We believe that the more consecutively identical items the sequence α has to the target sequence β, the higher the degree of match between α and β.

Given the target sequence β, and the sequence to be matched α, if there exists a common subsequence CS_i $(1 \leq i \leq n)$ of n itemsets occurring consecutively between the two, the match between α and β is defined as

$$\frac{\max\left(L(CS_1), \ldots, L(CS_n)\right)}{L(\beta)} \tag{15}$$

where $L()$ denotes the length of the sequence.

The pseudo-code of the WSPM algorithm proposed in this paper is shown in Algorithm 1.

In the above pseudo-code, the $\mathcal{WEP}_u$ function represents the weighted sequence pattern algorithm, which accepts the sequence pattern S, the projection database DS, and the minimum weight threshold w_{min} as inputs and outputs the full estimated high-weight sequence HAWS. The CalculateWeight function is used to compute the weights of the sequences, while the CalculateMatchingDegree function is used to compute the matching

Algorithm 1 WSPM: Weighted Sequential Pattern Mining

Inputs: Sequence database SDB, set of historical sequences SDB_t of target user u, minimum weight threshold w_{min}

Output: The next item to be accessed by the target user u

```
HAWS ← WEP_u(S, DS, w_min)                    ▷ Mine full estimated high weight sequence
for each β in SDB_t do
    for each α in HAWS do
        w(α) ← CalculateWeight(α)
        A(α, β) ← CalculateMatchingDegree(α, β)
    end for
    Recommend the item with the highest w(α) × A(α, β) to user u
end for
Return: recommended item for u
```

degree of two sequences. Finally, the recommendation algorithm recommends an item to the user u which is the item with the highest score based on the sequence weights and matching degree.

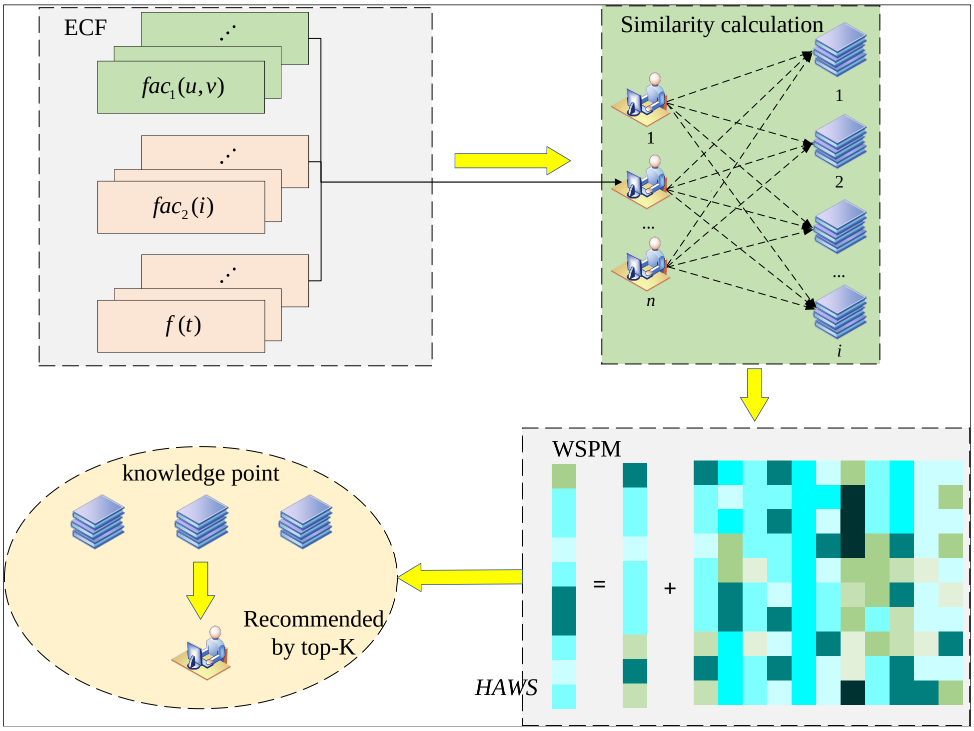

Figure 1. Overall framework of the ECF-WSPM methodology

4. **Knowledge point recommendation algorithm based on ECF-WSPM.** In this section, we will propose a new knowledge point recommendation algorithm based on ECF-WSPM by combining the previously introduced ECF and WSPM. The algorithm uses user rating data and behavioural sequence data to provide personalised knowledge point recommendation for users by calculating the similarity between users and mining user behavioural sequence patterns.

The overall framework of the ECF-WSPM method is shown in Figure 1. The ECF is responsible for the main work of user similarity calculation. The ECF calculates the similarity between users by introducing several correction factors. The correction factors include rating value difference correction factor, item popularity difference correction factor and time decay factor. The rating value difference correction factor adjusts the differences in users' ratings for the same item, making the similarity calculation more accurate. The popularity difference correction factor adjusts the weights according to the popularity of the items to avoid the excessive influence of popular items on the similarity calculation. The time decay factor adjusts the weights according to the time distribution of ratings to capture the changes in user interests over time. ECF improves the traditional cosine similarity and modified cosine similarity formulas, and combines the above correction factors to make the similarity calculation more reflective of the real similarity of user interests.

WSPM is responsible for mining behavioural sequence patterns and calculating sequence weights. The WSPM algorithm mines frequently occurring high-weighted sequence patterns from users' behavioural sequence data. These patterns reflect the user's behavioural patterns and interest preferences in different time periods. WSPM assigns weights to each item in the user's behavioural sequences, and the weights are calculated by considering ratings, item popularity and time factors. By calculating the total weight of each sequence pattern, important sequence patterns are identified as the basis for recommendation. WSPM generates all possible weighted sequence patterns by both item expansion and sequence expansion. The generated patterns are evaluated and frequent patterns with weights greater than or equal to a preset threshold are filtered out, which can be used for personalised recommendations.

In the integrated recommendation algorithm, the work of ECF and WSPM is combined to form a complete recommendation process: (1) Data Pre-Processing. Collect learners' learning behaviour data, including knowledge point access sequence, ratings, etc., and perform necessary cleaning and formatting. (2) ECF Similarity Calculation. Based on the traditional similarity calculation method, correction factors are introduced, such as the correction factor for the difference in user rating values and the difference in item popularity, as well as the time decay factor, to calculate the similarity between learners. (3) WSPM Sequence Construction. Using weighted sequence pattern mining technique, construct a sequence of knowledge points accessed by each learner, and assign a weight to each knowledge point in the sequence to reflect its importance in the recommendation process. (4) Sequence Pattern Mining. The WSPM algorithm is applied to mine frequent patterns in sequences, which represent learners' general learning paths and preferences. (5) Similar Learner Identification. According to the similarity obtained from ECF calculation, identify the group of learners who are similar to the target learners. (6) Knowledge Point Recommendation Generation. Combine the sequence patterns mined by WSPM and similar learners' preferences to generate a personalised list of knowledge point recommendations.

5. **Experiments and analysis of results.**

5.1. **Datasets.** In the field of education, there are several publicly available datasets that can be used for the study of knowledge recommendation algorithms, especially those based on user learning behaviour and knowledge points. These datasets can provide experimental data support for knowledge point recommendation algorithms based on ECF-WSPM. In this paper, three of these datasets are selected for training and testing of the algorithms.

ASSIST2009: The dataset was provided by the ASSISTMENTS online learning platform and used in several papers to evaluate the Knowledge Tracking Model. After deduplication, the dataset contains 325,673 interaction records from 4,151 students on 110 exercises.

ASSIST2015: dataset collected 708,631 records from 19,917 students on 100 knowledge points. This dataset has a higher number of records compared to the ASSIST2009 dataset, and the average number of student interactions is lower.

EdNet KT1 dataset: EdNet is a large-scale educational dataset containing multiple subsets, such as KT1, KT2, etc., covering data on students' behaviours in a wide range of learning environments, including answering questions, study time and correctness of answers.

The statistical information of the above three publicly available datasets is shown in Table 1.

Table 1. Statistical information on the 3 publicly available data sets

Statistical information	ASSIST2009	ASSIST2015	EdNet KT1
Number of answers	325,000	708,000	88,000
Number of learners	4,151	19,917	1,200
Number of question days	110	100	137
Number of knowledge points	110	100	74

5.2. **Experimental assessment indicators.** To evaluate the effectiveness of the proposed ECF-WSPM method, this paper compares the ECF-WSPM method with other comparative methods. Precision, Recall and F1 values are used as indicators to assess the effectiveness of each method. In this paper, 70% of the data submitted by each learner (test set) is selected from the test set to generate recommendation results. For each learner, top-K recommendations are made and the recommended knowledge points are compared with the remaining 30% of the submitted data (validation set) to verify the accuracy of the recommendations.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{16}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{17}$$

$$\text{F1} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{18}$$

where TP denotes the number of correctly predicted knowledge points in the recommendation validation set in the top-K recommendations, FP denotes the number of incorrectly predicted knowledge points in the recommendation validation set in the top-K recommendations, and FN denotes the number of knowledge points in the recommendation validation set in the other samples to be recommended.

5.3. **Performance comparison.** The proposed ECF-WSPM method is compared with three other collaborative filtering methods on the same dataset. KL-CF is a recommendation algorithm based on K-medoids clustering proposed by Deng et al. [27]. GA-CF is a recommendation method using genetic algorithm proposed by Alhijawi and Kilani [28]. R-CF is a robust collaborative filtering recommendation method using user item trust records proposed by Wang et al. [29].

In the actual recommendation process, 5 or 10 knowledge points are generally recommended for learners to choose. Therefore, this paper believes that according to the characteristics of online education, top-K takes 3, 5 and 7 to be more meaningful for practical recommendation. The statistics of Precision results of four different methods are shown in Table 2, which shows that the Precision index of various methods decreases with the increase of K value, and the ECF-WSPM method obtains the precision rate better than the other methods under different K values.

Table 2. The Precision of the four methods at different top-K values (%)

Data set	Top-K	KL-CF	GA-CF	R-CF	ECF-WSPM		
					$w_{min} = 0.3$	$w_{min} = 0.5$	$w_{min} = 0.7$
ASSIST2009	3	22.89	23.67	23.35	25.28	25.52	25.92
	5	16.49	17.37	18.58	19.69	19.78	19.88
	7	13.59	14.64	15.91	16.77	16.81	17.01
ASSIST2015	3	19.17	18.76	19.39	21.32	21.36	21.24
	5	13.99	13.77	15.33	16.63	16.65	16.63
	7	11.38	11.13	13.22	13.86	13.95	13.89
EdNet KT1	3	16.31	34.64	31.31	35.31	35.98	35.31
	5	12.37	25.17	23.17	25.17	25.37	25.57
	7	9.98	19.42	18.97	20.55	20.55	20.97

From Table 3, it can be seen that the Recall index of each method increases with the number of top-K recommendations, and the Recall of ECF-WSPM method can be optimal.

Table 3. The Recall of the four methods at different top-K values (%)

Data set	Top-K	KL-CF	GA-CF	R-CF	ECF-WSPM		
					=0.3	=0.5	=0.7
ASSIST2009	3	52.08	50.31	51.09	52.22	52.61	53.6
	5	58.44	56.06	61.5	60.56	60.84	61.63
	7	63.75	60.25	69.08	68.23	68.07	68.3
ASSIST2015	3	46.62	45.72	49.78	52.4	51.97	52.02
	5	53.31	52.33	61.12	62.91	62.36	62.74
	7	37.61	57.34	66.96	67.6	67.04	66.77
EdNet KT1	3	43.24	71.82	63.23	71.99	74.16	70.49
	5	50.98	80.56	75.98	79.44	79.81	81.14
	7	57.31	85.31	82.48	87.48	85.48	85.47

Table 4 records the F1 values of each method on the three datasets at different top-K values. The experimental results show that the ECF-WSPM method outperforms the other methods in terms of the F1 values recommended for different top-K.

Comprehensively analysing the above 3 tables, the ECF-WSPM recommendation method proposed in this paper outperforms the other methods on ASSIST-2009, ASSIST2015 and EdNet KT1 in each evaluation index. By comparing the above three tables, this paper concludes that the ECF-WSPM method has better performance when the number of recommendations $K = 3$ and the minimum weight threshold $w_{min} = 0.5$. Among them, on the EdNet KT1 dataset, the MDKT-GCCF method improves the results of Precision, Recall, and F1 by 18.36%, 29.81%, and 23.57%, respectively, compared to the KL-CF method, and by 1.28%, 2.17%, and 2.71%, respectively, compared to the GA-CF method.

Table 4. The F1 of the four methods at different top-K values (%)

Data set	Top-K	KL-CF	GA-CF	R-CF	ECF-WSPM		
					$w_{min} = 0.3$	$w_{min} = 0.5$	$w_{min} = 0.7$
ASSIST2009	3	22.89	23.67	23.35	25.28	25.52	25.92
	5	16.49	17.37	18.58	19.69	19.78	19.88
	7	13.59	14.64	15.91	16.77	16.81	17.01
ASSIST2015	3	19.17	18.76	19.39	21.32	21.36	21.24
	5	13.99	13.77	15.33	16.63	16.65	16.63
	7	11.38	11.13	13.22	13.86	13.95	13.89
EdNet KT1	3	16.31	34.64	31.31	35.31	35.98	35.31
	5	12.37	25.17	23.17	25.17	25.37	25.57
	7	9.98	19.42	18.97	20.55	20.55	20.97

6. **Conclusion.** A knowledge point recommendation based on ECF-WSPM is proposed. By comprehensively considering rating differences, item popularity and time factors, and combining high-weight pattern mining with user learning behaviour sequences, the ECF-WSPM method effectively alleviates the shortcomings of traditional collaborative filtering algorithms in dealing with rating sparsity and cold-start problems. Firstly, the user similarity calculation method is improved using ECF. By introducing rating value difference correction factor, item popularity difference correction factor and time decay factor. Second, combining with the WSPM technique, frequently occurring high-weighted sequence patterns are mined from users' learning behaviour sequences. These patterns not only reflect users' learning paths and preferences, but also reflect students' interests more accurately by assigning different weights, thus providing an important basis for personalised recommendation. The experimental results of this paper's method on several public education datasets show that the accuracy of the recommendation results of the ECF-WSPM method is significantly improved, especially in coping with the sparsity of ratings and the user cold-start problem. In addition, the method in this paper is able to respond to the dynamic changes of user interests in a timely manner and provide more personalised learning resource recommendations.

Acknowledgment. This work is supported by the Guangdong Provincial Department of Education Characteristic Specialty Project: "Information Management and Information System" (No. JXTD201901).

REFERENCES

[1] S. S. Khanal, P. Prasad, A. Alsadoon, and A. Maag, "A systematic review: machine learning based recommendation systems for e-learning," *Education and Information Technologies*, vol. 25, no. 4, pp. 2635-2664, 2020.

[2] Q. Guo, F. Zhuang, C. Qin, H. Zhu, X. Xie, H. Xiong, and Q. He, "A survey on knowledge graph-based recommender systems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 8, pp. 3549-3568, 2020.

[3] B. Walek, and V. Fojtik, "A hybrid recommender system for recommending relevant movies using an expert system," *Expert Systems with Applications*, vol. 158, 113452, 2020.

[4] U. Javed, K. Shaukat, I. A. Hameed, F. Iqbal, T. M. Alam, and S. Luo, "A review of content-based and context-based recommendation systems," *International Journal of Emerging Technologies in Learning*, vol. 16, no. 3, pp. 274-306, 2021.

[5] Z. Fayyaz, M. Ebrahimian, D. Nawara, A. Ibrahim, and R. Kashef, "Recommendation systems: Algorithms, challenges, metrics, and business opportunities," *Applied Sciences*, vol. 10, no. 21, 7748, 2020.

[6] T.-Y. Wu, J. C.-W. Lin, U. Yun, C.-H. Chen, G. Srivastava, and X. Lv, "An efficient algorithm for fuzzy frequent itemset mining," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 5, pp. 5787-5797, 2020.

[7] T.-Y. Wu, J. C.-W. Lin, Y. Zhang, and C.-H. Chen, "A Grid-Based Swarm Intelligence Algorithm for Privacy-Preserving Data Mining," *Applied Sciences*, vol. 9, no. 4, p. 774, 2019.

[8] Y. Ma, Y. Peng, and T.-Y. Wu, "Transfer learning model for false positive reduction in lymph node detection via sparse coding and deep learning," *Journal of Intelligent & Fuzzy Systems*, vol. 43, no. 2, pp. 2121-2133, 2022.

[9] C. Chand, A. Thakkar, and A. Ganatra, "Sequential pattern mining: Survey and current research challenges," *International Journal of Soft Computing and Engineering*, vol. 2, no. 1, pp. 185-193, 2012.

[10] J.-W. Han, J. Pei, and X.-F. Yan, "From sequential pattern mining to structured pattern mining: a pattern-growth approach," *Journal of Computer Science and Technology*, vol. 19, no. 3, pp. 257-279, 2004.

[11] K. I. Ghauth, and N. A. Abdullah, "Learning materials recommendation using good learners' ratings and content-based filtering," *Educational Technology Research and Development*, vol. 58, pp. 711-727, 2010.

[12] M. H. Aghdam, M. Analoui, and P. Kabiri, "Collaborative filtering using non-negative matrix factorisation," *Journal of Information Science*, vol. 43, no. 4, pp. 567-579, 2017.

[13] S. Muruganandam, and N. Srininvasan, "Personalised e-learning system using learner profile ontology and sequential pattern mining-based recommendation," *International Journal of Business Intelligence and Data Mining*, vol. 12, no. 1, pp. 78-93, 2017.

[14] H. Li, and D. Han, "A time-aware hybrid recommendation scheme combining content-based and collaborative filtering," *Frontiers of Computer Science*, vol. 15, no. 4, 154613, 2021.

[15] H. Liu, C. Zheng, D. Li, X. Shen, K. Lin, J. Wang, Z. Zhang, Z. Zhang, and N. N. Xiong, "EDMF: Efficient deep matrix factorization with review feature learning for industrial recommender system," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 4361-4371, 2021.

[16] P. Nitu, J. Coelho, and P. Madiraju, "Improvising personalized travel recommendation system with recency effects," *Big Data Mining and Analytics*, vol. 4, no. 3, pp. 139-154, 2021.

[17] B. Liu, "Based on intelligent advertising recommendation and abnormal advertising monitoring system in the field of machine learning," *International Journal of Computer Science and Information Technology*, vol. 1, no. 1, pp. 17-23, 2023.

[18] L. Wu, X. He, X. Wang, K. Zhang, and M. Wang, "A survey on accuracy-oriented neural recommendation: From collaborative filtering to information-rich recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 5, pp. 4425-4445, 2022.

[19] M. Srifi, A. Oussous, A. Ait Lahcen, and S. Mouline, "Recommender systems based on collaborative filtering using review texts—a survey," *Information*, vol. 11, no. 6, 317, 2020.

[20] J. Bobadilla, S. Alonso, and A. Hernando, "Deep learning architecture for collaborative filtering recommender systems," *Applied Sciences*, vol. 10, no. 7, 2441, 2020.

[21] A. Gazdar, and L. Hidri, "A new similarity measure for collaborative filtering based recommender systems," *Knowledge-Based Systems*, vol. 188, 105058, 2020.

[22] Z.-H. Zhang, F. Min, G.-S. Chen, S.-P. Shen, Z.-C. Wen, and X.-B. Zhou, "Tri-partition state alphabet-based sequential pattern for multivariate time series," *Cognitive Computation*, vol. 14, no. 6, pp. 1881-1899, 2022.

[23] X. Luo, H. Wu, Z. Wang, J. Wang, and D. Meng, "A novel approach to large-scale dynamically weighted directed network representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 9756-9773, 2021.

[24] W. Gan, J. C.-W. Lin, J. Zhang, P. Fournier-Viger, H.-C. Chao, and S. Y. Philip, "Fast utility mining on sequence data," *IEEE Transactions on Cybernetics*, vol. 51, no. 2, pp. 487-500, 2020.

[25] A. Onan, and M. A. Toçoğlu, "A term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification," *IEEE Access*, vol. 9, pp. 7701-7722, 2021.

[26] H. Phan, O. Y. Chén, M. C. Tran, P. Koch, A. Mertins, and M. De Vos, "XSleepNet: Multi-view sequential model for automatic sleep staging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5903-5915, 2021.

[27] Y. Wu, K. Li, G. Zhao, and X. Qian, "Personalized long-and short-term preference learning for next POI recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 4, pp. 1944-1957, 2020.

[28] J. Deng, J. Guo, and Y. Wang, "A Novel K-medoids clustering recommendation algorithm based on probability distribution for collaborative filtering," *Knowledge-Based Systems*, vol. 175, pp. 96-106, 2019.

[29] B. Alhijawi, and Y. Kilani, "A collaborative filtering recommender system using genetic algorithm," *Information Processing & Management*, vol. 57, no. 6, 102310, 2020.

[30] F. Wang, H. Zhu, G. Srivastava, S. Li, M. R. Khosravi, and L. Qi, "Robust collaborative filtering recommendation with user-item-trust records," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 4, pp. 986-996, 2021.