

Retraction

Retracted: Building an Open Sharing Platform for ELT Courses: The Example of MOOC Network

Advances in Multimedia

Received 15 August 2023; Accepted 15 August 2023; Published 16 August 2023

Copyright © 2023 Advances in Multimedia. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] N. Li and Y. Zhang, "Building an Open Sharing Platform for ELT Courses: The Example of MOOC Network," *Advances in Multimedia*, vol. 2022, Article ID 8793521, 9 pages, 2022.

Research Article

Building an Open Sharing Platform for ELT Courses: The Example of MOOC Network

Na Li  and Yitian Zhang 

Zhengzhou Preschool Education College, Zhengzhou 450000, China

Correspondence should be addressed to Na Li; lina02@zzpec.edu.cn

Received 6 September 2022; Revised 19 September 2022; Accepted 26 September 2022; Published 12 October 2022

Academic Editor: Tao Zhou

Copyright © 2022 Na Li and Yitian Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

English language teaching (ELT) has become an essential and indispensable part of primary education in today's increasingly frequent international exchanges, and MOOC, as a massive online teaching model integrating content and learning support services, is leading a pedagogical transformation, providing a historic opportunity to build an open sharing platform for ELT. However, developing an open sharing platform around MOOC networks is still challenging, especially for English courses, where various course designs significantly increase the difficulty of tracking students' learning status. Therefore, we introduce RNN-based sequence-to-sequence knowledge tracing models as the software foundation of the shared platform. The transformer model is further chosen to simulate students' historical learning trajectories to solve the problem of long-term dependency in traditional models. The research results have important theoretical and practical implications for building an open sharing platform for ELT courses.

1. Introduction

With the development of computer and Internet technologies in education, electronic online education has emerged and started to be widely used. One of the major online course formats that emerged from the Internet+Education is Massive Open Online Courses (MOOC). Platforms for taking massive online courses have several advantages over traditional forms of schooling. In the 2020 epidemic, online education is gaining traction in educational activities because of its online nature and gaining further attention from educational participants and researchers. Thanks to the convenience of information technology and the Internet, online education platforms can often support thousands of learners in the same course. This quantitative change has led to a qualitative shift that distinguishes MOOCs from traditional classroom formats in terms of scale and their online nature. However, MOOCs have some shortcomings compared to conventional education formats that have been fully developed. Compared to traditional classrooms, the failures of MOOC platforms are mainly due to their "online" and "large-scale" nature.

Regarding delivery, the existing online courses are not as practical as traditional classrooms regarding quality interaction between teachers and students. The MOOC platforms are not as effective as conventional classrooms in personalized instruction due to their large scale [1]. In addition, MOOC platforms have other disadvantages, such as the linearity of the course, student attention, teacher management, and some inherent design problems that need to be improved.

One of the researchers' concerns is the lack of personalized instruction on mooring platforms. While students participating in online education platforms are free to browse and choose their courses, they often lack the timely feedback and guidance that they would receive in a traditional classroom because the instructor's courses are usually delivered offline, and some courses have a large number of participants [2]. The main problem of this paper is the lack of personalized tutoring in online education platforms. The proposed educational solution is the intelligent tutoring system (ITS) [3]. The ITS is an ideal form of intelligent one-to-one teaching by simulating a human teacher and is a typical

application of AI technology in education. One of the expected functions of the system is to automatically develop learner-appropriate learning activities and instructional strategies, for which learner modeling is required [4].

Knowledge tracing is an important study in learner modeling. However, few studies discuss how to develop an open sharing platform for English course resources based on MOOC networks, which significantly hinders the prospect of using only technology in personalized English teaching. This paper proposes a new graph-based and transformer-based knowledge tracing model instead of the traditional recurrent neural network (RNN) for modeling students' historical learning trajectories to provide technical support for developing an open sharing platform for English teaching resources.

2. Literature Review

In the traditional classroom, teachers can only teach to the average level of students due to the inconsistent learning pace, making it difficult to effectively implement individualized and diversified teaching, making the possibility of "teaching to the student's abilities" minimal. Because it is challenging to teach in a targeted manner, the distribution of exercises is "too many but not precise"; students are faced with many recurring practices that they have already mastered, making learning efficiency plummet. The drawbacks mentioned above in traditional education can be solved by knowledge tracing. Knowledge tracing can provide students with personalized feedback [5] and suggestions for their learning path [6], allowing them to practice effectively on their weak knowledge points, preventing them from doing the problems they have already mastered blindly over and over again, which not only improves their learning efficiency but also stimulates their interest in learning new knowledge. This allows teachers to intervene scientifically in their students' learning process and truly personalize and scale up the teaching process. How to track students' learning process, assess their current mastery status, and recommend targeted exercises for them is a significant challenge in the field of knowledge tracing.

2.1. What Is Knowledge Tracing Model. Knowledge Tracing (KT), a fundamental task in an adaptive learning platform [7], can be described as the following problem. Given a learner's learning history,

$$X_{t-1} = (x_1, x_2, \dots, x_{t-1}), \quad (1)$$

$$x_i = (e_i, r_i), r_i \in \{0, 1\}. \quad (2)$$

$r_i \in \{0, 1\}$ indicates the correctness of the student's answer to exercise e_i . The purpose is to predict the probability that the student will answer question e_t correctly at time t ; that is as follows:

$$P(r_t = 1 | e_t, X_{t-1}) \quad (3)$$

Accurate knowledge tracing can recommend targeted

learning paths for students and provide exercises at the right difficulty level to improve student learning efficiency and engagement.

There are prominent manage-tracing-oriented researchers at home and abroad. The mainstream knowledge-tracing models are mainly divided into the following three categories: Bayesian knowledge-tracing, factorization, and deep learning knowledge-tracing models [8].

Bayesian knowledge tracing (BKT), essentially a Hidden Markov Model, assumes that the student's knowledge state is a set of known and unknown binary variables [9]. Since BKT models each knowledge state separately, it does not capture the relationship between different knowledge; moreover, BKT assumes that if a learner has mastered an understanding, they will keep getting and remembering that knowledge for the rest of the learning process.

Deep knowledge tracing (DKT) uses a long and short-term memory network (LSTM) to model students' learning states, which can predict students' future performance more accurately than the BKT model [10]. However, since DKT only uses one hidden variable to summarize students' knowledge states, it is difficult to track the shift of learners' knowledge states and explore the corresponding weaknesses [11]. Considering the complex forgetting mechanism of the human brain, DKTforget adds the following three forgetting factors to DKT: repetition time distance, sequence time distance, and a number of past attempts. To solve the problems in DKT, researchers proposed the dynamic key-value memory network (DKVMN) based on the memory enhancement network [12] by adding two external matrices, using the critical static matrix (Mk) for storing knowledge representations and the dynamic value matrix (Mv) for storing and updating the knowledge state representations corresponding to each knowledge.

Since traditional RNNs and their variants suffer from long-term dependency, there is a shortcoming of capturing dependent information in long-sequence problems of students, which the attention mechanism can mitigate. The attention mechanism focuses on the focus, which is equivalent to increasing the weight value of the focus part to improve the model's ability to process information. With the introduction of the transformer, the attention mechanism has become a hot topic of research. Due to its excellent performance, the attention mechanism is widely used in neural network models. In the field of knowledge tracing, self-attentive knowledge tracing (SAKT) applies the self-attention mechanism [13], which is set up in much the same way as the transformer, mapping the embedding e_t of the problem e_t to Q; the student's interaction, the binary $X_t = (e_t, r_t)$ is mapped to K and V, respectively. Context-aware attentive knowledge tracing (AKT) incorporates both attentional and self-attentive mechanisms into the model and uses the self-attentive mechanism to encode the embedding representations of the corresponding questions and knowledge states and applies the attentional mechanism to retrieve the similarity between the learner's historical learning records and the current exercise and to retrieve the similarity between the learner's history and the current exercise. The scaled dot product formula obtains the student's knowledge state score.

The sparsity of educational data also poses a significant challenge to the performance of the knowledge tracing model. Graph-based interaction model uses a graph convolutional neural network to fully explore the higher-order information between questions and knowledge, alleviating the problem caused by data sparsity [14]. The top K exercises that are most relevant to the current exercise are filtered according to the attention weight score. The learner's current knowledge state is derived by combining the knowledge mastery levels of the top K exercises.

2.2. The Rise of MOOC. The traditional teaching mode is the mainstream of school education. Still, with the development of the times and the advancement of technology, the drawbacks of this teaching mode are becoming more and more prominent, and the traditional English teaching mode is full of problems. First, the class time is fixed, single, and not flexible enough. The conventional teaching mode selects the learning time, specified time, and set place and lacks flexibility, which is not conducive to mobilizing students' learning autonomy. Second, the teaching content is boring; the teaching method is single, and the classroom lacks interaction and communication. In the traditional English classroom, the teacher is the center of the whole teaching activity, and the teacher transmits English knowledge to the students in a one-way manner. In contrast, the students receive it passively and without thinking.

The ability to express and apply language is neglected. The teacher and the students are unequal; the teacher is the master of knowledge; the "teacher's dignity" is emphasized, and there is not much communication and interaction between the teacher and the students. Students are afraid to speak English because they are afraid of making mistakes and being ridiculed, and our English is mostly dumb English. This lack of confidence is particularly evident among Chinese students. The third is the lack of a linguistic environment. It has been proven that a good learning environment has a positive effect on learning efficiency. In particular, the acquisition of language requires context. For most Chinese college students, the language environment under the traditional teaching mode is too narrow, and we lack a good English learning environment. Language needs to be communicative. If we want to learn a language well, we need more opportunities to communicate and practice. Through communication and interaction, we can open up all of our senses and experience the language firsthand. Fourth, there is a solid test-taking mentality. In China, in colleges and universities, we are often taught that learning is all about exams, and if you get a high score on an exam, you are a success, and if not, you are a failure. The direct consequence of this mentality is that the development of practical application skills is neglected. In learning English, the natural performance is to ignore the cultivation of practical application of English. In conclusion, the traditional English teaching model can no longer meet the diversified needs of modern society. We need to explore a new education model to eliminate the constraints of English learning and improve the overall English ability of college students, especially English listening and speaking ability [15].

MOOC is a large-scale and open online course, which is the inevitable result of "Internet+Education" in the era of education informatization. The first letter, "M," in "MOOC," stands for large-scale, which means that the MOOC courses are large and numerous; the middle two letters, "OO," stand for openness and online, and the last letter, "C," stands for the course. The middle two letters, "OO," stand for transparency and online, and the previous note, "C," stands for class, which means "large-scale, open and shared online courses." In essence, the catechism is a virtual online course that allows students to learn without time and space constraints and with greater flexibility and freedom. Compared to traditional methods, the catechism has significant features such as openness, large scale, and online learning. The catechism's transparency allows all students access to equitable educational resources and promotes the equity of modern education. The large-scale and online learning features enable students to learn English more efficiently without the old way of learning. In conclusion, education informatization is an inevitable trend in reforming English education and teaching in higher education. MU classes based on modern mobile Internet technology play an essential role in the reform of English teaching.

2.3. The Merits of MOOC. MOOC has the inherent characteristics of new information technology, such as big data, the Internet of Things, and artificial intelligence. With strong technical support, if universities can implement the "MOOC+English education" model and launch a series of public and free English learning courses with English teaching objectives and contents, it will undoubtedly enrich the English course resources of universities and promote the overall improvement of "teaching" and "learning." It will undoubtedly increase the English course resources of universities and promote the overall improvement of the "teaching" and "learning" levels. Some famous MOOC learning platforms in China have provided teachers and students with rich and high-quality online course resources, such as China University MOOC and Huawei MOOC. The Chinese University MOOC is an online learning platform of high-quality national courses jointly created by Netease Cloud Platform and Love Course Network and recognized by the relevant state departments. Students can freely browse and download free course resources according to their own learning needs. In addition, some of the top universities in China, such as Peking University and Tsinghua University, have also launched high-quality online course resources for students and faculty or the community, which makes the English learning process more convenient and efficient.

English MOOCs allow students to learn English in a more accessible and flexible way. Traditional English classes are limited by time and space, making it challenging to ensure that students can learn English at any time and in any environment. However, the mobile information technology-based English MOOC has broken through the time and space limitations of traditional English classroom teaching and has changed students' English learning methods dramatically [16]. Students no longer need to be constrained by external conditions such as place, time, and

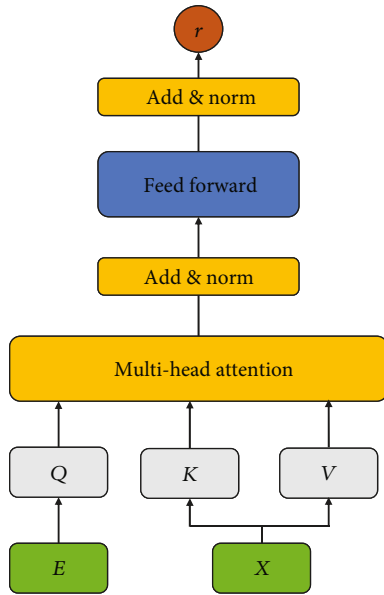


FIGURE 1: SAKT structure diagram.

environment to learn English. They can learn English anytime, anywhere, and are happy if they have the Internet and computers. With the help of mobile Internet and computers, college students can browse online English MOOC resources provided by major universities and choose their favorite course resources to study at any time. Students cannot only accumulate rich resources but also rely on the MOOC platform to consult and seek help online for challenging questions to improve their learning efficiency. In addition, the rich online resources allow students to change from passive learning to active learning, which is very beneficial to their English learning [17].

By introducing MOOCs into English teaching, universities can motivate students to learn independently. To encourage students to learn English, teachers need to make students interested in learning English and give them a sense of satisfaction and accomplishment from learning English. MOOC is a large-scale and systematic online open course based on information technology. The vast amount of online course resources enrich the content of English classroom teaching and motivates students to learn English. The enormous amount of online course resources enhances the range of English classroom teaching and mobilizes and stimulates students' interest in learning. During the learning process, students can repeatedly view the same course resources until they have fully absorbed them. Students can also ask for help online and communicate with master teachers if they do not understand something [18]. This online learning method dramatically enhances students' initiative and autonomy in learning English and gives them a sense of satisfaction and self-confidence.

However, for foreign language MOOC developers, it is a great challenge to demonstrate the scientific and systematic construction of courses to learners, and to change the current phenomenon of a small number of quality courses, teaching philosophy still focusing on knowledge transfer, weak teaching interaction, and low retention rate of students

in online learning. If the open sharing platform considers these problems, it can significantly alleviate students' difficulties in using English MOOCs online.

More importantly, few studies have discussed how to develop corresponding open sharing platforms around MOOC networks and ELT course resources, which significantly hinders the IT reform of ELT. Therefore, this paper proposes a new knowledge-tracking model to provide a potential technical path for developing an open sharing platform for English course resources.

3. Methodology

The LSTM is a variant of the RNN that allows the network to selectively add new information and forget old information by introducing new internal states and gating mechanisms, alleviating the above disadvantages of the RNN. However, since LSTM is still based on RNN in nature, it can handle sequences of up to 100 magnitudes, and it has been pointed out in the literature that LSTM does not entirely solve the problem of gradient disappearance in RNNs [1, 19]. In knowledge tracing, the learner's historical learning record is a long sequence problem, and the current learning state is highly dependent on the previous mastery level. To effectively capture the dependency information among long sequence exercises from the beginning, we propose a graph-based and transformer-based model for Knowledge Tracing (GTRKT), which abandons the traditional RNN sequence-to-sequence network model and uses a transformer to model students' historical learning trajectories.

3.1. Transformer and Forgetting Factor Matrix. In traditional RNNs, the input at time t needs to wait for the output from time $t-1$, which limits the parallel computation capability of the model. The ability of variant networks such as RNN still suffers from significant deficiencies for longer sequences. The proposal of transformer solves the above problem by bringing the idea of attention mechanism to the extreme. The authors abandoned the traditional RNN and its variants (LSTM, GRU, etc.) and used self-attention to compute the input and output of sequences. The transformer is originally used in the field of machine translation. Transformer was initially used in machine translation with good results, as the core of knowledge tracing is to model the long sequence learning paths of learners and explore the similarity between exercises, which is exactly the workflow of transformer. There has been a lot of work applying the transformer model to the knowledge tracing domain. SAKT is a transformer-based knowledge tracing model with a setup similar to a single encoder module in transformer, with the structure shown in Figure 1, mapping the embedding E of the exercises to Q and the learner interactions X to K and V .

In a subsequent study on the application of transformer in the knowledge tracing domain, it was argued that the attention layer of the SAKT model is too shallow to capture the dynamic learning changes of students, and that SAKT does not apply self-attention to exercises and interactions but feeds embedded features directly into the attention

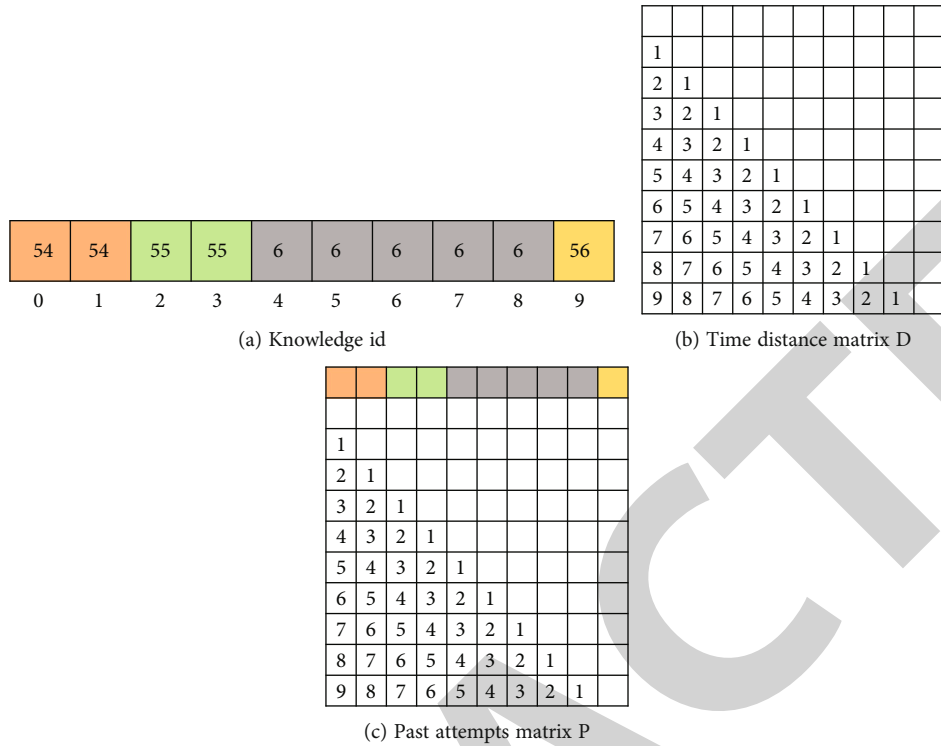


FIGURE 2: Forgetting factor matrices.

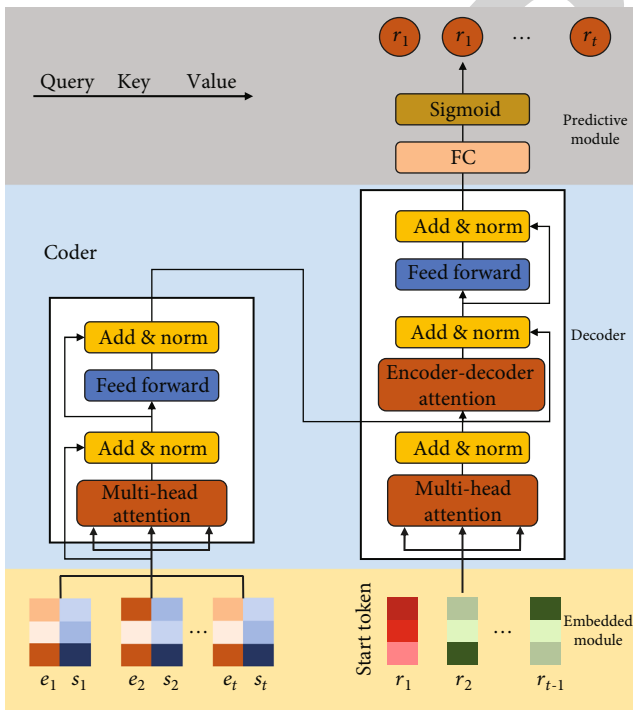


FIGURE 3: GTRKT model structure diagram.

sublayer. To solve the above problems, the SAINT model is proposed. The main idea of SAINT is also based on the transformer, which consists of two parts, an encoder and decoder, and its internal settings are consistent with the transformer. In SAINT, the embedding of the problem is

trained with the model. It is embedded to obtain the final vector representation. The GTRKT proposed in this chapter improves on this by modeling the relationship between exercises and knowledge using R-GCN and learning the embedding representations of the activities and knowledge as the input to the encoder. In addition, the students' forgetting behavior is taken into account. Two forgetting factors, time distance and some past attempts, are added to the attention distribution calculated using the scaled dot product formula.

Figure 2 details the process of calculating the two forgetting factor matrices. In particular, Figure 2(a) shows the knowledge graph of 10 consecutive questions for a random student, which shows that the learner first practiced Knowledge 54 twice, then Knowledge 55 twice, then Knowledge 6 five times, and finally Knowledge 56 once. Since the main task of KT is to predict the probability of answering the current exercise correctly based on the student's past performance, only the matrix values that satisfy the condition $\tau < t$ are considered in the calculation of the forgetting factor matrix. Figure 2(b) shows the computed results of the time distance matrix D , and Figure 2(c) shows the computed results of the past attempt count matrix P . For matrix D , $D_{t,\tau}$ represents the time interval between the student's answer to question e_t and question e_τ . When $t = 1$, the value in the first row (starting from 0) of the D matrix is 1 because the historical learning trajectory is practicing knowledge 54 at $t = 0$. When $t = 2$, the values in the second row (starting from 0) of the D matrix are 2 and 1 because the historical learning trajectory is practicing knowledge 54 at $t = 0$ and practicing knowledge 54 at $t = 1$, respectively, so the values in the second row (starting from 0) of the D matrix are 2 and 1, respectively. For matrix P , $P_{t,\tau}$ represents the total

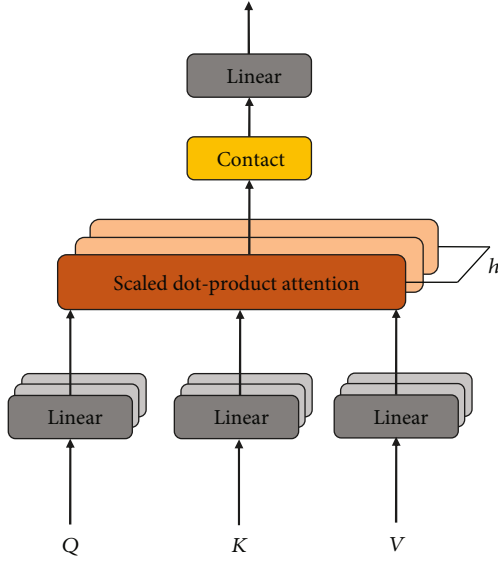


FIGURE 4: Multiheaded attention network.

0.5	0.2	0.8	0.5	$-\infty$	$-\infty$	1.0	0	0
1.3	1.2	0.7	1.3	1.2	$-\infty$	0.525	0.475	0
0.2	1.4	0.9	0.2	1.4	0.9	0.130	0.159	0.711
Origin			Mask			Softmax		

FIGURE 5: Masking operations.

number of exercises practiced before time e_τ that have the same knowledge as problem e_τ . When $t = 1$, the learner has practiced knowledge 54 once, so the value of the first row (starting from 0) of the P matrix is 1. When $t = 2$, the learner has practiced knowledge 54 twice, so the values of the second row (starting from 0) of the P matrix are 1 and 2, respectively.

Considering the forgetting curve of the human brain, the influence values between e_t and e_τ usually decay exponentially with time; in addition, considering that repetition is beneficial to deepen the memory of the human brain, the influence values of e_t and e_τ increase with the increase of the number of repetitions of the learner. Combining these two scenarios, the values of the distribution of attentional mechanisms are redefined as follows:

$$\alpha_{t,\tau} = \text{Softmax} \left(\frac{e^{-\theta D_{t-\tau}} \cdot e^{\beta D_\tau} \cdot Q_t K_t^T}{\sqrt{d_k}} \right) \in [0, 1]. \quad (4)$$

For this problem Q_t , the final attentional score is as follows:

$$A_t = \sum_{\tau} \alpha_{t,\tau} v_t. \quad (5)$$

3.2. GTRKT Model. We propose a GTRKT model, which differs from other transformer-based knowledge tracing

models by improving the scaling dot product formula of the basic attention mechanism by adding time distance and the number of past attempts of the learner. GTRKT uses RGCN to model the problem-knowledge relationship. As shown in Figure 3, the GTRKT model consists of three main modules. The first part is the embedding module, which uses RGCN to update the node embedding of exercises and knowledge. The second part is the transformer module, the main body of which is the encoder and decoder. They both consist of N identical layers, except that the decoder adds an encoder-decoder attention layer on top of the encoder. The third part is the prediction module, which is mainly a fully connected layer with *Sigmoid*.

The attention network takes Q , K , and V as inputs, which represent the *query*, *key*, and *value*, respectively. Figure 4 shows the schematic diagram of the multiheaded attention network, and Equations ((6)–(10)) show the overall computation process. The proposed multiheaded attention network allows the model to learn different behaviors based on the same attention mechanism given the same Q , K , and V . Applying the multiheaded attention network to the field of knowledge tracing is equivalent to making a comprehensive evaluation of students' learning in several different time dimensions based on a unified evaluation criterion.

$$Q_i = QW_i^O, \quad (6)$$

$$K_i = KW_i^K, \quad (7)$$

$$V_i = VW_i^V, \quad (8)$$

$$\text{head}_i = \text{Attention}(Q_i, K_i, V_i) = \text{Softmax} \left(\text{mask} \left(\frac{Q_i K_i^T}{\sqrt{d_k}} \right) \right) V_i, \quad (9)$$

$$\text{Multihead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) W^O. \quad (10)$$

Unlike a single attention, a multiheaded attention network is equivalent to allowing *query*, *key*, and *value* to undergo different linear transformations, combining different subspaces of the input parameters, using the scaled dot product formula, and inputting the h sets of linearly changed values into the attention convergence in parallel, with the final output being obtained by stitching the h attention heads together through a fully connected layer.

In the computation of the attention network of GTRKT, a mask mechanism is used to prevent the current position from paying attention to the subsequent positions. The implementation is shown in Figure 5, where the upper triangular part of the matrix $Q_i K_i^T$ is replaced by $-\infty$ after multiplying the dot product and is subjected to Softmax operation, which is used to zero the attention weight of the upper triangular position.

The feed forward neural networks (FNN) is applied to the output of the multiheaded attention sublayer, and *ReLU* is chosen as the activation function to increase the

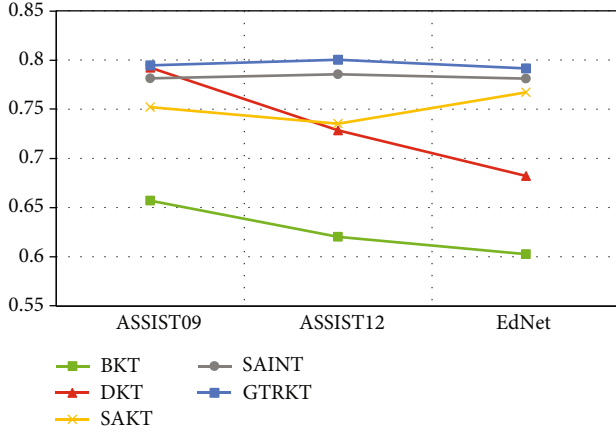


FIGURE 6: Comparison of different model.

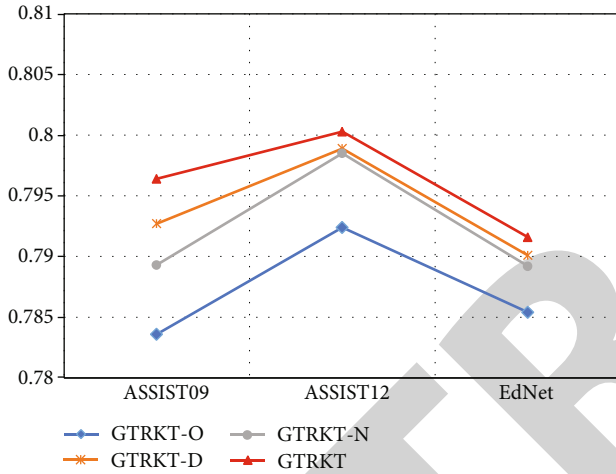


FIGURE 7: Comparison of different model with improved attention mechanism.

nonlinearity of the model.

$$F = \text{FFN}(M) = \text{ReLU}(MW_1 + b_1)W_2 + b_2, \quad (11)$$

where M is the input to the feedforward neural network, i.e., the output from the previous self-attentive sublayer; F represents the final output of the feedforward neural network, and W_1 , W_2 , b_2 are the learnable parameters.

The encoder consists of N identical layers. Each layer consists of two sub-layers: a self-attentive sub-layer and a feedforward neural network sub-layer. In each sublayer, a residual connection is used, which can be used to propagate features from the lower layers to the higher layers. In the context of knowledge tracing, the use of residual connections facilitates the propagation of students' recent exercises embedding to the last layer, allowing the model to make full use of the information at the lower layers. In addition, in order to avoid gradient explosion and disappearance in the training of the deep network, the different inputs are layer normalized, which has a good stabilizing and accelerating effect on the training of the network.

The decoder and encoder setup is similar and consists of N identical layers. Three sublayers in each layer are added with layer normalization and residual connectivity. The first layer is the self-attentive sublayer, whose output is denoted as S_1 ; the middle layer is the encoder-decoder attentive sublayer, whose output is denoted as S_2 ; and the tail layer is the feed forward neural network sublayer, which is the final output of the individual encoder O .

The output of the decoder is passed through a fully connected layer network with a *Sigmoid* function to obtain a sequence of learning state prediction results for the learner r_1, r_2, \dots, r_t .

$$L = -\sum_t (r_t \log r_t + (1 - r_t) \log (1 - r_t)). \quad (12)$$

4. Discussion and Results

ASSIST09 is a dataset of educational information collected by the ASSISTment online tutoring platform between 2009 and 2010, which contains multiple knowledge questions, i.e., a question that has various knowledge for the question. EdNet is a dataset of 131,441,538 learning interactions from 784,309 students collected by Santa since 2017 and is the largest publicly available education dataset today. Because the EdNet dataset is too large, 5,000 learners were randomly selected from the EdNetKT1 dataset for the experiment. The questions need to be preprocessed according to the Question file in the Contents folder, and the correct answer and tags are mapped to the corresponding questions in the EdNetKT1 dataset.

Figure 6 shows the experimental results of the GTRKT model compared with other benchmark KT models. The experimental data indicate that GTRKT outperforms the other KT models in all three datasets regarding AUC performance. The ASSIST12 dataset's GTRKT model is superior to other KT models. Considering that transformer does not have the problem of long-term dependence when dealing with long sequence problems, it can effectively extract the support between long interval problems, and the ASSIST12 dataset has the most extended average sequence of learned interactions per student compared to the other two datasets. Therefore, GTRKT performs better in the ASSIST12 dataset than DKT and traditional BKT modeled using LSTM networks. In the transformer-based KT model, the performance of the SAINT model is improved significantly in the three datasets compared with SAKT because the self-attentive mechanism is not applied to the exercises and interactions in SAKT. Still, only their embedded features are directly used as the input of the attention sublayer; that is, the number of attention layers in SAKT is too shallow, and the dependencies between the exercises are not well learned. The dependencies between activities are not well known. Our proposed GTRKT model shows a significant improvement over the SAINT model, with AUC performance improving by 1.68%, 1.88%, and 1.34% in the three datasets, respectively, with an average performance improvement of 1.63%, demonstrating the effectiveness of the proposed graph embedding and forgetting mechanism, which is

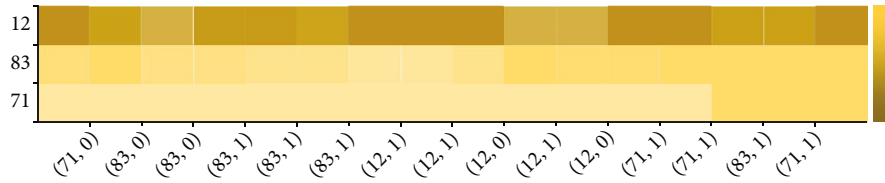


FIGURE 8: Heat map of sequence prediction results.

beneficial for enhancing the model’s performance in predicting students’ future performance.

To demonstrate the effectiveness of the forgetting mechanism, the following four sets of ablation experiments were conducted (Figure 7). GTRKTO indicates that the original scaled dot product formula is used; GTRKTD suggests that the time distance factor is added, and GTRKTN demonstrates the number of times the learner has tried in the past is added. GTRKT indicates that the above two forgetting factors are added. The experimental results show that the time distance and the number of past attempts of the learner help to improve the AUC performance of the model, and the time distance factor has a higher effect on the performance improvement than the number of past attempts of the learner factor.

Finally, we used a random sample of students in the ASSIST09 dataset to represent the sequential prediction results of the learner’s 20 consecutive learning records in the form of a heat map, as shown in Figure 8. The learner answered 20 straight questions, containing a total of 3 knowledge points. The vertical coordinates of the heat map represent the corresponding knowledge ids, and the horizontal coordinates are binary (s_i and r_i), representing the questions answered by the learner and the correctness of the answers, where the ids of the questions have been mapped to the corresponding knowledge. The darker the color of the square in the heat map, the closer it is to 1, which means the higher the predicted correctness of the learner’s answer to the question. In most cases, the student’s prior learning history directly influences the student’s current state of knowledge. As students repeatedly practice a knowledge point, the square’s color at their location will gradually deepen if they are correct and lighten if incorrect. From the sequential practice sequence of Knowledge 12 (black box in the figure), we can see that the learners’ knowledge state of Knowledge 12 constantly changes dynamically, and the model’s prediction is more consistent with the learners’ actual responses. By looking at the heat map of the sequences, we can see that the GTRKT model predicts the students’ knowledge state in a more realistic way, proving the model’s validity.

In sum, a large amount of experimental data proves that the GTRKT model outperforms the current mainstream knowledge tracing models on several datasets and has higher interpretability.

5. Conclusion

Education is a future-oriented endeavor, and the demands of modern society are becoming increasingly diverse. Some

reports have pointed out that personalized learning is one of the biggest challenges hindering technological progress. In the era of big data, based on the maturity of integrated technologies such as the Internet of Things and cloud computing, MOOCs based on big data technologies can analyze and mine educational data, design and develop courses suitable for students, realize personalized education in the digital environment, improve students’ thinking ability, innovation ability, and self-learning ability, and make students grow toward personalization. The development of personalized education is not only beneficial to the development of customized education. The development of personalized education is not only conducive to the development of students’ potential but also to the cultivation of innovative talents, which is especially important in English education.

The mission of the new era is to promote the deep integration of information technology and education. The development of cloud computing, social network media, and human-computer interaction IT technologies has led to the birth of MOOC, a prominent MOOC representative of the deep integration of information technology and higher education. It has created a new education model in which the needs of students’ personal and social realities are genuinely taken into account. MOOC reflects the educational values of independent learning, personalized learning, and lifelong learning and delivers high-quality educational resources to all corners of the world through information technology and network technology. Unlike traditional learning methods, MOOC is more concerned with the complete learning experience of students, and its emergence provides an effective way for students to personalize their English learning. Still, few studies have discussed how MOOC provides an open sharing platform for students’ English learning at the technological level.

Therefore, this paper proposes a GTRKT, which models students’ knowledge states using the transformer model compared to the RNN-based sequence-to-sequence knowledge tracing model. GTRKT uses RGCN to learn the embeddings of questions, extracting higher-order information between questions and knowledge better than other models that use ordinary embedding vectors for representation. GTRKT can remove higher-order information between questions and expertise than other models using embedding vectors. In addition, two forgetting factors, time distance and the number of attempts by the user, are added to the attention formula of the scaled dot product. In sum, we propose a new knowledge tracking model using graph embedding, attention mechanism, and transformer to improve the model’s performance in predicting students’ future English learning status.

This paper has some shortcomings, such as using a web dataset to test the performance of our proposed model. Still, given the deviation of the web dataset from the actual dataset, it is possible that the validity of the performance of the model in this paper needs to be further tested. We also suggest that future research consider incorporating more factors into the knowledge-tracking model as well as considering learning representations of the rich textual information contained in the questions and processing these heterogeneous data with techniques from fields such as image recognition and natural language to explore the relevance between the exercises better. These technological updates will provide technical support for developing an open sharing platform for ELT resources.

Data Availability

The datasets used during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

- [1] N. A. Albelbisi, "Development and validation of the MOOC success scale (MOOC-SS)," *Education and Information Technologies*, vol. 25, no. 5, pp. 4535–4555, 2020.
- [2] H. M. Dai, T. Teo, and N. A. Rappa, "Understanding continuance intention among MOOC participants: the role of habit and MOOC performance," *Computers in Human Behavior*, vol. 112, article 106455, 2020.
- [3] B. Wu, "Influence of MOOC learners discussion forum social interactions on online reviews of MOOC," *Education and Information Technologies*, vol. 26, no. 3, pp. 3483–3496, 2021.
- [4] A. A. Mubarak, S. A. Ahmed, and H. Cao, "MOOC-ASV: analytical statistical visual model of learners' interaction in videos of MOOC courses," *Interactive Learning Environments*, vol. 26, pp. 1–16, 2021.
- [5] L. Meng, M. Zhang, W. Zhang, and Y. Chu, "CS-BKT: introducing item relationship to the Bayesian knowledge tracing model," *Interactive Learning Environments*, vol. 29, no. 8, pp. 1393–1403, 2021.
- [6] K. Zhang and Y. Yao, "A three learning states Bayesian knowledge tracing model," *Knowledge-Based Systems*, vol. 148, pp. 189–201, 2018.
- [7] W. Gan, Y. Sun, X. Peng, and Y. Sun, "Modeling learner's dynamic knowledge construction procedure and cognitive item difficulty for knowledge tracing," *Applied Intelligence*, vol. 50, no. 11, pp. 3894–3912, 2020.
- [8] X. Wang, Z. Zheng, J. Zhu, and W. Yu, "What is wrong with deep knowledge tracing? Attention-based knowledge tracing," *Applied Intelligence*, vol. 52, pp. 1–12, 2022.
- [9] S. Liu, J. Yu, Q. Li et al., "Ability boosted knowledge tracing," *Information Sciences*, vol. 596, pp. 567–587, 2022.
- [10] R. Pelánek, "Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques," *User Modeling and User-Adapted Interaction*, vol. 27, no. 3-5, pp. 313–350, 2017.
- [11] P. I. Pavlik, L. G. Eglington, and L. M. Harrell-Williams, "Logistic knowledge tracing: a constrained framework for learner modeling," *IEEE Transactions on Learning Technologies*, vol. 14, no. 5, pp. 624–639, 2021.
- [12] B. Brahim and A. Lotfi, "A traces based system helping to assess knowledge level in e-learning system," *Journal of King Saud University-Computer and Information Sciences*, vol. 32, no. 8, pp. 977–986, 2020.
- [13] S. Liu, R. Zou, J. Sun et al., "A hierarchical memory network for knowledge tracing," *Expert Systems with Applications*, vol. 177, article 114935, 2021.
- [14] X. Song, J. Li, Q. Lei, W. Zhao, Y. Chen, and A. Mian, "Bi-CLKT: bi-graph contrastive learning based knowledge tracing," *Knowledge-Based Systems*, vol. 241, article 108274, 2022.
- [15] B. Jiang, S. Wu, C. Yin, and H. Zhang, "Knowledge tracing within single programming practice using problem-solving process data," *IEEE Transactions on Learning Technologies*, vol. 13, no. 4, pp. 822–832, 2020.
- [16] M. F. Asli, M. Hamzah, A. A. A. Ibrahim, and E. Ayub, "Problem characterization for visual analytics in MOOC learner's support monitoring: a case of Malaysian MOOC," *Heliyon*, vol. 6, no. 12, article e05733, 2020.
- [17] P. Guo, "MOOC and SPOC, which one is better?," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 8, pp. 5961–5967, 2017.
- [18] R. L. Moore and C. Wang, "Influence of learner motivational dispositions on MOOC completion," *Journal of Computing in Higher Education*, vol. 33, no. 1, pp. 121–134, 2021.
- [19] X. Zhao, "Implementation of English ICAI MOOC system based on BP neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 18, pp. 1–10, 2021.