

Multi-type MOOCs Recommendation: Leveraging Deep Multi-Relational Representation and Hierarchical Reasoning

Ye Zhang¹, Yanqi Gao¹, Dongjie Wang^{2*}, Yupeng Zhou¹, Jinlong He¹, Zhaoyang Sun¹, Minghao Yin^{1*}

¹ College of Information Science and Technology, Northeast Normal University

² Electrical Engineering and Computer Science Department, University of Kansas

zhangy923@nenu.edu.cn, gaoyq557@nenu.edu.cn, wangdongjie@ku.edu, zhouyp605@nenu.edu.cn, hejl104@nenu.edu.cn, suncy355@nenu.edu.cn, ymh@nenu.edu.cn

Abstract

Massive open online courses (MOOCs) recommendation provides online courses tailored to learners' individual preferences. Existing literature is limited by: 1) Ignoring the interrelations among courses, knowledge concepts, and videos, which leads to suboptimal recommendation performance; 2) Neglecting the hierarchical interactions between learners and components like courses, knowledge concepts, and videos, which makes it difficult to capture learners' intentions accurately. To address them, we propose a novel multi-type MOOCs recommendation framework, which enables multi-type educational content recommendations. This framework includes two important components: multi-relational representation and hierarchical reasoning. Regarding multi-relational representation, we first create two static course-relational and knowledge concept-relational graphs based on domain knowledge and construct a dynamic video-relational graph using learners' browsing historical sequences. Then, we capture the interactions among different components by learning the corresponding embeddings via graph neural networks. Regarding hierarchical reasoning, we implement a hierarchical beam search strategy to narrow down the candidate courses, knowledge concepts, and videos by calculating joint probability. Finally, we introduce an optional layer to increase the diversity and reasonableness of video recommendations by estimating learners' intentions. Extensive experiments are conducted to show the effectiveness, robustness, and interpretability of our method.

Introduction

MOOCs platforms (e.g. Coursera, edX, and Knewton) have revolutionized the accessibility of educational areas, making high-quality learning resources available to global audiences and encouraging lifelong learning. Customized and multi-type recommendations can significantly enhance learners' interest, improve learning efficiency, and broaden the diversity of their learning experiences.

Existing MOOCs recommendations are typically classified into two categories: 1) Content-based recommendation, which utilizes sequential recommendation models to analyze learners' preferences and recommends educational content (e.g. courses, knowledge concepts, or videos) based

on historical learning records (Tian and Liu 2021; Wang et al. 2022; Zhang et al. 2023; Gong et al. 2023; Wen et al. 2023). 2) Meta-path-based recommendation, which often constructs a knowledge graph to indicate the relationship between learners and video contents to produce explainable and logical recommendation results (Jiang et al. 2023; Li et al. 2023; Alatrash et al. 2024; Liang et al. 2023).

However, prior studies are not sufficient. Content-based methods typically focus solely on users' preferences and their historical activities. They often tend to offer recommendations that are narrowly aligned with past interactions and stated preferences (Wang et al. 2021). Although these methods are personalized, they lack depth in addressing the comprehensive needs of learners. Thus, the effectiveness of these recommendations is suboptimal. Meta-path-based methods tend to overlook the multi-type interactions between learners and various educational components such as courses, knowledge concepts, and videos. This oversight makes it difficult for these methods to precisely capture learners' intentions and preferences (Wang et al. 2023a). These studies typically focus on recommendations for single-type of educational content, neglecting the interconnectivity among various aspects of learning.

Thus, we explore multi-type MOOCs recommendation, focusing on accurately capturing learners' intentions. There are two research challenges: 1) **How to represent and capture the relationships among course, knowledge concept, and video during recommendation?** In the educational domain, the three components are interrelated. Effectively recommending any one of them can enhance our comprehension of learners' interests, thereby improving the performance of recommendations in the other two areas. 2) **How to grasp the sparse, multi-type interaction data between learners and various educational contents to produce more explainable results?** Courses, knowledge concepts, and videos exhibit hierarchical structural relationships. By reasoning from higher-level to lower-level elements, the accuracy of the recommendations and explainability can be significantly enhanced.

To address these challenges, we propose a Multi-Type MOOCs Recommendation framework (**TOME**) including multi-relational representation and hierarchical reasoning. **In the multi-relational representation stage**, we first con-

*Corresponding author

struct course-relational and knowledge concept-relational static graphs based on domain knowledge respectively. Then, we create a dynamic video-relational graph using learners’ historical browsing records. Moreover, we employ graph neural networks to preserve the information of the three graphs into three distinct embeddings. The learned embeddings are integrated through a multi-relation adaptive LSTM network, which adaptively adjusts weights across multiple relations and also incorporates the temporal information of learner behaviors. **In the hierarchical reasoning stage**, we develop a hierarchical beam search strategy to consider the multi-level relationships among courses, knowledge concepts, and videos. More specifically, we should select the top- β candidate items at each level. The selection probability at the lower level is influenced by the selections made at the higher level, which involves calculating the joint probability of items from preceding levels as the predictive probability. By integrating the hierarchical relationships among these levels, we effectively capture the complex interactions between learners and diverse educational content. This approach significantly enhances both the explainability and accuracy of our recommendations. Moreover, in the MOOCs recommendation domain, learners often watch new videos to clarify unclear questions. Thus, we develop an optional layer with personalized and diversified modes to enhance the accuracy and reasonableness of video recommendations. If the recommended video is watched by learners, we activate the personalization mode to analyze its probability of being rewatched. Otherwise, we opt for the diversification mode to recommend unexplored items.

Our main contributions are: 1) We propose a two-stage multi-type MOOCs recommendation framework including multi-relational representation and hierarchical reasoning. 2) We use graph neural networks and a multi-relation adaptive LSTM to capture and integrate the interrelations among courses, knowledge concepts, and videos. 3) We introduce a hierarchical beam search strategy to grasp the interactions between learners and educational content to provide more explainable and accurate recommendations. 4) We develop an optional layer to balance exploitation and exploration in video recommendations to enhance recommendation diversity and aid learners in exploring new knowledge areas. 5) We introduce a new evaluation metric for knowledge concept recommendation by considering multiple relevant concepts to objectively assess learning behaviors. 6) We conduct extensive experiments on real-world datasets to demonstrate the effectiveness and generalization of our framework.

Definitions and Problem Statement

Problem statement

Suppose that there is a learner u , along with $|V|$ MOOC videos, $|K|$ knowledge concepts and $|C|$ courses. Let v_t is the video watching at time t for the learner. Furthermore, each video establishes unique relationships with both knowledge concepts and courses. Our task is to predict the next video that the learner will watch, the potential course they may select, and the knowledge concepts he is likely to explore based on his watching record, course-video correlations, and knowledge concept-video correlations.

Definition 1 Course-relational Graph \mathcal{G}_C is defined with the nodes of videos, and the edges indicating if videos belong to the same course.

Definition 2 Knowledge Concept-relational Graph \mathcal{G}_K consists of video nodes, with edges reflecting the Jaccard similarity of knowledge concepts covered by different videos.

Definition 3 Video-relational Graph is \mathcal{G}_V^T evolves dynamically over time, with the nodes of videos, and the edges representing interactions among videos within the time period T , extending from $t - l$ to t .

Method

Framework overview

Figure 1 shows the overview of TOME, which includes two stages: multi-relational representation and hierarchical reasoning. In the first stage, we first input the learner’s current learning content, including courses, knowledge concepts, and videos into an embedding layer to get the corresponding embeddings. After that, we construct a course-relational graph, a knowledge concept-relational graph, and a video-relational graph. Graph neural networks are then employed to preserve the knowledge of these graphs in embeddings respectively. To integrate these embeddings, we develop a multi-relation adaptive LSTM to customize the relational weights and incorporate the temporal patterns of learners’ behaviors. In the second stage, we use the learned embeddings of the previous stage to recommend educational content at three different levels. A new hierarchical beam search strategy is implemented to narrow down the recommendation scope from course to knowledge concept to video. For each level, the probability of recommending the content is determined by calculating the joint probability that takes into account the recommendation from the preceding level. Additionally, we develop an optional layer with two modes: personalization and diversification to enhance the diversity and accuracy of recommendation results. The personalization mode focuses on providing videos based on the learner’s interests and learning process, while the diversification mode aims to recommend more new and related educational content.

Multi-relational representation

The multi-relational representation needs to capture and integrate three types of graphs: course graph-relational \mathcal{G}_C , knowledge concept-relational graph \mathcal{G}_K and video-relational graph \mathcal{G}_V^T . At each time step t , given a learner u , the previous learned course c_{t-1} , the knowledge concept k_{t-1} and video v_{t-1} , an embedding layer E_W is employed to map these entities into trainable vector embedding:

$$e_{t-1}^c, e_{t-1}^k, e_{t-1}^v = E_W(c_{t-1}, k_{t-1}, v_{t-1}), \quad (1)$$

$$\mathbf{W} \in \{\mathbf{W}_C, \mathbf{W}_K, \mathbf{W}_V\}$$

where $\mathbf{W}_C \in \mathbb{R}^{|C| \times \delta}$, $\mathbf{W}_K \in \mathbb{R}^{|K| \times \delta}$, $\mathbf{W}_V \in \mathbb{R}^{|V| \times \delta}$ are the courses, knowledge concepts and videos weight matrices respectively. δ is the embedding dimension. e_{t-1}^c , e_{t-1}^k and e_{t-1}^v are the embeddings for the course learned, the knowledge concept studied and the video watched at time $t - 1$ with respectively.

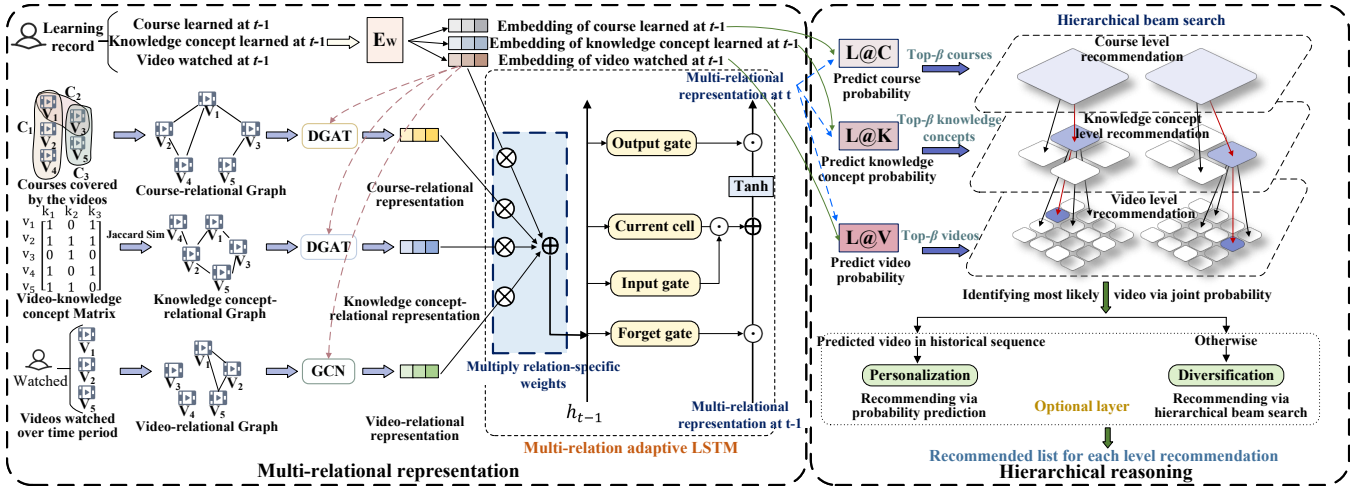


Figure 1: An overview of TOME. First, we construct three relational graphs and employ various graph neural networks to obtain representations. Then, we use a multi-relation adaptive LSTM to integrate them to capture the intricate interrelations among various educational contents. After that, we introduce a hierarchical reasoning strategy to recommend top- β educational content at each level. Moreover, we develop an optional layer to enhance the diversity and popularity of recommendations.

In \mathcal{G}_C and \mathcal{G}_K , it is particularly important to accentuate the intricate relationships among adjacent nodes. Hence, two Dimensional Graph Attention Networks (DGATs) (Lim et al. 2020) are employed to obtain the representations of these two relational graphs, denoted as $e_t^{\mathcal{G}_C}$ and $e_t^{\mathcal{G}_K}$, which are accomplished through Eq (4) and (5). The detail of DGAT are presented in Eq (2) and (3).

$$\alpha_{v_{t-1},j} = \frac{\exp(LR(\mathbf{a}[\mathbf{W}_p e_{t-1}^v \oplus \mathbf{W}_p e^{v_j}]))}{\sum_{e^{v_j} \in \mathcal{N}_{\mathcal{G}}(v_{t-1})} \exp(LR(\mathbf{a}[\mathbf{W}_p e_{t-1}^v \oplus \mathbf{W}_p e^{v_j}]))} \quad (2)$$

$$p_t = \sum_{v_j \in \mathcal{N}_{\mathcal{G}}(v_{t-1})} \alpha_{v_{t-1},j} \odot \mathbf{W}_p e^{v_j} \quad (3)$$

$$e_t^{\mathcal{G}_C} = DGAT(e_{t-1}^v, \mathcal{G}_C) \quad (4)$$

$$e_t^{\mathcal{G}_K} = DGAT(e_{t-1}^v, \mathcal{G}_K) \quad (5)$$

where LR is the LeakyRelu function, $\mathbf{W}_p \in \mathbb{R}^{\delta \times \delta}$ is an input projection, and a linear layer \mathbf{a} parameterized with $\mathbf{W}_a \in \mathbb{R}^{2 \cdot \delta \times \delta}$ to predict the attention weights $\alpha_{v_{t-1},j} \in \mathbb{R}^{\delta}$ between history video input v_{t-1} and each video of its closed neighbourhood $v_j \in \mathcal{N}_{\mathcal{G}}(v_{t-1})$ from \mathcal{G}_C and \mathcal{G}_K . e_{t-1}^v and e^{v_j} are the embeddings obtained from Eq (1). The attention weight $\alpha_{v_{t-1},j}$ is applied to calculate the weighted sum of its corresponding neighbors, ultimately generating the hidden representation $p_t \in \mathbb{R}^{\delta}$.

We employ Graph Convolutional Network (GCN) (Kipf and Welling 2016), which is particularly adept at handling sparse graphs, to learn the representation of \mathcal{G}_V^T , as shown in Eq (6).

$$e_t^{\mathcal{G}_V^T} = GCN(e_{t-1}^v, \mathcal{G}_V^T) \quad (6)$$

Furthermore, a multi-relation adaptive LSTM is designed to integrate representations of various relations while simul-

taneously capturing their individual distinctiveness. The specific implementation is described as follows:

$$i_t = \sigma(\mathbf{W}_i e_t^v + \mathbf{U}_i h_{t-1} + \mathbf{V}_i e_t^{\mathcal{G}_C} + \mathbf{Z}_i e_t^{\mathcal{G}_K} + \mathbf{Q}_i e_t^{\mathcal{G}_V^T} + \mathbf{b}_i) \quad (7)$$

$$f_t = \sigma(\mathbf{W}_f e_t^v + \mathbf{U}_f h_{t-1} + \mathbf{V}_f e_t^{\mathcal{G}_C} + \mathbf{Z}_f e_t^{\mathcal{G}_K} + \mathbf{Q}_f e_t^{\mathcal{G}_V^T} + \mathbf{b}_f) \quad (8)$$

$$o_t = \sigma(\mathbf{W}_o e_t^v + \mathbf{U}_o h_{t-1} + \mathbf{V}_o e_t^{\mathcal{G}_C} + \mathbf{Z}_o e_t^{\mathcal{G}_K} + \mathbf{Q}_o e_t^{\mathcal{G}_V^T} + \mathbf{b}_o) \quad (9)$$

$$c_t = \tanh(\mathbf{W}_c e_t^v + \mathbf{U}_c h_{t-1} + \mathbf{V}_c e_t^{\mathcal{G}_C} + \mathbf{Z}_c e_t^{\mathcal{G}_K} + \mathbf{Q}_c e_t^{\mathcal{G}_V^T} + \mathbf{b}_c) \quad (10)$$

$$e_t = o_t \odot \tanh(f_t \odot e_{t-1} + i_t \odot c_t) \quad (11)$$

where $\mathbf{W}, \mathbf{V}, \mathbf{Z}, \mathbf{Q}, \mathbf{U} \in \mathbb{R}^{\delta}$ are the weights for $e_t^v, e_t^{\mathcal{G}_C}, e_t^{\mathcal{G}_K}, e_t^{\mathcal{G}_V^T}$, respectively, and h_{t-1} is the hidden state, each of δ dimensions. $i_t, f_t, o_t, c_t \in \mathbb{R}^{\delta}$ represent the input gate, forget gate, output gate, and the current cell state. Finally, we obtain the multi-relational representation e_t at time t .

Hierarchical reasoning

In hierarchical reasoning stage, multi-type MOOCs recommendation is designed across three levels: course level recommendation, knowledge concept level recommendation and video level recommendation. Due to the uniform method of calculating probabilities across each recommendation level, the probability of recommending each entity at each level tk is predicted using Eq (12).

$$tk = \text{softmax}(DL_{\mathbf{W}}(DO(e_t \oplus \tilde{e}_{t-1}))) \quad (12)$$

where $e_t \in \mathbb{R}^{\delta}$ is the hidden embedding in Eq (10), \tilde{e}_{t-1} is the embedding of current level. e_t as a common embedding is concatenated with \tilde{e}_{t-1} . DO is the dropout layer, DL is projected onto the $|V|$ videos as a linear layer parameterized with $\mathbf{W} \in \mathbb{R}^{\delta + \delta \times |V|}$. The softmax normalization computes the conditional probability of the current level.

Hierarchical beam search. Since courses, knowledge concepts and videos are interrelated, we use hierarchical beam search to utilize the joint probabilities of the top- β entities from both current and previous levels to calculate prediction results across different levels, where β is the beam width. The process of traversing and sorting the sum of each level across all entities proves to be inefficient. As a result, we employ hierarchical beam search, which mitigates the impact of sparse interaction and significantly diminishes the search space. Specifically, given a continuous level of recommendation distribution from $\{tk_C, tk_V, tk_K\}$, such as (tk_C, tk_K) and (tk_K, tk_V) , the calculation method for each consecutive recommendation is consistent. For instance, taking the prediction of the top- β knowledge concept as an example, which is calculated by Eq (13).

$$tk_K^\beta = \log(tk_{C,b}^\beta) + \log(tk_{K,j}), tk_{K,j} \in \mathcal{N}(tk_{C,b}^\beta), \quad (13)$$

$$b \in \{1, 2, \dots, \beta\}$$

where tk_C^β and tk_K^β are the top- β probability of recommending entities for the input level recommendation of tk_C and tk_K respectively. They represent the sum of the logarithmic probabilities of all entities traversed within their respective searching path. In the knowledge concept level, $tk_{K,j} \in \mathcal{N}(tk_{C,b}^\beta)$ represents the neighborhood knowledge concepts that are directly connected to the top- β courses. By calculating the sum of the log probabilities of the knowledge concepts associated with j within these course neighborhoods, the probability of top- β knowledge concepts can be determined through ranking.

We show an example of the hierarchical beam search with $\beta = 2$ (purple areas are searched) in Figure 1. Firstly, we search the top-2 courses by ranking probability of at course level. Additionally, we only compute the top-2 knowledge concepts by Eq (13) which is associated with these two courses. Furthermore, we compute the top-2 videos containing the knowledge concepts searched. Finally, we can calculate the recommended distribution among the three levels through the joint probability calculation.

Optional layer. Due to the distinct nature of the video level recommendation compared to another two level recommendations, learners who haven't grasped the content of a video tend to prefer videos with clearer explanations rather than re-watching the same video. Similarly, those who have understood the content seek new knowledge. Therefore, we develop an optional layer including personalization mode and diversification mode to enrich the diversity of the recommended entities and lower the likelihood of recommending repetitive content. If the video predicted as the most likely to be watched through hierarchical beam search is found in the historical sequence, select the personalization mode; otherwise, opt for the diversification mode. In personalization mode, the recommended list of each level is obtained by ranking the probability in Eq (12). In diversification mode, the recommended list of each level is obtained by sorting the probability through hierarchical beam search.

Optimization

For each level recommendation, the cross-entropy loss is $\mathcal{L}_{L@D} = -\sum_{n=1}^N y_n^{L@D} \log(\hat{y}_n^{L@D})$, $D \in \{C, K, V\}$, where $y_n^{L@D}$ is the actual label and $\hat{y}_n^{L@D}$ is the predicted label across course, knowledge concept and video recommendations. For the n -th training sample, N is the number of training samples, and the overall loss $\mathcal{L} = \frac{1}{|D|} \sum \mathcal{L}_{L@D}$ is calculated using equal weights to be non-biased towards any level recommendation.

Experiments

Experimental setup

Datasets. We conduct experiments on two real-world datasets (e.g. Computer, MOOC3), obtained from *Xuetang X¹* MOOCs platform. Table 1 shows the data statistics.

Dataset	Computer	MOOC3
#Learner	14860	4166
#Video	3107	4040
#Course	105	137
#Knowledge Concept	1263	5815

Table 1: Dataset key statistics *w.r.t* Computer and MOOC3.

Evaluation metric. During the evaluation process, we employ distinct metrics tailored to different levels of recommendation tasks. For video-level and course-level recommendations, we evaluate the performance in terms of the following metrics: $HR@5$, $NDCG@5$ and MRR , as described in (Wu et al. 2023). To objectively evaluate multiple knowledge concept recommendations, we introduce a novel metric $FACC$, which is designed to flexibly assess the accuracy of varying numbers of knowledge concept labels for different learners. The calculation equation is as follows:

$$FACC = \frac{1}{N_{test}} \sum_{n=1}^{N_{test}} \frac{|tl_n \cap r_n|}{|tl_n|} \quad (14)$$

where N_{test} is the number of test data, r represents the list of recommended knowledge concepts, and tl denotes the list of ground truth knowledge concepts. The length of r is matched with the length of tl , which corresponds to the knowledge concepts learned by learners next time. We have also incorporated the N^2 evaluation metrics (Cheng et al. 2013; Feng et al. 2020), which focuses on recommending entities that have not been interacted with in historical records, to assess the performance of TOME in recommending previously unconnected content. Such N^2 -based metrics are employed to demonstrate the recommendation diversity of TOME instead of comparing with other baselines.

Baseline algorithms. We compare TOME with 9 outstanding single-type sequential recommendations and 3 widely multi-task recommendations. In the single-type sequential recommendation area, we respectively conduct the course, knowledge concept, and video recommendation tasks. The

¹<https://www.xuetangx.com>

Methods	Computer							MOOC3						
	Course			Video			Knowledge concept	Course			Video			Knowledge concept
	HR@5	NDCG@5	MRR	HR@5	NDCG@5	MRR	FACC	HR@5	NDCG@5	MRR	HR@5	NDCG@5	MRR	FACC
COTREC	0.2616	0.1524	0.1398	0.0350	0.0204	0.0235	0.0130	0.0708	0.0467	0.0518	0.0168	0.0112	0.0118	0.0013
lightGCN	0.2997	0.0810	0.1770	0.0197	0.0124	0.0151	0.0216	0.1587	0.0422	0.1086	0.0182	0.0114	0.0188	0.0072
SR-GNN	0.0742	0.0162	0.0649	0.0404	0.0088	0.0228	0.0252	0.1829	0.0384	0.0937	0.0266	0.0062	0.0152	0.0084
CoSeRec	0.6452	0.3126	0.2437	0.0571	0.0375	0.0389	0.2678	0.0113	0.0074	0.0075	0.0398	0.0241	0.0257	0.0275
SRMA	0.6710	0.3172	0.2413	0.6600	0.5583	0.5366	0.0672	0.0120	0.0078	0.0076	0.0130	0.0083	0.0087	0.0295
S3Rec	0.7587	0.3807	0.2867	0.5843	0.4722	0.4462	0.0299	0.7038	0.3924	0.3220	0.6930	0.6000	0.5771	0.0277
TiCoSeRec	0.8410	0.4975	0.3981	0.0668	0.0395	0.0442	0.0817	0.8012	0.5516	0.4813	0.0852	0.0540	0.0571	0.0231
IntGNN	0.9271	0.8920	0.8845	0.3770	0.2219	0.1933	0.0290	0.9126	0.5948	0.8898	0.0718	0.0458	0.0468	0.0212
MCLRec	0.9336	0.8956	0.8873	0.6717	0.6114	0.6017	0.0054	0.9306	0.9047	0.8992	0.7585	0.7088	0.6981	0.0028
MMoE	0.2909	0.1612	0.1606	0.0586	0.0321	0.0470	0.1944	0.3937	0.2615	0.2318	0.0018	0.0010	0.0009	0.0991
SNR	0.1153	0.0542	0.0596	0.0027	0.0017	0.0017	0.0146	0.1151	0.0499	0.0649	0.0028	0.0021	0.0019	0.0905
PLE	0.1960	0.1059	0.1054	0.0056	0.0033	0.0040	0.0641	0.3583	0.1986	0.1943	0.0048	0.0028	0.0041	0.0986
TOME	0.9484	0.9071	0.8967	0.7707	0.7245	0.7155	0.6600	0.9427	0.9246	0.9216	0.7959	0.7586	0.7514	0.4703

Table 2: Overall Performance for course, video and knowledge concept recommendations on Computer and MOOC3 datasets.

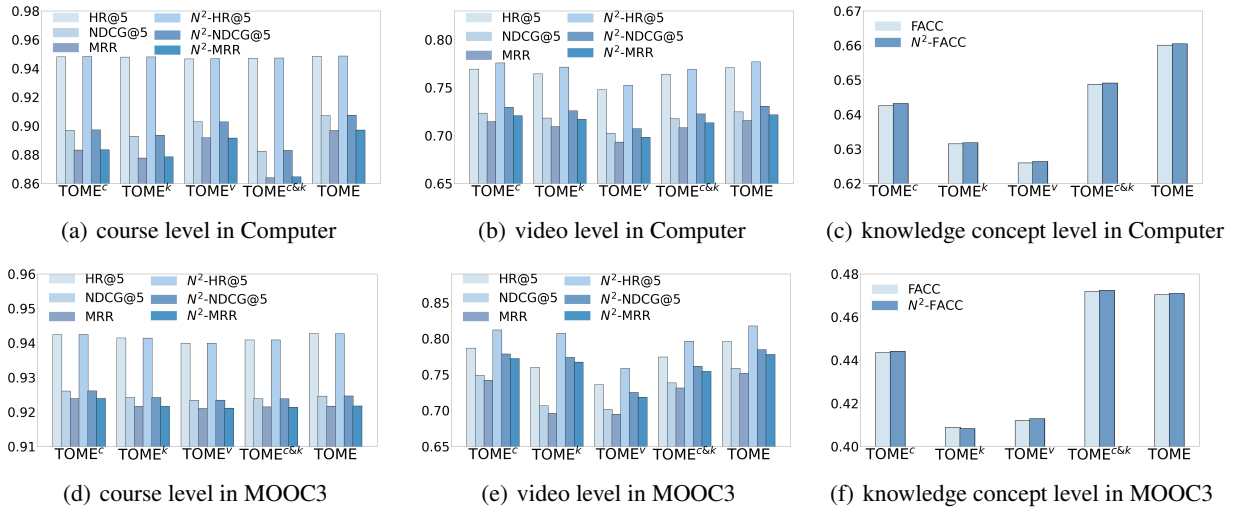


Figure 2: Analysis of the impact of different relational graphs on three-level recommendations across Computer and MOOC3.

development of the multi-task MOOCs recommendation domain is hindered by a focus on two sides: predicting browsing durations and recommending single-item types. It overlooks the need for a broader, integrated multi-task framework. Thus, we utilize the representations trained by TOME as the input features for the multi-task recommendation to fairly validate the effectiveness of our approach.

1. Single-type recommendation methods

- **COTREC** (Xia et al. 2021) combines self-supervised learning with collaborative training in a graph-based framework to enhance session-based recommendation.
- **lightGCN** (He et al. 2020) proposes a simplified GCN design that only includes the neighborhood aggregation component in GCN.
- **SR-GNN** (Wu et al. 2019) advances session-based recommendations by transforming session sequences into graph data to detect complex item interactions.
- **CoSeRec** (Liu et al. 2021) enhances sequential recommendation by integrating robust data augmentation, contrastive self-supervised learning for sequence agreement, and multi-task training.
- **SRMA** (Liu et al. 2022) uses neuron masking, layer dropping, and encoder complementing to enhance data representation and model robustness.

- **S3Rec** (Zhou et al. 2020) exploits self-supervision from data correlations and mutual information maximization to learn complex relationships among items.
- **TiCoSeRec** (Dang et al. 2023) designs five operators to control sequence data, converting non-uniform to uniform sequences considering time interval variance.
- **IntGNN** (Xu et al. 2023) mines the potential user intent by counting the frequency of an item appeared in a session and the interval between item re-interactions.
- **MCLRec** (Qin et al. 2023) employs meta-learning to refine model augmenters and uses contrastive regularization to prevent overly similar recommendations.

2. Multi-task recommendation methods

- **MMoE** (Ma et al. 2018) enhances multi-task learning by allocating expert submodels to tasks, with a customized gating network optimizing each task.
- **SNR** (Ma et al. 2019) effectively boosts multi-task learning by turning shared hidden layers into learnable and flexible sub-networks.
- **PLE** (Tang et al. 2020) elevates efficiency in joint representation learning and task routing by separating shared and task-specific items for semantic extraction.

Experimental results

Overall performance. In this experiment, we compare the performance of TOME against baseline models across courses, knowledge concepts, and videos using two real datasets. Table 2 shows the comparison results in terms of all evaluation metrics across Computer and MOOC3 datasets. We can find that TOME significantly outperforms other baselines across all types of recommendations. The underlying driver for this observation is that TOME can comprehensively capture the learners’ intent through the information integration of multi-relational representation and accurately recommend relevant learning content via hierarchical reasoning. Additionally, we notice that for all models, the performance of knowledge concept recommendations on the Computer dataset significantly surpasses that on the MOOC3 dataset. A potential reason for this observation is that the Computer dataset includes more detailed knowledge concepts interactions, resulting in better recommendation performance. Moreover, another interesting observation is that multi-task recommendation approaches do not stand out across all datasets. A possible reason is that such approaches often focus on predicting browsing duration and single-type items and ignore the relationships among different learning contents, leading to weak behavior in recommendations. In summary, this experiment underscores the effectiveness of TOME in multi-type MOOCs recommendations.

Analysis of the influence of different relational graphs in representation. In this experiment, we investigate the impact of three different relational graph representations in our video recommendation model. We develop five model variants of TOME: $TOME^C$, focusing exclusively on the course graph representation; $TOME^K$, solely maintaining the knowledge concept graph representation; $TOME^V$, keeping only the video graph representation; and $TOME^{C\&K}$, integrating both course and knowledge concept graph representations. Figure 2 shows the comparison results. We observe that the performance of each single-relational graph varies. Except in course-level recommendations, where the number of courses is relatively limited, TOME generally outperforms other model variants. This suggests that integrating multiple relational graph representations helps to better capture learners’ interests, leading to improved recommendations. Furthermore, TOME tends to perform better than other variants in terms of N^2 -based evaluation metrics, although the improvement is less pronounced in course-level recommendations. This indicates that the optional layer in its configuration generally enhances recommendation diversity. In conclusion, this experiment demonstrates that each technical relational graph representation in TOME is indispensable to comprehending learners’ intents for better recommendations.

Analysis of the impact of beam width β in hierarchical reasoning. In order to analyze the contribution of beam width β , we report the performance when $\beta \in \{1, 10, 100\}$ in hierarchical beam search process at hierarchical reasoning stage. Figure 3 illustrates the recommendation results for different β on course level, video level and knowledge concept level. The results ending with V and C correspond to

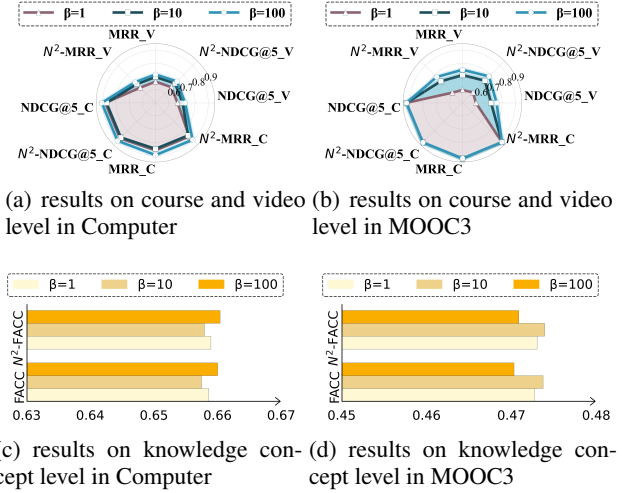


Figure 3: Analysis of the effect of different beam width β .

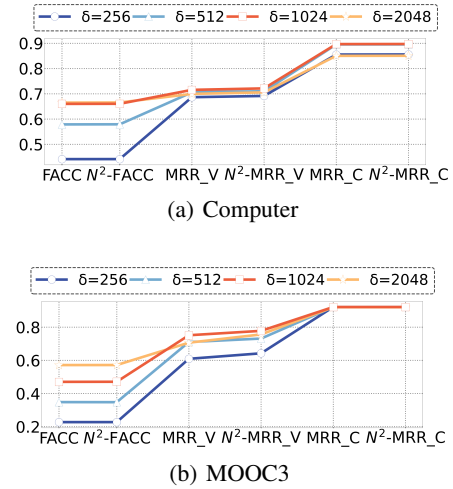


Figure 4: Analysis of the impact of various dimensions δ .

recommendations at the video level and course level, respectively. It is evident that $\beta = 100$ achieves superior performance across the three levels of recommendations—course, video, and knowledge concept—on most datasets. However, in the Computer dataset, $\beta = 10$ achieves the best performance at the knowledge concept level. The difference between $\beta = 10$ and $\beta = 100$ is minimal, which indicates that higher values of β maintain consistent effectiveness in mitigating the challenges posed by the sparsity of interaction data, while slightly smaller β values can also perform optimally in specific scenarios. Therefore, the hierarchical beam search proves to be a highly robust and adaptable strategy for significantly improving recommendation performance under varying levels of data sparsity and interaction patterns.

Effect of different embedding dimensions. In this experiment, we investigate the impact of different embedding dimension settings during the multi-relational representation stage. Experiments are conducted with various dimensions (specifically 256, 512, 1024, and 2048). As illustrated in Figure 4, it is evident that the smaller the representational dimension, the more information is lost, resulting in inferior results. However, an overly large dimension of representa-

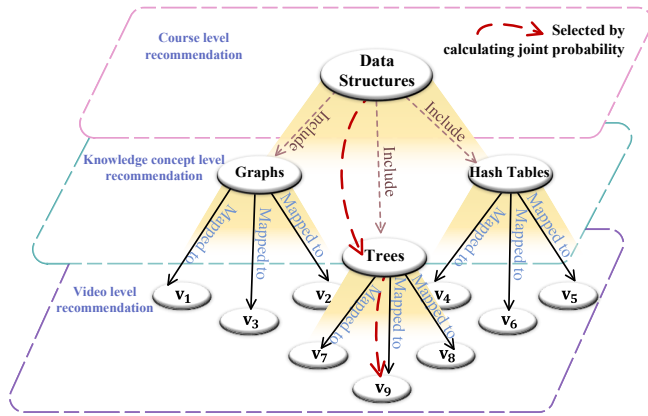


Figure 5: An example of multi-type top-1 recommendations at course level, knowledge concept level and video level. There is an inherent hierarchical relationship among three levels. First, the “Trees” knowledge concept is recommended by calculating the joint probability with the course and related knowledge concepts. Then, v_9 is recommended based on joint probabilities across three levels.

tion results in a surplus of redundant information. Therefore, 1024 is the best embedding dimension to make TOME achieve the best results across all three-level recommendations. Additionally, among these three-level recommendations, the course-level recommendation is also the easiest to yield the best results. A potential reason for this observation is the larger number of videos and knowledge concepts, making these two-level recommendations relatively more challenging. Moreover, under the same representational dimension, the results of N^2 are quite close to the normal results, and only significant differences are observed in video-level recommendations.

Case study. Figure 5 presents a case study that showcases the hierarchical reasoning of TOME during the recommendation process. This case study uses a specific example, which involves one course, three knowledge concepts, and nine videos for three-level top-1 recommendations. There exists inherent hierarchical relations among courses, knowledge concepts, and videos. As depicted in this Figure, the course of Data Structures includes three knowledge concepts: Graphs, Trees, and Hash Tables, each of which maps to three videos. Firstly, TOME acquires representations of learners across multiple relations during the multi-relation representation stage. Subsequently, in hierarchical reasoning stage, it calculates probabilities to predict the course that the learner is likely to learn. In this context, the Data Structures course is used as an example of the course most likely to be learned. Following this, the joint probability calculation of the course and predicted knowledge concepts is used to identify the recommended knowledge concept of “Trees”. Then, the joint probability across three levels of course, knowledge concept, and video, is utilized to predict the video that is most likely to watch. Ultimately, TOME yields a multi-type MOOCs recommendation that simultaneously recommends courses, knowledge concepts, and videos. This hierarchical reasoning mechanism can not only improve interpretability but also enhance recommendation accuracy.

Conclusion

We propose a novel multi-type MOOCs recommendation named TOME, comprising two stages: multi-relational representation and hierarchical reasoning. The multi-relational representation stage is designed to capture relational among courses, knowledge concepts, and videos to enhance recommendation performance. Initially, we construct two static graphs which are course relational graph and knowledge concept-relational graph, along with a dynamic graph for video-relational graph, to capture video representation within multi-relational context. Subsequently, we employ a multi-relation adaptive LSTM to better integrate representations of different relations. To address the sparsity of course interaction data, knowledge concept interaction data, and video interaction data, we adopt a hierarchical beam search in hierarchical reasoning stage. This involves computing the joint probabilities of contents at both current and past levels, thereby integrating the logical relationships across the three-level recommendations to enhance the interpretability of our recommendation results. Furthermore, to ensure the diversity of our recommendations, we have introduced an optional layer. This not only increases the variety of recommendations but also aims to enhance learners’ interest in exploring new knowledge. Experimental results on two real-world datasets indicate the effectiveness of TOME.

Related Work

Existing MOOCs recommendation methods can be classified into two categories: 1) content-based recommendation; 2) meta-path-based recommendation. The content-based MOOCs recommendations (Lin et al. 2021; Zhang et al. 2019; Jiang, Pardos, and Wei 2019; Shaw and Patra 2022; Lin et al. 2022) typically utilize sequential recommendation models that analyze learners’ historical learning records to suggest content based on preferences or similarities to history learning records. However, due to these methods conduct recommendations based on learners’ historical records, they inadequately comprehend learning intentions and produce many repetitive education contents recommendations. The meta-path-based MOOCs recommendations (Hao, Li, and Bai 2023; Gong et al. 2020; Wang et al. 2023b; Zhou et al. 2022) construct a heterogeneous information network including learners and video content. Employing random walks to capture meta-path information, they recommend interpretable and coherent content. However, these methods overlook interactions between learners and various content types, hindering the capture of learning intentions and preferences. Both approaches focus on single-type recommendations, limiting the ability to provide multi-type educational content and impeding learning progress. To overcome the limitations of existing works, we propose a novel multi-type MOOCs recommendation via multi-relational representation and hierarchical reasoning. This framework utilizes the multi-relational representation module to sufficiently capture the intricate relations among different educational contents. Then, it employs the hierarchical reasoning module to precisely capture the learners’ learning preferences to enhance the accuracy and interpretability of recommendations.

Acknowledgments

This work was supported by the Natural Science Research Foundation of Jilin Province of China under Grant No. YDZJ202201ZYTS423, NSFC under Grant No. (62206045, 62106040, 62306149), Jilin Province Philosophy and social science planning project/major project under Grant No.2023ZD15, the Fundamental Research Funds for the Central Universities under Grant No.(2412022QD041, 2412022ZD016, 2412022QD040), Jilin Education Department Project under Grant No. (JJKH20231319KJ, JJKH20231320KJ, JJKH20241430KJ), Jilin Science and Technology Association under Grant No. QT202320, National Social Science Foundation in 2024 under Grant No.24BMZ101, 2023 Project of the 14th Five-Year Plan for Scientific Research of the State Language Commission under Grant NO. YB145-73, and the Startup Foundation for Introducing Talent of NUIST under Grant No. 2023r045.

References

- Alatrash, R.; Chatti, M. A.; Ain, Q. U.; Fang, Y.; Joarder, S.; and Siepmann, C. 2024. ConceptGCN: Knowledge concept recommendation in MOOCs based on knowledge graph convolutional networks and SBERT. *Computers and Education: Artificial Intelligence*, 6: 100193.
- Cheng, C.; Yang, H.; Lyu, M. R.; and King, I. 2013. Where you like to go next: Successive point-of-interest recommendation. In *Twenty-Third international joint conference on Artificial Intelligence*.
- Dang, Y.; Yang, E.; Guo, G.; Jiang, L.; Wang, X.; Xu, X.; Sun, Q.; and Liu, H. 2023. Uniform Sequence Better: Time Interval Aware Data Augmentation for Sequential Recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, 4225–4232.
- Feng, S.; Tran, L. V.; Cong, G.; Chen, L.; Li, J.; and Li, F. 2020. Hme: A hyperbolic metric embedding approach for next-poi recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on research and development in information retrieval*, 1429–1438.
- Gong, J.; Wan, Y.; Liu, Y.; Li, X.; Zhao, Y.; Wang, C.; Lin, Y.; Fang, X.; Feng, W.; Zhang, J.; et al. 2023. Reinforced moocs concept recommendation in heterogeneous information networks. *ACM Transactions on the Web*, 17(3): 1–27.
- Gong, J.; Wang, S.; Wang, J.; Feng, W.; Peng, H.; Tang, J.; and Yu, P. S. 2020. Attentional graph convolutional networks for knowledge concept recommendation in moocs in a heterogeneous view. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval*, 79–88.
- Hao, P.; Li, Y.; and Bai, C. 2023. Meta-relationship for course recommendation in MOOCs. *Multimedia Systems*, 29(1): 235–246.
- He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; and Wang, M. 2020. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, 639–648.
- Jiang, L.; Liu, K.; Wang, Y.; Wang, D.; Wang, P.; Fu, Y.; and Yin, M. 2023. Reinforced Explainable Knowledge Concept Recommendation in MOOCs. *ACM Transactions on Intelligent Systems and Technology*, 14(3): 1–20.
- Jiang, W.; Pardos, Z. A.; and Wei, Q. 2019. Goal-based course recommendation. In *Proceedings of the 9th international conference on learning analytics & knowledge*, 36–45.
- Kipf, T. N.; and Welling, M. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.
- Li, J.; Li, S.; Zhao, Y.; Guo, L.; Hao, F.; Ren, M.; and Li, K. 2023. Predicting Dropouts Before Enrollments in MOOCs: An Explainable and Self-supervised Model. *IEEE Transactions on Services Computing*.
- Liang, Z.; Mu, L.; Chen, J.; and Xie, Q. 2023. Graph path fusion and reinforcement reasoning for recommendation in MOOCs. *Education and Information Technologies*, 28(1): 525–545.
- Lim, N.; Hooi, B.; Ng, S.-K.; Wang, X.; Goh, Y. L.; Weng, R.; and Varadarajan, J. 2020. STP-UDGAT: Spatial-temporal-preference user dimensional graph attention network for next POI recommendation. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 845–854.
- Lin, Y.; Feng, S.; Lin, F.; Zeng, W.; Liu, Y.; and Wu, P. 2021. Adaptive course recommendation in MOOCs. *Knowledge-Based Systems*, 224: 107085.
- Lin, Y.; Lin, F.; Zeng, W.; Xiahou, J.; Li, L.; Wu, P.; Liu, Y.; and Miao, C. 2022. Hierarchical reinforcement learning with dynamic recurrent mechanism for course recommendation. *Knowledge-Based Systems*, 244: 108546.
- Liu, Z.; Chen, Y.; Li, J.; Luo, M.; Yu, P. S.; and Xiong, C. 2022. Improving contrastive learning with model augmentation. *arXiv preprint arXiv:2203.15508*.
- Liu, Z.; Chen, Y.; Li, J.; Yu, P. S.; McAuley, J.; and Xiong, C. 2021. Contrastive self-supervised sequential recommendation with robust augmentation. *arXiv preprint arXiv:2108.06479*.
- Ma, J.; Zhao, Z.; Chen, J.; Li, A.; Hong, L.; and Chi, E. H. 2019. Snr: Sub-network routing for flexible parameter sharing in multi-task learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 216–223.
- Ma, J.; Zhao, Z.; Yi, X.; Chen, J.; Hong, L.; and Chi, E. H. 2018. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 1930–1939.
- Qin, X.; Yuan, H.; Zhao, P.; Fang, J.; Zhuang, F.; Liu, G.; Liu, Y.; and Sheng, V. 2023. Meta-optimized Contrastive Learning for Sequential Recommendation. *arXiv preprint arXiv:2304.07763*.
- Shaw, R.; and Patra, B. K. 2022. Cognitive-aware lecture video recommendation system using brain signal in flipped learning pedagogy. *Expert Systems with Applications*, 207: 118057.

- Tang, H.; Liu, J.; Zhao, M.; and Gong, X. 2020. Progressive layered extraction (ple): A novel multi-task learning (mtl) model for personalized recommendations. In *Proceedings of the 14th ACM Conference on Recommender Systems*, 269–278.
- Tian, X.; and Liu, F. 2021. Capacity tracing-enhanced course recommendation in moocs. *IEEE Transactions on Learning Technologies*, 14(3): 313–321.
- Wang, D.; Chen, Z.; Ni, J.; Tong, L.; Wang, Z.; Fu, Y.; and Chen, H. 2023a. Hierarchical graph neural networks for causal discovery and root cause localization. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 1531–1540.
- Wang, D.; Wang, P.; Liu, K.; Zhou, Y.; Hughes, C. E.; and Fu, Y. 2021. Reinforced imitative graph representation learning for mobile user profiling: An adversarial training perspective. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 4410–4417.
- Wang, X.; Jia, L.; Guo, L.; and Liu, F. 2023b. Multi-aspect heterogeneous information network for MOOC knowledge concept recommendation. *Applied Intelligence*, 53(10): 11951–11965.
- Wang, X.; Ma, W.; Guo, L.; Jiang, H.; Liu, F.; and Xu, C. 2022. Hgnn: Hyperedge-based graph neural network for mooc course recommendation. *Information Processing & Management*, 59(3): 102938.
- Wen, J.; Qin, Y.; Su, X.; and Hu, S. 2023. Video recommendation method based on multi-attention mechanism and heterogeneous information network. In *2023 4th International Conference on Computer Engineering and Application (ICCEA)*, 669–674. IEEE.
- Wu, C.; Wu, F.; Huang, Y.; and Xie, X. 2023. Personalized news recommendation: Methods and challenges. *ACM Transactions on Information Systems*, 41(1): 1–50.
- Wu, S.; Tang, Y.; Zhu, Y.; Wang, L.; Xie, X.; and Tan, T. 2019. Session-based recommendation with graph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, 346–353.
- Xia, X.; Yin, H.; Yu, J.; Shao, Y.; and Cui, L. 2021. Self-supervised graph co-training for session-based recommendation. In *Proceedings of the 30th ACM international conference on information & knowledge management*, 2180–2190.
- Xu, G.; Yang, J.; Guo, J.; Huang, Z.; and Zhang, B. 2023. Int-GNN: A User Intention Aware Graph Neural Network for Session-Based Recommendation. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. IEEE.
- Zhang, H.; Shen, X.; Yi, B.; Wang, W.; and Feng, Y. 2023. KGAN: Knowledge grouping aggregation network for course recommendation in MOOCs. *Expert Systems with Applications*, 211: 118344.
- Zhang, J.; Hao, B.; Chen, B.; Li, C.; Chen, H.; and Sun, J. 2019. Hierarchical reinforcement learning for course recommendation in MOOCs. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, 435–442.
- Zhou, J.; Du, Y.; Song, B.; Li, X.; Zhao, A.; and Xiao, X. 2022. MOOC video recommendation method based on meta-path attention mechanism. *Journal of Computer Applications*, 42(6): 1808.
- Zhou, K.; Wang, H.; Zhao, W. X.; Zhu, Y.; Wang, S.; Zhang, F.; Wang, Z.; and Wen, J.-R. 2020. S3-rec: Self-supervised learning for sequential recommendation with mutual information maximization. In *Proceedings of the 29th ACM international conference on information & knowledge management*, 1893–1902.