Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 19 No. 11 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijim.v19i11.56057

PAPER

Intelligent Education Based on Mobile Learning: Transitioning from Traditional Classrooms to Adaptive Learning Environments

Jinqiu Wang (✉)

Dean's Office of Anhui Sanlian University, Hefei, China

luyimiao0@126.com

ABSTRACT

With the rapid advancement of information technology and the widespread adoption of mobile devices, a paradigm shift has been observed in education—from traditional classroom-based instruction to more flexible and personalized mobile learning environments. Mobile learning has not only eliminated spatial and temporal constraints but has also enabled new application scenarios for intelligent education. However, existing learning content recommendation systems exhibit notable limitations in addressing the dynamically evolving nature of learning environments and individualized learning needs. A predominant focus on students' static characteristics and historical learning behaviors has resulted in the neglect of the dynamic changes in the learning environment and students' time management. Consequently, optimizing the timing of content delivery to accommodate individual learning requirements and contextual variability has emerged as a critical research challenge. To address this issue, a learning content recommendation model tailored for mobile learning environments was proposed in this study. The model comprises three main components: an encoding layer, a Transformer layer, and a prediction layer. The encoding layer combines the graph structure of the mobile learning network with the learning content recommendation problem by encoding student and interactive learning content nodes and corresponding time information. The Transformer layer adjusts the temporal influence of each node and aggregates the embeddings of both nodes and time. The prediction layer leverages the output embeddings from the Transformer layer—infused with temporal features—to perform learning content delivery prediction. Through the construction and optimization of this model, the objective is to enhance the precision and efficiency of content delivery, thereby improving educational quality and learning outcomes in mobile learning environments.

KEYWORDS

mobile learning, intelligent education, learning content recommendation, delivery timing optimization, transformer model, learning environment transition

Wang, J. (2025). Intelligent Education Based on Mobile Learning: Transitioning from Traditional Classrooms to Adaptive Learning Environments. *International Journal of Interactive Mobile Technologies (iJIM)*, 19(11), pp. 51–65. https://doi.org/10.3991/ijim.v19i11.56057

Article submitted 2025-02-15. Revision uploaded 2025-04-16. Final acceptance 2025-04-18.

1 INTRODUCTION

The rapid advancement of information technology, coupled with the widespread proliferation of mobile devices, has brought about a profound transformation in the field of education [1, 2]. A gradual shift has been observed from traditional classroom-based instruction to more flexible and personalized mobile learning paradigms [3–5]. Mobile learning has not only eliminated the conventional constraints of time and space [6], enabling students to engage in learning anytime and anywhere [7], but has also provided new technological means and application scenarios for the advancement of intelligent education [8, 9]. Within this context, the optimization of learning content delivery timing through intelligent technologies—adapted to both individual learning needs and dynamically changing learning environments—has emerged as a critical topic in contemporary educational research.

The optimization of timing for intelligent learning content recommendation has been recognized as a key strategy for enhancing instructional efficiency within mobile learning environments [10, 11]. Accurate and timely content delivery has been shown to improve student engagement and learning efficiency [12], while personalized learning pathways and real-time feedback [13] further support learning under optimal cognitive conditions. Research on intelligent education enabled by mobile learning is therefore of substantial theoretical significance, while also offering practical contributions to educational innovation. Such research supports the promotion of educational equity and quality, providing both scientific grounding and technological support for personalized education and lifelong learning.

Despite notable progress in learning content recommendation and timing optimization, several limitations remain in existing research [10, 14]. For instance, Gan and Ma [15] focused solely on students' static characteristics and historical learning behaviors, overlooking the dynamic nature of learning environments and individual time management. Moreover, most current recommendation systems are based on traditional collaborative filtering or content-based algorithms [16, 17], which are limited in their ability to incorporate complex interaction data and temporal features. These methods often underperform in diverse and rapidly evolving learning scenarios, highlighting an urgent need for further refinement and enhancement.

To address these shortcomings, a learning content recommendation model was proposed in this study that incorporates the dynamics of learning environment transitions and supports timing optimization within mobile learning contexts. The proposed model consists of three primary components. First, the encoding layer encodes student and interactive content nodes and corresponding temporal information, thereby integrating the graph structure of the mobile learning network with the content recommendation task. Second, the Transformer layer adjusts the temporal influence of each node and aggregates both node and time embeddings. Finally, the prediction layer utilizes the output embeddings from the Transformer layer—infused with temporal features—to perform learning content delivery prediction. Through the construction and optimization of this model, more accurate and efficient content delivery is expected to be achieved, ultimately enhancing the quality of education and learning outcomes within mobile learning environments.

2 TIMING OPTIMIZATION FOR LEARNING CONTENT RECOMMENDATION CONSIDERING LEARNING ENVIRONMENT TRANSITIONS

The rise of mobile learning and the advancement of intelligent education have enabled learning to occur beyond the boundaries of traditional classrooms, allowing for anytime, anywhere access. However, significant variations may exist in students' cognitive states and learning needs across different times and environments. Therefore, the dynamic adjustment of learning content delivery timing—ensuring alignment with actual learning conditions and individual states—has become a critical challenge in the pursuit of intelligent education. Optimizing the timing of content recommendation not only improves learning efficiency but also enhances student engagement and proactivity, thereby contributing to overall improvements in learning outcomes.

To address this challenge, a learning content recommendation model was developed in this study that considers learning environment transitions by integrating a continuous-time transformer (ContiFormer) with a mobile interaction-based graph neural network (GNN). By incorporating a Transformer model enriched with temporal information, the temporal dimension was fully leveraged to capture behavioral patterns and fluctuations in learning states across different time points. In order to further improve the model's adaptability to temporal dynamics, a temporal transformation graph was introduced in this study to model students' state transitions over time and the influence of these transitions on learning content recommendations. Moreover, to enhance both computational efficiency and prediction accuracy, residual network structures and downsampling modules were incorporated. These architectural components not only optimize computational performance but also strengthen the model's ability to capture complex behavioral patterns.

2.1 Embedding layer

Two types of embeddings were incorporated into the model: long-term embeddings and continuous-time embeddings. Long-term embeddings can capture the static characteristics of both students and learning content, including individual learning preferences, historical learning behavior, and the fundamental attributes of the learning materials. In contrast, continuous-time embeddings account for the dynamic influence of temporal factors on both student behavior and learning content features. By modeling variations along the temporal dimension, a more precise understanding of students' evolving learning needs across different time spans can be achieved, thereby enabling more accurate timing in learning content recommendation.

a) Long-term student/interactive content embeddings: Long-term embeddings were introduced into the model framework to parameterize the feature representations of students and learning content as low-dimensional vectors. These vectors are capable of capturing rich semantic information and inter-relational patterns. Within the time-aware Transformer model, these embeddings serve as node features and were optimized to model the overall structural information. The embeddings of student and learning content nodes r_i and r_u denote the personalized learning preferences of students and the attributes of learning content, respectively. Initially, these embeddings were stored in a global embedding table R. During training, they were continuously updated using backpropagation and gradient descent optimization algorithms to reflect the most recent learning behaviors and content features.

b) Continuous-time embeddings: Continuous-time embeddings can map temporal information into a high-dimensional space, enabling the representation of temporal spans through dot products of time embeddings. Specifically, temporal influence is defined as a function $\psi(s_1 - s_2)$, where s_1 and s_2 denote distinct timestamps at which learning behaviors occur. In this formulation, temporal influence is directly modeled via the dot product of corresponding time embeddings, allowing the model to capture the impact of time spans on learning behaviors, rather than relying solely on absolute temporal differences. The introduction of continuous-time embeddings enables the model to encode temporal spans through time embeddings and to infer future learning behaviors accordingly. Let S represent the time kernel and $\cdot$ denote the dot product operation, and it leads to:

$$\psi(s_1 - s_2) = S(s_1, s_2) = \varphi(s_1) \cdot \varphi(s_2) \tag{1}$$

Based on Bochner's theorem, temporal influence between any pair of time points can be captured via the dot product of their respective time embeddings. This formulation allows the temporal representation to be applied flexibly at arbitrary time points, enabling the model to adapt to dynamic temporal variations across diverse learning environments. Through this mechanism, both the evolution of student learning behaviors over time and the prediction of future learning needs and optimal content delivery timing can be effectively supported. To explicitly indicate temporal characteristics, the learnable parameter was defined as $\mu = \left[\mu_1, \ldots, \mu_{f_S}\right]^T$, where f_S denotes the dimension. The continuous-time embedding is then formulated as:

$$\Theta(s) \rightarrow \sqrt{\frac{1}{f_S}} z\left[COS(\mu_1 s), SIN(\mu_1 s), \ldots, COS(\mu_{f_S} s), SIN(\mu_{f_S} s)\right]^S \tag{2}$$

2.2 Transformer layer

The transformer layer with integrated temporal information was designed to unify student/content embeddings with time embeddings. This unified approach allows time information to participate in the model's calculations along with other features, thereby capturing a more comprehensive representation of students' evolving learning behaviors. By incorporating temporal influence into the Transformer layer, patterns in behavioral changes over time can be effectively identified. For instance, as students transition from classroom-based instruction to adaptive learning environments, their learning behaviors and content needs are likely to shift. Through the integration of time embeddings, the model is enabled to consider such temporal transitions, thereby supporting more personalized and timely content recommendations. Moreover, the introduction of the time factor into the attention module allows for dynamic weighting across different time points, enabling the model to identify the most salient learning content and interactive behaviors at specific moments. As a result, the influence of temporal dynamics on student behavior can be more precisely captured, enhancing the impact of content delivery. For example, learning preferences that exhibit temporal specificity—such as increased engagement with certain content types during particular periods—can be captured and used to guide content recommendation, thereby improving both learning outcomes and student satisfaction. In addition, by aggregating student data to compute interaction importance scores, the model's adaptability and flexibility in complex learning environments can be further enhanced. A schematic illustration of the Transformer layer's input processing is presented in Figure 1.

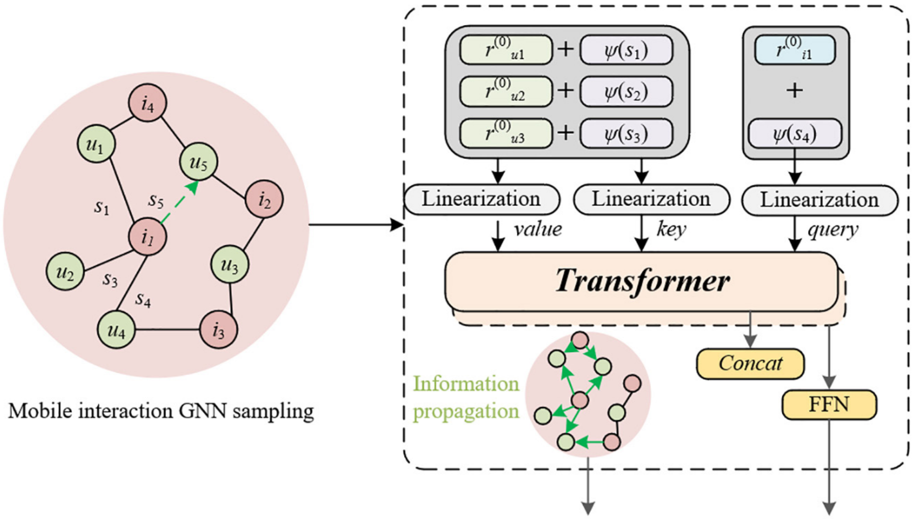

Fig. 1. Schematic representation of input processing in the Transformer layer

The core of information construction within the Transformer layer lies in the integration of long-term node embeddings and time embeddings as the input. In this context, long-term node embeddings represent persistent features of students or learning content over the entire learning process, whereas time embeddings capture the dynamic characteristics associated with specific time points. By combining these two types of embeddings, the model is enabled to jointly account for both long-term learning behaviors and present learning demands. This dual consideration allows each Transformer layer to process inputs in a manner that captures both temporal dynamics and long-term behavioral patterns, thus enabling more precise and adaptive content recommendations. For example, when a student transitions from a traditional classroom to an adaptive learning environment, substantial changes may occur in behavioral patterns and content needs. The integration of time embeddings allows the model to detect such changes in real time and adjust recommendation strategies accordingly. Let $m = 1, 2, \ldots, M$. In addition, $g_i^{(m-1)}(s)$ represents the information of student i at time s in layer m–1; $r_i^{(m-1)}(s) \in R^f$ denotes the time embedding of student i; $\Theta(s)$ represents the time embedding; and || denotes the concatenation operator. The query input for student i at time s in layer m is then defined as:

$$g_i^{(m-1)}(s) = r_i^{(m-1)}(s) \,||\, \Theta(s) \tag{3}$$

In addition to node processing, information construction also propagates information from students' neighboring nodes through the graph structure. In this graph, nodes correspond to students or learning content, and edges represent associations or interactive relationships. A GNN framework was employed to facilitate the propagation of information across nodes, thereby enabling feature sharing and integration. This allows student learning behaviors to be influenced not only by their individual characteristics and time-specific factors but also by peers who have similar learning behaviors or interests. Through the graph-based propagation mechanism, information from neighboring nodes can be aggregated to improve understanding and predictive accuracy for each target node. In practical terms, this means that when a student transitions from classroom-based learning to an adaptive learning environment, their prior classroom behavior—as well as the behavior of their peers—can be utilized by the model to support the recommendation of more personalized and timely learning content. $T_u, V_i(s) = \{(u, s_t) |\ (i, u, s_t) \in R_s WI\ s_t < s\}$ that interact with i before time s were randomly sampled. For each pair (u, s_t), the input information at layer m is defined as:

$$g_u^{(m-1)}(s_t) = r_u^{(m-1)}(s_t) \,||\, \Theta(s_t) \tag{4}$$

Let $g_u(s_t)$ represent the information associated with learning content u at time s_t, and let $r_u(s_t)$ denote its corresponding time representation. Specifically, when $m = 1$, $r_u^{(0)}(s_t)$ is equivalent to R_u, which denotes the long-term item embedding. For $m > 1$, the embedding is derived from the output of the preceding Transformer layer with temporal integration.

To facilitate information propagation, the model needs to calculate the linear aggregation of all sampled interaction source information. During this process, each student's neighbor nodes $V_i(s)$ represent other students or learning content that have interacted with the target student i at time s. By sampling from these neighboring nodes, the model is able to gather diverse collaborative information and temporal dynamics. A weighted summation was then computed over these sampled neighbors to generate a unified representation through linear aggregation. This step not only enables the effective fusion of information from different neighbors but also emphasizes those interaction sources that exert greater influence on the target student at the current time point through a weighting mechanism. For instance, frequent classroom interactions with specific peers or content may continue to exert significant influence as students transition into adaptive learning environments. Through linear aggregation, such critical interaction data can be effectively integrated by the model to construct a temporally adaptive embedding. The linear aggregation is defined as:

$$r_{V_i}^{(m)}(s) = \sum\nolimits_{(u,s_t)\in V_i(s)} \tau_s^i(u,s_t) Q_n^{(m)} g_u^{(m-1)}(s_t) \tag{5}$$

Traditional attention mechanisms typically emphasize the relevance between interactive learning content while overlooking the role of temporal information. To address this limitation, a temporal attention mechanism was introduced in this study, which simultaneously considers both neighbor interactions and the temporal attributes of interaction edges, thereby enhancing performance in the temporal dimension. By learning weights $\tau_s^i(u,s_t)$, this attention mechanism combines the degree of association between students and interactive learning content and temporal information. The weights $\tau_s^i(u,s_t)$ consider the direct interaction between students and the learning content while incorporating the corresponding time information. This is critical for assessing the significance of historical learning interactions, as students' learning behavior and content preferences may change over time. For example, a student's level of interest in a particular topic may fluctuate across different time periods. By applying the temporal attention mechanism, the model dynamically adjusts content recommendations to better align with evolving learning habits and needs. Assuming the linear transformation matrices are given by $Q_j^{(m)}$ and $Q_w^{(m)}$, the temporal attention weights $\tau_s^i(u,s_t)$ can be calculated as:

$$\tau_s^i(u,s_t) = \frac{1}{\sqrt{f + f_S}} \left(Q_j^{(m)} g_u^{(m-1)}(s_t) \right)^S Q_w^{(m)} g_i^{(m-1)}(s) \tag{6}$$

By removing the scalar factor and transformation matrices, the right-hand side of the previous expression can be equivalently written as:

$$r_i^{(m-1)}(s)\cdot r_u^{(m-1)}(s_t)+\Theta(s)\cdot\Theta(s_t) \tag{7}$$

where the first term represents the degree of association between the student and the interactive learning content, reflecting the student's interest and engagement level with the specific content, and the second term models the temporal information, capturing how the student's learning behavior evolves over time. Through a multi-level attention mechanism, the integration of temporal information and collaborative signals is progressively reinforced across layers, thereby enabling the model to more precisely capture the temporal dynamics of student behavior. This hierarchical temporal attention mechanism enhances the model's ability to interpret and predict students' learning behavior within complex temporal sequences. Subsequently, the attention weights of the sampled interactions were normalized using a softmax function:

$$\tau_s^i(u,s_t)=\frac{\exp\left(\tau_s^i(u,s_t)\right)}{\sum_{U(u',s_T)\in V_i(s)}\exp\left(\tau_s^i(u',s_t')\right)} \tag{8}$$

Additionally, the specific attention values can be computed by aggregating all sampled interaction information. More specifically, all sampled interaction data were compacted into a matrix $G^{(m-1)}V_i(s)\in R^{(f+f_S)\times T}$, which includes the interaction information of all neighboring nodes across different time points. A linear transformation was then applied to compute the key, value, and query matrices, defined respectively as $J_i^{(m-1)}(s)=Q_j^{(m)}G_{Vi}^{(m-1)}(s)$, $N_i^{(m-1)}(s)=Q_j^{(m)}G_{Vi}^{(m-1)}(s)$, and $w_i^{(m-1)}(s)=Q_j^{(m)}g^{(m-1)}(s)$. These matrices were utilized within the attention module, where attention weights were calculated through dot product operations, and the weighted interaction information was obtained. This procedure ensures that all sampled temporal sequence data are comprehensively considered, thereby supporting more accurate learning content recommendations in the temporal dimension. For clarity of presentation, by omitting superscripts and the time *t* and by combining Equations (6) and (8), Equation (5) can be rewritten as:

$$r^{V_i}=N_i\cdot soft\max\left(\frac{J_i^S w_i}{\sqrt{f+f_S}}\right) \tag{9}$$

A dot-product attention mechanism was employed within the Transformer architecture, and multi-head attention was adopted to concatenate the outputs of each head for aggregation. Unlike conventional self-attention, the attention mechanism in this framework is time-aware and specifically designed to model both the interactions between students and learning content and the temporal context of those interactions. The multi-head attention mechanism enables the model to capture diverse temporal dynamics across different attention heads, thereby facilitating a more comprehensive understanding of student learning behaviors. This design allows the model to better adapt to the transition from traditional classroom-based learning to adaptive learning environments and optimize the timing of learning content delivery. For instance, when students shift from classroom instruction to an adaptive learning setting, the model can dynamically adjust both the sequence and timing of content recommendations based on their learning behaviors at different time periods, thereby improving the accuracy and effectiveness of personalized recommendations.

To further enhance its adaptability to the classroom-to-adaptive learning transition, the Transformer layer captures diverse temporal behavioral features through the multi-head attention mechanism. The resulting weighted interaction information was concatenated with the query information and processed through a feedforward neural network (FFN). The FFN consists of two linear transformation layers, with one layer employing a ReLU activation function to introduce non-linearity, enabling the model to better capture complex learning behavior patterns. Through this structure, the model is capable of fine-grained temporal modeling of learning behaviors, ultimately generating time-aware final embeddings for further learning content recommendation. Let $r_i^{(m)}(s)$ denote the time embedding of student i in layer m; then the representation is given by:

$$r_i^{(m)}(s) = FFN\left(r_{Vi}^{(m)}(s) \,||\, g_i^{(m-1)}(s)\right) \tag{10}$$

2.3 Model prediction

The prediction mechanism of the proposed learning content recommendation model is primarily based on the stacked architecture of Transformer layers enhanced with temporal dynamics and accurately capturing fine-grained interaction features between students and learning content along the time dimension. By stacking M time-aware Transformer layers, final time-specific embeddings $r_i^{(M)}(s)$ and $r_u^{(M)}(s)$ were generated for students and learning content at specific time points, forming a solid foundation for learning content recommendation. The overall model architecture is illustrated in Figure 2.

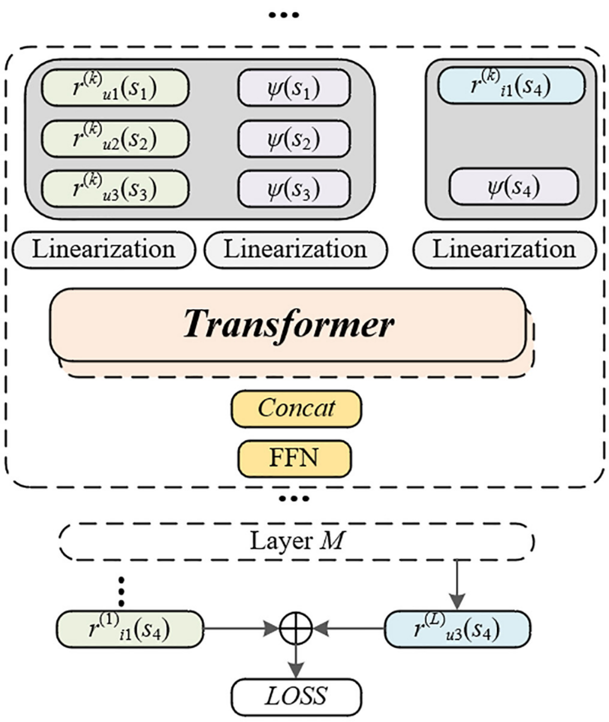

Fig. 2. Overall architecture of the learning content recommendation model

Initially, the model processes the interaction information between student i and learning content u at time s using L Transformer layers that incorporate temporal information. Each Transformer layer employs a temporal attention mechanism to perform weighted aggregation over interaction information from different time points, producing an embedding that reflects the current time point. After multiple stacked layers, the final output embeddings $r_i^{(M)}(s)$ and $r_u^{(M)}(s)$ were obtained, representing the ultimate dynamic features of the student and the learning content at time t. This capturing of temporal dynamics enables not only the prediction of recommended content at the current time point but also the inference of student–content embeddings at any arbitrary timestamp, thus supporting comprehensive optimization of learning content recommendation.

Subsequently, the final embeddings were utilized to compute the recommendation scores for all interactive learning content. These scores were computed based on the inner product or other similarity metrics between the embeddings of the student and the learning content. Let $e(i, u, s)$ denote the recommendation score for content u being recommended to student i at time s; then the computation is defined as:

$$e(i,u,s) = r_i^{(M)}(s) \cdot r_u^{(M)}(s) \tag{11}$$

By computing recommendation scores for all candidate learning content and ranking them accordingly, a recommendation list can be generated to optimize the timing of content delivery. This approach enables the model to account for the temporal dynamics of learning content while dynamically adjusting recommendation strategies to better align with individual learning progress and needs.

To ensure end-to-end optimization, all components of the model were designed to be differentiable, allowing the entire architecture to be trained using the Adam optimization algorithm. The Adam optimizer, which incorporates both first-order and second-order gradient information, adaptively adjusts the learning rate during training, thereby improving training efficiency and stability. Furthermore, the model was optimized by a pairwise Bayesian Personalized Ranking (BPR) loss function. The BPR function loss is suitable for implicit feedback scenarios, optimizing the performance of top-N recommendations by maximizing the difference in predicted scores between observed and unobserved interactions. Let P_S represent the training samples, $\delta(\cdot)$ denote the sigmoid function, and φ represent the learnable parameters. The BPR loss function is defined as:

$$M_{yoe} = \sum_{(i,u,k,s)\in P_S} -\log\delta(e(i,u,s) - e(i,k,s)) + \eta \left\|\varphi\right\|_2^2 \tag{12}$$

Finally, model parameters were continuously adjusted through backpropagation during the training process to minimize the BPR loss function, thus optimizing the learning content recommendation effect. This end-to-end optimization framework enables the model to dynamically adapt to changing learning environments and to deliver the most appropriate content at the most effective time, thereby improving students' learning efficiency and effectiveness. Throughout the transition from traditional classroom-based instruction to adaptive learning environments, the model is capable of adjusting its recommendation strategy by leveraging temporal dynamics and actual student needs. This ensures that the timing of content delivery is optimized, ultimately enabling a more personalized and efficient learning experience.

3 RESULTS AND DISCUSSION

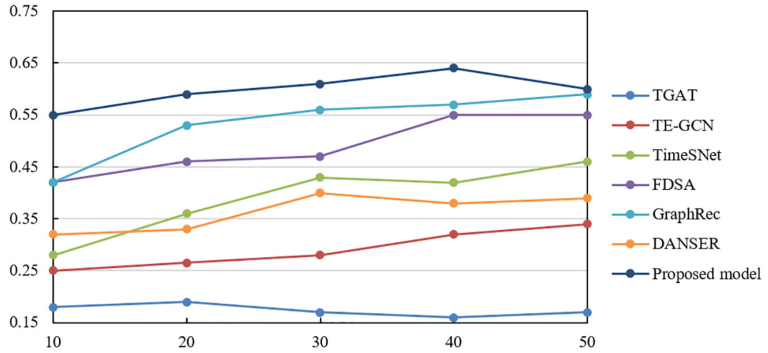

Fig. 3. NDCG@10 performance on the training set

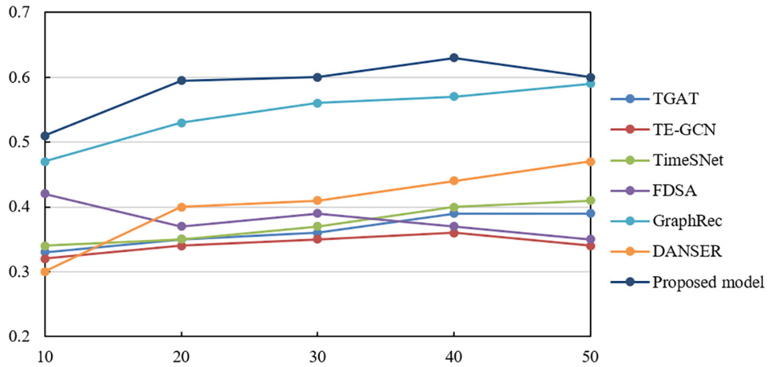

Fig. 4. NDCG@10 performance on the validation set

As observed in the data presented in Figures 3 and 4, the proposed model consistently outperformed other models on both the training and validation sets. On the training set, the NDCG@10 values achieved by the proposed model were 0.55, 0.59, 0.61, 0.64, and 0.62 across varying training set sizes, exhibiting a clear upward trend. These results were significantly higher than those of other baseline models. In particular, at a sample size of 50, the model attained the highest NDCG@10 value of 0.62, surpassing GraphRec (0.59), FDSA (0.55), and DANSER (0.39). The model also demonstrated strong performance on the validation set, with NDCG@10 values recorded at 0.51, 0.595, 0.6, 0.63, and 0.6. At the 50-sample mark, the proposed model continued to outperform competitors, with notable performance gaps observed between the proposed model and others such as GraphRec (0.59) and FDSA (0.35). These findings indicate that the proposed learning content recommendation model—designed for mobile learning environments—significantly improved the timing optimization of content delivery. The integration of student and interactive learning content nodes and temporal information within the model architecture was shown to enable more accurate prediction of learner needs, thereby enhancing recommendation system efficiency. In particular, the consistent increase in NDCG@10 scores on both the training and validation sets highlights the model's advantage in handling sequential data and delivering personalized recommendations. Compared with baseline methods, the proposed model exhibited a superior capacity to adjust for the influence of temporal factors in content recommendation, resulting in more precise delivery of learning content. These results underscore the model's strong practical application value and potential.

The ablation analysis presented in Table 1 reveals the contribution of each model component to Recall@10. The full model consistently achieved superior performance across all five datasets, with particularly notable results on the MovieLens and SAINT+ datasets, where Recall@10 values reached 0.5842 and 0.3652, respectively—clearly outperforming all alternative configurations. When the long-term embedding was removed, a substantial decline in Recall@10 was observed across all datasets, with the most pronounced drop of 0.0634 recorded on the Mobile Learning Effectiveness dataset. Similarly, the exclusion of the continuous-time embedding and self-attention mechanism resulted in noticeable performance degradation. For example, removing the continuous-time embedding led to decreases of 0.0985 and 0.1125 on the UKP-SQuAD and PSLC DataShop datasets, respectively. Further reductions in performance were recorded when the GNN and BPR loss functions were excluded. The absence of the GNN caused the most significant drop in the PSLC DataShop dataset (0.0224↓). These ablation results underscore the critical role of each component in enhancing model performance, particularly the long-term embedding, continuous-time embedding, and self-attention mechanism. The elimination of these elements resulted in considerable degradation in Recall@10, indicating that they are essential for capturing learners' long-term interests, time series information, and relationships between nodes. In addition, the removal of GNN and the BPR loss function also led to a significant decrease in performance, indicating that graph structures can effectively capture the relationship between learners and content when processing learning content recommendations, while the BPR loss function helps optimize the accuracy of recommendations. These results fully demonstrate the rationality of the model design and the practical value of each module in optimizing the timing of learning content recommendation.

Table 1. Ablation analysis on five datasets (Recall@10)

	UKP-SQuAD	PSLC DataShop	MovieLens	SAINT+	Mobile Learning Effectiveness Dataset
Full model	0.3562	0.2125	0.3521	0.5842	0.3215
Removal of the long-term embedding	0.0026↓	0.02236↓	0.0054↓	0.0052↓	0.0634↓
Removal of the continuous-time embedding	0.0985↓	0.1125	0.1125↓	0.3652	0.3125
Removal of the self-attention mechanism	0.0745↓	0.0936↓	0.0923↓	0.3512	0.2635
Removal of GNN	0.0378↓	0.0224↓	0.0215↓	0.0735↓	0.0874↓
Removal of the BPR loss function	0.0123↓	0.0235↓	0.0017↓	0.0336↓	0.0612↓

Table 2. Variants in temporal information construction

	UKP-SQuAD	PSLC DataShop	MovieLens	SAINT+	Mobile Learning Effectiveness Dataset
Proposed model	0.3526	0.2136	0.35262	0.5841	0.3215
Without temporal information (student)	0.0112	0.0124	0.0114	0.0125	0.1426
Without temporal information (learning content)	0.1124	0.09562	0.0826	0.2635	0.2236

The data presented in Table 2 demonstrate the critical role of temporal information construction in shaping the performance of the model. The full model consistently achieved high Recall@10 scores across all five datasets, particularly on the MovieLens and SAINT+ datasets, where scores reached 0.5841 and 0.3215, respectively—indicating strong recommendation performance. However, a substantial decline in Recall@10 scores was observed across all five datasets when student-specific temporal information was excluded. This was especially pronounced on the Mobile Learning Effectiveness Dataset, where Recall@10 decreased by 0.1426, underscoring the essential role of students' temporal information in ensuring recommendation accuracy. Similarly, the exclusion of temporal information associated with learning content also led to performance degradation. Although the impact was less severe, decreases were still evident, particularly on the UKP-SQuAD and PSLC DataShop datasets. These findings affirm that the construction of temporal information—especially for both students and learning content—is a key factor in improving the precision of learning content delivery. The ablation analysis of temporal information variants confirms that the temporal information of both students and learning content must be considered when addressing learning content recommendation tasks. The absence of either type of temporal information significantly impairs recommendation performance, with the lack of student temporal information having the most substantial impact on both accuracy and personalization. By effectively integrating the temporal characteristics of both students and learning content, the full model demonstrated an enhanced ability to capture learners' evolving needs, thereby enabling more precise content delivery. Although the exclusion of content time information had a relatively smaller impact, its effect still highlighted the broader importance of temporal modeling in personalized recommendation systems.

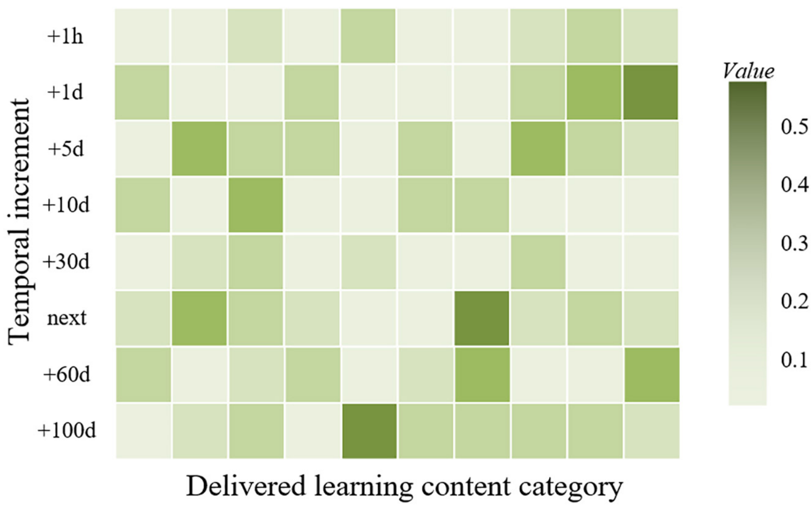

Fig. 5. Visualization of temporal attention weights

Significant variation in delivery adaptability across different learning content categories at various time points was observed when analyzed from the perspective of temporal increments, as shown in Figure 5. For instance, the category "mobile fragmented learning" exhibited elevated values under short-term temporal increments such as "+1h" and "+1d," aligning with user preferences in mobile learning environments for timely and lightweight content. In contrast, "hands-on practice courses" demonstrated a pronounced peak at the "next" time point, indicating a strong suitability for real-time delivery based on the current learning scenario. Moreover, categories such as "professional qualification certification" and "immersive virtual learning" were found to exhibit higher values under longer-term temporal increments, including "+30d" and "+100d." This pattern suggests that more systematic, long-duration learning content is better suited for planned, future-oriented content delivery strategies. This data distribution reflects the strong correlation between learning content and temporal scenarios, verifying the model's ability to capture temporal information in the study.

4 CONCLUSION

A learning content recommendation model tailored to mobile learning environments was proposed to optimize delivery timing and enhance educational quality and learning outcomes. The model architecture comprises three core components: an encoding layer, a Transformer layer, and a prediction layer. The encoding layer was designed to encode student and interactive content nodes and temporal information within a graph-based structure to address the learning content recommendation task. The Transformer layer adjusts inter-node relationships based on temporal influence and aggregates node and time embeddings. Finally, the prediction layer utilizes the temporally integrated embeddings to perform content delivery prediction. Experimental results demonstrated that the proposed model consistently outperformed baseline approaches across multiple datasets, particularly in terms of recommendation accuracy. Ablation studies further validated the effectiveness of each model component. Notably, the long-term embedding, temporal information, and GNN were identified as critical contributors to recommendation performance.

The study presents both theoretical and practical contributions to the field of learning content delivery timing optimization in mobile learning environments. By incorporating temporal information, individualized student data, and the graph structure into the learning content recommendation model, this study provides new ideas for personalized education recommendations, which can significantly improve the accuracy and efficiency of recommendations. However, this study has certain limitations. The model's performance is highly dependent on temporal information, and its generalization capability may be restricted on smaller-scale datasets. Additionally, although historical behavior and the time factor have been effectively leveraged, deeper exploration of learner-specific attributes remains an area for improvement. Future research should aim to enhance generalization performance, explore more granular learner personalization features, and integrate multimodal learning content, further advancing the development of intelligent education toward greater personalization and efficiency.

5 ACKNOWLEDGEMENT

This paper was funded by the Anhui Provincial Higher Education Quality Engineering Project, "Current Status and Improvement Strategies for Blended Teaching Management in Application-Oriented Undergraduate Institutions Under the New Educational Development Context" (Project No.: 2022jyxm526).

6 REFERENCES

[1] A. Rodríguez-Ríos *et al.*, "Information and communication technology, mobile devices, and medical education," *Journal of Medical Systems*, vol. 44, p. 90, 2020. https://doi.org/10.1007/s10916-020-01559-w

[2] Y. Zhang, "Machine learning-based personalized learning path decision-making method on intelligent education platforms," *International Journal of Interactive Mobile Technologies (iJIM),* vol. 18, no. 16, pp. 68–82, 2024. https://doi.org/10.3991/ijim.v18i16.51009

[3] L. F. Chu, M. J. Erlendson, J. S. Sun, H. L. Alva, and A. M. Clemenson, "Mobile computing in medical education: Opportunities and challenges," *Current Opinion in Anesthesiology*, vol. 25, no. 6, pp. 699–718, 2012. https://doi.org/10.1097/ACO.0b013e32835a25f1

[4] F. D. C. Velázquez and G. M. Méndez, "Augmented reality and mobile devices: A binominal methodological resource for inclusive education (SDG 4). An example in secondary education," *Sustainability*, vol. 10, no. 10, p. 3446, 2018. https://doi.org/10.3390/su10103446

[5] K. Baba, N.-E. El Faddouli, and N. Cheimanoff, "Mobile-optimized AI-driven personalized learning: A case study at Mohammed VI polytechnic university," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 4, pp. 81–96, 2024. https://doi.org/10.3991/ijim.v18i04.46547

[6] T. D. Aungst and P. Belliveau, "Leveraging mobile smart devices to improve interprofessional communications in inpatient practice setting: A literature review," *Journal of Interprofessional Care*, vol. 29, no. 6, pp. 570–578, 2015. https://doi.org/10.3109/13561820.2015.1049339

[7] I. Jahnke and J. Liebscher, "Three types of integrated course designs for using mobile technologies to support creativity in higher education," *Computers & Education*, vol. 146, p. 103782, 2020. https://doi.org/10.1016/j.compedu.2019.103782

[8] E. Baran, "A review of research on mobile learning in teacher education," *Journal of Educational Technology & Society*, vol. 17, no. 4, pp. 17–32, 2014.

[9] A. B. Feroz Khan and S. R. A. Samad, "Evaluating online learning adaptability in students using machine learning-based techniques: A novel analytical approach," *Education Science and Management*, vol. 2, no. 1, pp. 25–34, 2024. https://doi.org/10.56578/esm020103

[10] L. G. Morrison *et al.*, "The effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: An exploratory trial," *PLOS One*, vol. 13, no. 5, p. e0169162, 2018. https://doi.org/10.1371/journal.pone.0169162

[11] N. Nouman, Z. A. Shaikh, and S. Wasi, "A novel personalized learning framework with interactive e-Mentoring," *IEEE Access*, vol. 12, pp. 10428–10458, 2024. https://doi.org/10.1109/ACCESS.2024.3354167

[12] V. B. Semwal, K. Mondal, and G. C. Nandi, "Robust and accurate feature selection for humanoid push recovery and classification: Deep learning approach," *Neural Computing and Applications*, vol. 28, pp. 565–574, 2017. https://doi.org/10.1007/s00521-015-2089-3

[13] K. M. Briedis, A. Djelouah, M. Meyer, I. McGonigal, M. Gross, and C. Schroers, "Neural frame interpolation for rendered content," *ACM Transactions on Graphics*, vol. 40, no. 6, pp. 1–13, 2021. https://doi.org/10.1145/3478513.3480553

[14] D. K. Krishnappa, M. Zink, C. Griwodz, and P. Halvorsen, "Cache-centric video recommendation: An approach to improve the efficiency of YouTube caches," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 11, no. 4, pp. 1–20, 2015. https://doi.org/10.1145/2716310

[15] M. Gan and Y. Ma, "DeepInteract: Multi-view features interactive learning for sequential recommendation," *Expert Systems with Applications*, vol. 204, p. 117305, 2022. https://doi.org/10.1016/j.eswa.2022.117305

[16] Y. H. Lee, C. P. Wei, P. J. H. Hu, T. H. Cheng, and C. W. Lan, "Small Clues Tell: A collaborative expansion approach for effective content-based recommendations," *Journal of Organizational Computing and Electronic Commerce*, vol. 30, no. 2, pp. 111–128, 2020. https://doi.org/10.1080/10919392.2020.1718056

[17] G. Lekakos and P. Caravelas, "A hybrid approach for movie recommendation," *Multimedia Tools and Applications*, vol. 36, pp. 55–70, 2008. https://doi.org/10.1007/s11042-006-0082-7

7 AUTHOR

Jinqiu Wang is currently serving as an Associate Professor at Anhui Sanlian University, holds a Master's degree in Landscape Plants and Ornamental Horticulture from Anhui Agricultural University. Her academic expertise encompasses Design Studies and Educational Management. With a robust research portfolio, she has published 5 scholarly papers and contributed to 1 provincial-level research project. Her pedagogical accomplishments include receiving 2 Provincial-Level Teaching Achievement Awards and completing 5 teaching-research initiatives. Notably, she has mentored students in disciplinary competitions, with her mentees attaining the distinction of provincial second prize at the highest level (E-mail: luyimiao0@126.com).