EdGCL: Disentangling Social and Cognitive Homophily in Graph-Based Educational Recommender Systems

Qingqing Liang[1,2], **Chunyang Wang**[1*], **Peiwei Xia**[1], **Yanan Zeng**[1], **Xin Liu**[1], **Xuesong Lu**[1*], **Aoying Zhou**[1]

[1]School of Data Science and Engineering, East China Normal University
[2]College of Computer Science and Technology, Guizhou University
{qqliang.dase, peiwei.xia, yann_zeng}@stu.ecnu.edu.cn, {cywang, xliu, xslu, ayzhou}@dase.ecnu.edu.cn

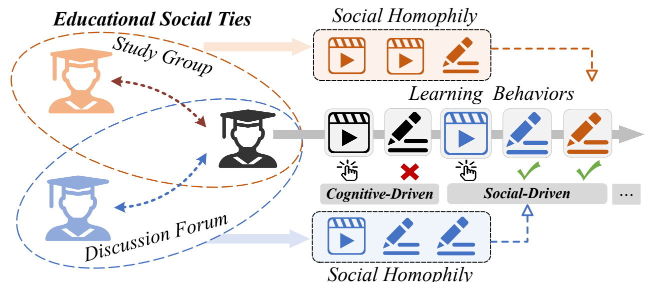
Abstract

Educational recommendation systems have been a fundamental component for alleviating learning disorientation in self-paced learning. While existing studies mainly leverage cognitive theories to guide learning motivation modeling, they critically overlook the role of social influences. Through empirical analysis, we identify social homophily as an additional driver of learning behaviors, i.e., learners tend to adopt resources validated by their social cohort. However, two challenges impede effective social homophily modeling: (1) the absence and sparsity of predefined social relations in online education, and (2) the deep entanglement of social homophily with cognitive homophily in behavioral data. To tackle these challenges, we propose a graph-based framework EdGCL that explicitly disentangles social homophily and cognitive homophily. EdGCL infers implicit social relations from learners' social behaviors and encodes them via a graph transformer, generating social-view representations. Simultaneously, it constructs a heterogeneous learning graph to model cognitive homophily, which is enhanced by a type-aware aggregator and cognitive diagnosis loss. To ensure the semantic distinctiveness of dual-view homophily modeling, a cross-view contrastive disentanglement mechanism is designed to pull intra-view representations closer while pushing inter-view representations away. Evaluation on two real-world educational datasets demonstrates the superior recommendation performance of EdGCL, highlighting the necessity of dual homophily modeling for understanding the motivations behind learning behaviors.

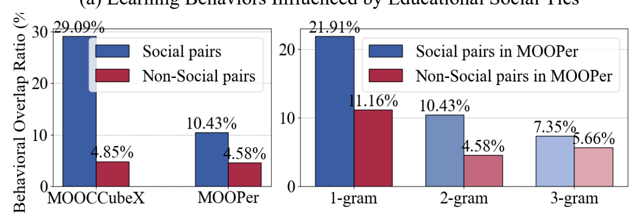
Code — https://github.com/DaSESmartEdu/EdGCL

Introduction

Educational recommender systems aim to alleviate the problems of cognitive overload (Van Merrienboer and Ayres 2005) and learning disorientation (Chen 2008) due to the overwhelming volume of resources in online learning platforms such as MOOCs and Intelligent Tutoring Systems. Current educational recommendation methods primarily adhere to the cognitive-guided paradigm, aiming to promote continuous progress in learners' mastery of complete knowledge structure. Various cognitive and pedagogical theories, such as forgetting curve (Shen et al. 2021; Bai et al. 2025) and difficulty consistency (Zhang et al. 2024), have been extensively studied to interpret motivations behind learning behaviors. However, we argue that existing studies largely overlook the role of social influences in shaping learning motivations. In educational contexts, learners usually engage in various social activities, such as *discussion forums*, *peer assistance*, *study groups*, *resource-sharing communities* and so on. Based on these activities, educational social networks can be established to reveal peer influence, which acts as a powerful catalyst in broadening learners' awareness of available learning resources.



(a) Learning Behaviors Influenced by Educational Social Ties



(b) Learning Behavioral Overlap Analysis of Social Homophily

Figure 1: The illustration and analysis of social homophily that influences learning behaviors in online education.

Crucially, educational social networks reveal **social homophily** among learners, i.e., the inherent tendency to adopt resources highly validated by one's immediate social cohort. As illustrated in Figure 1(a), besides personal cognitive needs, learning behaviors of students can be frequently driven by social ties with other learners. For instance, after incorrectly answering a question (depicted in black), the student is likely to join a relevant online discussion forum and follow the high-quality resource suggestions from so-

*Corresponding authors

cial friends as subsequent learning behaviors (depicted in blue and orange). To substantiate the social homophily phenomenon, we conduct an empirical analysis on real-world educational datasets from two online learning platforms, namely, MOOCCubeX (Yu et al. 2021) and MOOPer (Liu et al. 2021). As shown in Figure 1(b), students with social ties (i.e., social pairs) demonstrate significantly higher behavioral overlap than their isolated counterparts, i.e., students without social connections. More precisely, analyzing from 1-gram to 3-gram behavior subsequences of MOOPer shows that higher behavioral overlap ratios consistently exist among socially-tied students. Therefore, social homophily modeling is essential for developing a holistic understanding of learning motivations of students to enhance educational recommender systems.

Social homophily modeling has been studied in social recommendation (Tang, Hu, and Liu 2013), which exploits social relations among users. Despite substantial progress in social recommendation, we identify two critical challenges in adapting it to the education domain. First, the absence and sparsity of predefined social relations make it non-trivial to extract and represent social homophily in online education. Unlike explicit groups (e.g., classes or teams) prevalent in offline education environments (Liu et al. 2023; Yu et al. 2024) or readily available social networks in other domains (Sharma et al. 2024), online education typically lacks well-labeled social connections, or the connections are sparse, rendering it highly susceptible to the over-smoothing problem (Chen et al. 2020). Second, social homophily and cognitive homophily (reflected from learning activities, such as video-watching and question-answering), as two intertwined drivers of learning behaviors, are deeply coupled within behavioral data. To handle the intricate intertwining, semantic disentanglement is necessary to derive informative learner representations of different views.

To tackle the aforementioned challenges, we propose a novel graph-based educational recommendation framework, namely EdGCL, that explicitly disentangles social-view and cognitive-view homophily modeling. From the social view, EdGCL draws on learners' social behaviors, e.g, discussion forum co-occurrence, to infer implicit social relations. A graph transformer module is adopted to encode the homogeneous social graph, aggregating individual-level and community-level representations. From the cognitive view, EdGCL constructs a heterogeneous learning graph among learners and learning resources, such as courses, videos and questions. To adapt graph transformer to heterogeneous graph, we modify it by designing a heterogeneous tokenizer and a type-aware neighborhood aggregator. Furthermore, to clearly distinguish the dual-view homophily, we design a cross-view contrastive disentanglement mechanism to refine the view-specific representations. In addition, two auxiliary tasks, social link prediction and cognitive diagnosis, are added to guarantee the capacity of mining social strengths and knowledge proficiency, respectively.

Our main contributions are summarized as follows:

- We emphasize the importance of both social and cognitive homophily modeling for understanding motivations behind learning behaviors. To the best of our knowledge, our work is the first to consider social homophily modeling in educational recommender systems.
- We effectively decouple learning motivations with dual-view homophily modeling using graph transformers, and design a cross-view contrastive disentanglement mechanism and a multi-task training scheme to precisely refine social-view and cognitive-view representations.
- We conduct extensive experiments on two real-world online education datasets. Experimental results show the superior performance of EdGCL over comparative models. The ablation and visualization analysis further verifies the effectiveness of our disentanglement mechanism.

Problem Definition

In online educational platforms, let $\mathcal{V}_u = \{u_i\}_{i=1}^N$ be a set of students and $\mathcal{V}_r = \{r_j\}_{j=1}^M$ be the set of heterogeneous learning resources, e.g., courses, videos and questions. Students' behaviors can be broadly divided into two categories:

Definition 1 (**Learning Behaviors**). For a student u, the learning behaviors are denoted as $\{l_1^u, l_2^u, ..., l_t^u\}$, where each term $l_t^u = (r_t, c_t, t)$ represents that the student u has learned the recourse r_t of the type c_t at the time t.

Definition 2 (**Social Behaviors**). Besides interactions with resources, social ties or social activities also exist among students, e.g., discussing on forums, studying in groups. For a student u, the social behaviors are denoted as $\{s_1^u, s_2^u, ..., s_t^u\}$, where each term $s_t^u = (u_t, t)$ represents that a social activity occurs among u and u_t at the time t.

Problem Statement. Given learning behaviors $\{l_1^u, l_2^u, ..., l_t^u\}$ and social behaviors $\{s_1^u, s_2^u, ..., s_t^u\}$ of the student u, the goal is to predict the likelihood $\hat{y}_{ur}$ that u prefers to learn each unobserved resource r, and generate a top-k resource list that u will be most likely to learn.

Method

As illustrated in Figure 2, we present the details of our proposed EdGCL approach, which mainly consists of three components: (1) Behavior-driven graph construction module, which leverages both social and learning behaviors of students to establish corresponding graph structures. (2) Dual-view homophily modeling module, which adopts advanced graph transformers to mine high-order homophily from both social and cognitive views, respectively. (3) Cross-view disentanglement module, which explicitly decouples dual views via an auxiliary cognitive diagnosis task and a social link prediction task, and implicitly enhances semantic disentanglement by designing cross-view contrastive training loss. By fusing dual-view learner representations, top-k recommendation lists of candidate educational resources will be generated via estimated preference scores.

Behavior-Driven Graph Construction

The core foundation of graph-based recommendation lies in the accurate and meaningful graph structure among users and items. To effectively discern multifaceted learning motivations, we convert student behaviors into a homogeneous social graph and a heterogeneous learning graph, providing

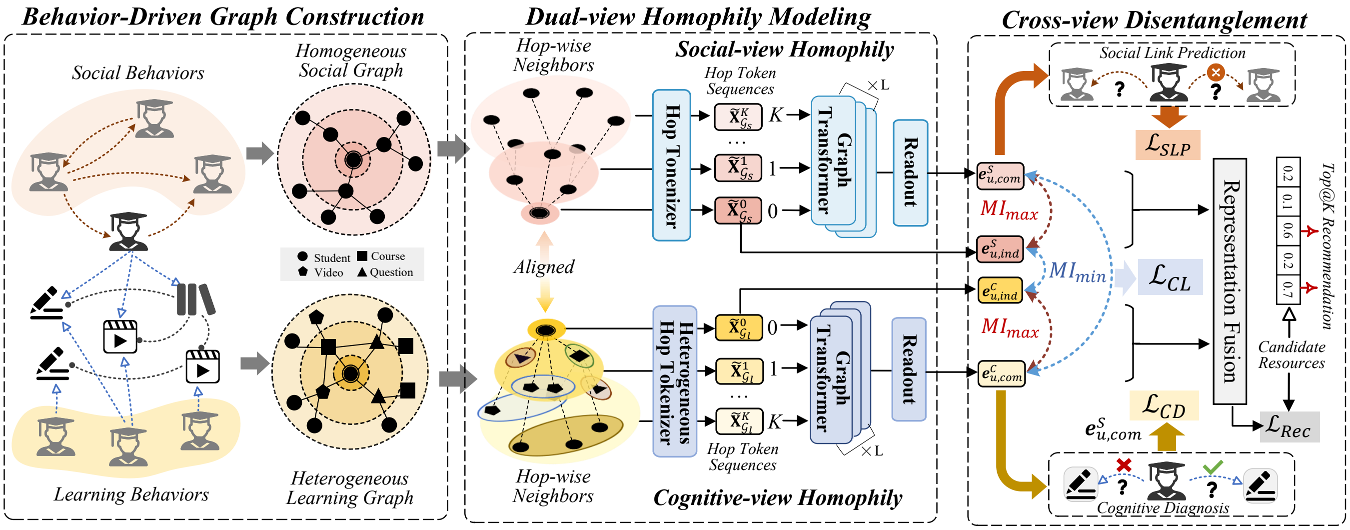


Figure 2: The overall architecture of EdGCL. Both social graph and learning graph are constructed for dual-view homophily modeling. The cross-view disentanglement module enhances semantical representation disentanglement across dual views.

distinct signals of social and cognitive views for subsequent homophily modeling and disentangling.

Homogeneous Social Graph. The social graph is formally defined as $\mathcal{G}_s = \{\mathcal{V}_u, \mathcal{E}_s, \mathcal{A}_s\}$, where $\mathcal{V}_u = \{u_i\}_{i=1}^{N}$ represents the set of student nodes, and $\mathcal{E}_s \subseteq \mathcal{V}_u \times \mathcal{V}_u$ denotes the set of social edges connecting them. Each entry in adjacency matrix $\mathcal{A}_s(i,j)$ signifies either an explicit social tie (e.g., membership in study groups) or an implicit social relation inferred from behavioral signals, such as co-participation in discussion forums. For example, if students u_i and u_j frequently participate in discussions on the same question, their social connection strengths will be enhanced.

Heterogeneous Learning Graph. To encode diverse learning behaviors with different types of resources, a heterogeneous learning graph is constructed as $\mathcal{G}_l = \{\mathcal{V}, \mathcal{E}_l, \mathcal{A}_l, \mathcal{R}_l\}$, where $\mathcal{V} = \mathcal{V}_u \cup \mathcal{V}_r$ denotes the union set of student nodes and resource nodes, and $\mathcal{E}_l$ denotes the set of student-resource interaction relations. To clearly represent node heterogeneity, each node $v \in \mathcal{V}$ is attached with a type mapping function $\phi(v) : \mathcal{V} \rightarrow \mathcal{R}_l$, with $\mathcal{R}_l$ being the set of node types (e.g., student, course, video and question). The weighted adjacency matrix $\mathcal{A}_l$ encodes normalized interaction frequencies, with each entry reflecting the cognitive engagement intensity between student u_i and resource r_k.

Dual-view Homophily Modeling

To model the homophily among similar students, effectively aggregating information from high-order neighbors via a powerful graph representation learning module is crucial. On the one hand, the inherently sparse structure of educational social networks heightens vulnerability to the over-smoothing problem during deep message propagation. On the other hand, the heterogeneous node and relation types in the learning graph introduce distinct semantics across different types of learning resources. To address these two issues, we adopt graph transformers as our foundational homophily modeling module, which inherently enhances robustness against over-smoothing via attentive multi-hop aggregation. We further adapt the multi-hop tokenizer to accommodate the heterogeneous node types within each hop.

Multi-hop Tokenizer. To utilize graph transformers, hop-level tokenization is a critical step that converts graph-structured data into sequential multi-hop tokens suitable for Transformer-based architectures. Formally, following prior works (Ying et al. 2021; Chen et al. 2023), we incorporate the Laplacian positional embeddings derived from two graphs, denoted as $\mathbf{P}_{\mathcal{G}_s}, \mathbf{P}_{\mathcal{G}_l} \in \mathbb{R}^{N \times d_1}$, to encode node structural information. Then, attribute matrix $\mathbf{X} \in \mathbb{R}^{N \times d_2}$ is concatenated to obtain initial node representations:

$$\tilde{\mathbf{X}}_{\mathcal{G}_s} = \mathbf{X} || \mathbf{P}_{\mathcal{G}_s}, \tilde{\mathbf{X}}_{\mathcal{G}_l} = \mathbf{X} || \mathbf{P}_{\mathcal{G}_l}. \quad (1)$$

Next, a hop-level token sequence is constructed via a multi-hop propagation process to capture high-order dense subgraph structures. With the homogeneous social graph, for each central student node $u \in \mathcal{G}_s$, we recursively aggregate its k-hop neighbors in a layer-wise manner, as follows:

$$\tilde{\mathbf{x}}_u^{k+1} = \sum_{n \in \mathcal{N}(u;\mathcal{G}_s)} \frac{1}{\sqrt{D_u D_n}} \cdot \tilde{\mathbf{x}}_n^k, \quad (2)$$

where $\mathcal{N}(u;\mathcal{G}_s)$ denotes the neighbors of u in $\mathcal{G}_s$, and $\frac{1}{\sqrt{D_u D_n}}$ is a symmetric normalization term based on degrees. By repeating this process up to K times, we obtain the hop-level token sequences of all nodes, denoted as $\mathbf{S}_{\mathcal{G}_s} = \{\tilde{\mathbf{X}}^0_{\mathcal{G}_s}, \tilde{\mathbf{X}}^1_{\mathcal{G}_s}, \cdots, \tilde{\mathbf{X}}^K_{\mathcal{G}_s}\} \in \mathbb{R}^{N \times (K+1) \times d}$. Here, $\tilde{\mathbf{X}}^0_{\mathcal{G}_s}$ is the individual-level representations of all central nodes themselves, and $\tilde{\mathbf{X}}^K_{\mathcal{G}_s}$ increasingly encodes broader structural contexts from their k-hop neighbors in the social graph.

To discern semantical differences across learning behaviors with heterogeneous resources, we modify the tokenizer with type-aware neighborhood aggregation by grouping k-hop neighbors by different node types to form (k, t)-hop neighbors. Then, multi-hop propagation process of the learning graph $\mathcal{G}_l$ can be adaptively conducted as follows:

$$\begin{aligned} \tilde{\mathbf{x}}_{u,\phi(n)}^{k+1} &= \sum_{n\in\mathcal{N}(u;\phi(n);\mathcal{G}_l)} \frac{1}{\sqrt{D_v D_n}} \cdot \tilde{\mathbf{x}}_n^k, \\ \tilde{\mathbf{x}}_u^{k+1} &= \sum_{\phi(n)\in\mathcal{R}_l} \alpha_{\phi(n)} \cdot \tilde{\mathbf{x}}_{u,\phi(n)}^{k+1}, \end{aligned} \tag{3}$$

where $\alpha_{\phi(n)}$ is achieved via type-aware attention mechanism. Finally, the hop-level token sequence of learning graph $\mathcal{G}_l$ is obtained as $\mathbf{S}_{\mathcal{G}_l} = \{\tilde{\mathbf{X}}_{\mathcal{G}_l}^0, \tilde{\mathbf{X}}_{\mathcal{G}_l}^1, \cdots, \tilde{\mathbf{X}}_{\mathcal{G}_l}^K\}$.

Transformer Encoder. After tokenizing graph data into hop-level token sequences, we apply transformer encoders with the same structures to learn social-view and cognitive-view representations, respectively. Formally, taking the social view as an example, we linearly transform the input token sequences into $\mathbf{H}_{\mathcal{G}_s}^0 = \mathbf{S}_{\mathcal{G}_s}\mathbf{W}$, and then stack multiple transformer layers of Multi-Head Attention (MHA) and Feed Forward Network (FFN) as follows:

$$\begin{aligned} \tilde{\mathbf{H}}_{\mathcal{G}_s}^{(l+1)} &= \text{MHA}(\text{LN}(\mathbf{H}_{\mathcal{G}_s}^l)) + \mathbf{H}_{\mathcal{G}_s}^l, \\ \mathbf{H}_{\mathcal{G}_s}^{(l+1)} &= \text{FFN}(\text{LN}(\tilde{\mathbf{H}}_{\mathcal{G}_s}^{(l+1)})) + \tilde{\mathbf{H}}_{\mathcal{G}_s}^{(l+1)}, \end{aligned} \tag{4}$$

where LN denotes layer normalization. After stacking L layers, we can obtain $\mathbf{H}_{\mathcal{G}_s}^L = \{\mathbf{H}_{\mathcal{G}_s}^{L,0}, \cdots, \mathbf{H}_{\mathcal{G}_s}^{L,K}\}$ from social graph $\mathcal{G}_s$. Meanwhile, $\mathbf{H}_{\mathcal{G}_l}^L = \{\mathbf{H}_{\mathcal{G}_l}^{L,0}, \cdots, \mathbf{H}_{\mathcal{G}_l}^{L,K}\}$ can also be similarly obtained from learning graph $\mathcal{G}_l$.

Overall, for a given student u, we take the center node token as its individual-level representations and aggregate remaining neighborhood tokens into its community-level representations via attentive readout as follows:

$$\begin{aligned} \mathbf{e}_{u,ind}^S = \mathbf{H}_{\mathcal{G}_s}^{L,0}[u], \mathbf{e}_{u,com}^S &= \sum_{k=1}^K \alpha_k \cdot \mathbf{H}_{\mathcal{G}_s}^{L,k}[u], \\ \mathbf{e}_{u,ind}^C = \mathbf{H}_{\mathcal{C}_s}^{L,0}[u], \mathbf{e}_{u,com}^C &= \sum_{k=1}^K \alpha_k \cdot \mathbf{H}_u^{L,k}[u]. \end{aligned} \tag{5}$$

Cross-view Disentanglement

We further design a cross-view disentanglement module which refines representations in Eq.(5) to extract view-specific representations while suppressing irrelevant or overlapping information. Both explicit multi-task driven disentanglement and implicit contrastive disentanglement are designed to jointly ensure the decoupling effect.

Multi-task Driven Disentanglement. To explicitly supervise the disentanglement process, one feasible way is to attach representations to corresponding downstream tasks and perform multi-task learning. For enriching social-view representations, social link prediction task is suitable to exclusively focus on extracting social-side semantics. Therefore, we additionally refine $\mathbf{e}_{u,ind}^S$ and $\mathbf{e}_{u,com}^S$ via an auxiliary social link prediction loss, which is defined as follows:

$$\mathcal{L}_{SLP} = -\sum\nolimits_{(u_i,u_j)} \text{BCELoss}(y_{u_i,u_j}, \mathcal{F}_S(\mathbf{e}_{u_i}^S; \mathbf{e}_{u_j}^S)), \tag{6}$$

where $\mathbf{e}_{u_i}^S = [\mathbf{e}_{u_i,ind}^S \| \mathbf{e}_{u_i,com}^S]$ and $\mathcal{F}_S$ can be any link prediction function. $y_{u_i,u_j} \in \{0, 1\}$ indicates whether the social relation among the given student pair exists or not.

For enriching cognitive-view representations, we resort to the cognitive diagnosis task (Shen et al. 2024), specifically designed to diagnose students' knowledge mastery based on learning behaviors with responses, i.e., correctness or incorrectness. Formally, $\mathbf{e}_{u,ind}^C$ and $\mathbf{e}_{u,com}^C$ are combined to predict the performance on question-typed resources:

$$\mathcal{L}_{CD} = -\sum_{(u_i,r_q)} \text{BCELoss}(y_{u_i,r_q}, \mathcal{F}_C(\mathbf{e}_{u_i}^C - \mathbf{e}_{r_q}^C)), \tag{7}$$

where $\mathbf{e}_{u_i}^C = [\mathbf{e}_{u_i,ind}^C \| \mathbf{e}_{u_i,com}^C]$, and $\mathbf{e}_{r_q}^C$ is the question representation obtained from learning graph transformer. Noting that the input of cognitive diagnosis function $\mathcal{F}_C$ is commonly modeled as $\mathbf{e}_{u_i}^C - \mathbf{e}_{r_q}^C$, measuring the gap between question difficulty and students' cognitive level.

Disentangled Contrastive Learning. The disentanglement effect of the aforementioned auxiliary tasks heavily relies on the quality of supervision signals, leading to limitations in situations where low-quality social labels or sparse question-answering behaviors are provided. To get rid of over-reliance on supervision signals, we propose to utilize self-supervised learning (SSL) (Jing et al. 2023) to achieve implicit yet effective representation disentanglement.

Specifically, we adopt the mutual information (MI) regularization mechanism to construct contrastive learning optimization objectives. The core idea is to select positive pairs whose MI should be maximized and negative pairs whose MI should be minimized. Specifically, we assign intra-view positive pairs as mutual information maximization terms to maintain semantical consistency within each view. In contrast, we assign cross-view negative pairs as mutual information minimization terms to ensure semantical distinctiveness across views. As shown in Figure 2, given two student u_i and u_j, their individual-level and community-level representations constitute the following contrastive pairs:

$$\begin{aligned} MI_{max}^{u_i} &= \left\{(\mathbf{e}_{u_i,ind}^S, \mathbf{e}_{u_i,com}^S), (\mathbf{e}_{u_i,ind}^C, \mathbf{e}_{u_i,com}^C)\right\}, \\ MI_{min}^{u_i} &= \left\{(\mathbf{e}_{u_i,ind}^S, \mathbf{e}_{u_i,ind}^C), (\mathbf{e}_{u_i,com}^S, \mathbf{e}_{u_i,com}^C)\right\}, \\ MI_{min}^{u_i,u_j} &= \left\{\begin{matrix}(\mathbf{e}_{u_i,ind}^S, \mathbf{e}_{u_j,ind}^C), (\mathbf{e}_{u_i,com}^S, \mathbf{e}_{u_j,com}^C) \\ (\mathbf{e}_{u_i,ind}^C, \mathbf{e}_{u_j,ind}^S), (\mathbf{e}_{u_i,com}^C, \mathbf{e}_{u_j,com}^S)\end{matrix}\right\}. \end{aligned} \tag{8}$$

Then, our cross-view contrastive optimization objective based on InfoNCE (Hjelm et al. 2018) is defined as follows:

$$\mathcal{L}_{CL} = \sum_{u_i\in\mathcal{V}_u} -log\frac{\sum_{\mathbf{pos}} exp(f(\mathbf{pos})/\tau)}{\sum_{\mathbf{neg}} exp(f(\mathbf{neg})/\tau)}, \tag{9}$$

where $f(\cdot)$ is a function measuring the mutual information contained in representations, for which we adopt cosine similarity. τ is a softmax temperature hyperparameter. Here, **pos** and **neg** denote the positive and negative pairs drawn from $MI_{max}^{u_i}$ and $\{MI_{min}^{u_i} \bigcup \sum_{u_j} MI_{min}^{u_i,u_j}\}$, respectively.

Through the cross-view disentanglement module, disentangled social-view and cognitive-view representations are obtained. To make a comprehensive recommendation decision, EdGCL adaptively fuses representations to get holistic student and resource representations, as follows:

$$\begin{aligned}
\mathbf{e}_u^S &= g^S \cdot \mathbf{e}_{u,ind}^S + (1 - g^S) \cdot \mathbf{e}_{u,com}^S, \\
\mathbf{e}_u^C &= g^C \cdot \mathbf{e}_{u,ind}^C + (1 - g^C) \cdot \mathbf{e}_{u,com}^C, \\
\mathbf{e}_u &= \mathcal{H}_u(\mathbf{e}_u^S, \mathbf{e}_u^C), \mathbf{e}_r = \mathcal{H}_r(\mathbf{e}_r^C),
\end{aligned} \tag{10}$$

where g^S and g^C are learnable fusion gates, and $\mathcal{H}_u, \mathcal{H}_r$ are the student encoder and resource encoder. We employ the dot-product to forecast $\hat{y}_{u,r} = \mathbf{e}_u^\intercal \cdot \mathbf{e}_r$ and adopt the Bayesian Personalized Ranking (BPR) loss to encourage a higher prediction than a sampled unobserved counterpart:

$$\mathcal{L}_{REC} = \sum_{(u,r^+,r^-)\in\mathcal{O}} -\ln \sigma(\hat{y}_{u,r^+} - \hat{y}_{u,r^-}) + \lambda\|\Theta\|^2, \tag{11}$$

where $\mathcal{O} = \{(u, r^+, r^-)|(u, r^+) \in \mathbf{Y}^+, (u, r^-) \in \mathbf{Y}^-\}$ is the pair-style samples, $\mathbf{Y}^+, \mathbf{Y}^-$ denotes the positive and unobserved interactions, respectively. λ controls the L_2 regularization strength to prevent over-fitting.

Overall, the total joint loss function is defined as follows:

$$\mathcal{L} = \mathcal{L}_{REC} + \alpha \cdot \mathcal{L}_{CD} + \beta \cdot \mathcal{L}_{SLP} + \gamma \cdot \mathcal{L}_{CL}, \tag{12}$$

where α, β, and γ are hyperparameters that control the contributions of different disentanglement components.

Performance Evaluation

Experiential Settings

Datasets and Evaluation Metrics. We conduct experiments on two real-world educational datasets, **MOOCCubeX** (Yu et al. 2021) and **MOOPer** (Liu et al. 2021), because they contain both learning and social behaviors of students. We filter out students with fewer than 5 learning behaviors and retain only those who engage in both types of behaviors. After that, we split the learning sequences into training, validation, and testing sets in a ratio of 6:2:2. Focusing on item ranking, we adopt Recall@K, NDCG@K and MRR as the evaluation metrics, where K is set to 5 and 10 following the conventions. In line with recent studies (Yang et al. 2024b,a), we use an all-ranking protocol that ranks each positive item with all non-interacted items. Dataset statistics are summarized in Table 1.

Dataset	MOOCCubeX	MOOPer
Students	10,893	6,057
Resources	66,891	3,784
Social Edges	3,589,654	1,316,667
Social Graph Density	3.03%	3.59%
Learning Edges	2,150,392	422,602
Learning Graph Density	0.29%	1.63%
Avg. Learning Sequence Length	191	68

Table 1: Statistics of datasets.

Comparative Methods. We compare EdGCL with 10 representative recommendation models: KGCL (Yang et al. 2022), KGRec (Yang et al. 2023), ACKRec (Gong et al. 2020), GCARec (Yu et al. 2023), CL-KCRec (Gu et al. 2024), DSL (Wang, Xia, and Huang 2023), RecDiff (Li, Xia, and Huang 2024), HGSR (Yang et al. 2024b), HDSR (Hu et al. 2025), TALLRec (Bao et al. 2023). The first two are general GNN-based recommendation models, the next three are educational recommendation models using graphs, the following four are recent social recommendation models using graphs, and the last one is an LLM-based recommendation model, where we select DeepSeek-R1-Distill-Qwen-7B (Guo et al. 2025) as the backbone LLM.

Implementation Details. The EdGCL model is implemented in PyTorch and optimized using the Adam optimizer. The learning rate is set to $1\mathrm{e}^{-3}$ with a weight decay of $1\mathrm{e}^{-5}$ for MOOPer and $1\mathrm{e}^{-7}$ for MOOCCubeX. The hop length K, defining the graph neighborhood range, is set to 4 for MOOCCubeX and to 5 for MOOPer. The node feature dimensions are set to 128 for MOOPer and 64 for MOOCCubeX, and we use the 15 eigenvectors of the Laplacian matrix as the positional encodings, which are fed into a 3-layer Transformer encoder. For auxiliary tasks, we conduct a grid search on the loss weights and set α, β, γ to 0.01, 0.01, $5\mathrm{e}^{-5}$ for MOOCCubeX and 0.05, 0.1, $1\mathrm{e}^{-4}$ for MOOPer. The temperature in contrastive learning is set to 0.05. The temperature of DeepSeek is set to 0.1. All experiments are conducted on an NVIDIA A800 GPU with 80GB memory.

Overall Performance

Table 2 presents the overall performance comparison between EdGCL and comparative baselines. All Results are averaged over 4 runs with different random seeds. We derive the following observations: (1) EdGCL consistently outperforms all baselines across all metrics. Compared with the best-performing baselines (underlined), it achieves significant improvements on both datasets for all metrics. (2) Compared to the first two general graph-based recommendation methods and the next three graph-based educational recommendation methods, EdGCL explicitly incorporates social behaviors to enrich student semantic representations, enabling better capturing of peer influence that is critical in educational settings. (3) Compared to the following four graph-based social recommendation methods, EdGCL leverages Graph Transformer to more effectively capture high-order homophily patterns in both cognitive and social views. (4) TALLRec performs generally worse than other baselines, indicating that LLMs do not necessarily adapt well to specific tasks like recommendation. In summary, EdGCL achieves the new SOTA performance by disentangling social and cognitive homophily.

Ablation Study

To better understand the contribution of each component in EdGCL, we conduct ablation studies from two perspectives: modeling components and disentanglement objectives.

Modeling Perspective. We first analyze the effectiveness of the core modeling components in EdGCL by design-

Model	MOOCCubeX					MOOPer				
	Recall@5	Recall@10	NDCG@5	NDCG@10	MRR	Recall@5	Recall@10	NDCG@5	NDCG@10	MRR
KGCL [2022]	0.2875	0.4816	0.3465	0.4973	0.2591	0.3898	0.5486	0.4093	0.5007	0.2809
KGRec [2023]	0.3259	0.4953	0.3939	0.5468	0.2814	0.4409	0.5479	0.4801	0.5450	0.3469
ACKRec [2020]	0.2757	0.4566	0.3736	0.5057	0.2619	0.4110	0.5215	0.4457	0.5130	0.3025
GCARec [2023]	0.2426	0.4344	0.3663	0.5127	0.2762	0.3889	0.5526	0.3904	0.5161	0.2906
CL-KCRec [2024]	0.3397	0.5368	0.4141	0.5779	0.2794	0.4358	0.5834	0.4672	0.5531	0.3502
DSL [2023]	0.2781	0.4474	0.3375	0.4813	0.2556	0.3730	0.5019	0.3893	0.4649	0.2741
RecDiff [2024]	0.2917	0.4549	0.3694	0.5242	0.2605	0.4115	0.5499	0.4369	0.5217	0.3013
HGSR [2024]	0.3359	0.5214	0.3928	0.5534	0.2688	0.4312	0.5405	0.4716	0.5392	0.3381
HDSR [2025]	0.3403	0.5391	0.4277	0.5693	0.2865	0.4506	0.5605	0.4824	0.5596	0.3483
TALLRec [2023]	0.2808	0.4953	0.1757	0.2444	0.2020	0.2622	0.4891	0.1596	0.2319	0.1889
EdGCL	**0.3596*** (↑ 0.0193)	**0.5576*** (↑ 0.0185)	**0.4369*** (↑ 0.0092)	**0.5867*** (↑ 0.0088)	**0.3051*** (↑ 0.0186)	**0.4755*** (↑ 0.0249)	**0.6136*** (↑ 0.0302)	**0.5014*** (↑ 0.0190)	**0.5797*** (↑ 0.0201)	**0.3674*** (↑ 0.0172)

Table 2: Recommendation performance comparison between EdGCL and existing models. The best results are bold and the second-best results are underlined. * indicates significant difference (p-value < 0.01 in the t-test).

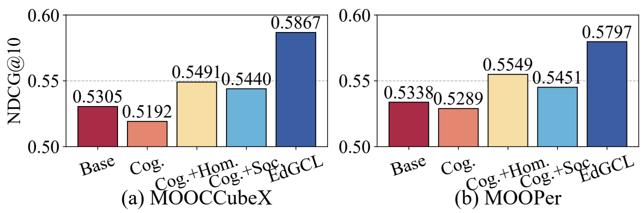


Figure 3: Ablation study from modeling perspectives.

Model	MOOCCubeX			MOOPer		
	R@10	N@10	MRR	R@10	N@10	MRR
EdGCL	**0.6136**	**0.5797**	**0.3674**	**0.5576**	**0.5867**	**0.3051**
w/o CL	0.6023	0.5659	0.3595	0.5507	0.5793	0.3030
w/o SLP	0.5936	0.5594	0.3538	0.5391	0.5686	0.2939
w/o CD	0.5809	0.5435	0.3423	0.5489	0.5790	0.3015

Table 3: Ablation study from disentanglement perspectives.

ing four ablated variants: *Base* learns unified representations from a single graph constructed with both learning and social behaviors. *Cog.* learns cognitive representations solely from learning behaviors, without incorporating social behaviors. *Cog.+Hom.* extends *Cog.* by modeling individual- and community-level cognitive representations, but still based only on learning behaviors. *Cog.+Soc.* disentangles cognitive and social views without considering representations at both individual and community levels. The results are shown in Figure 3. EdGCL outperforms all variants, demonstrating the necessity of jointly modeling social and cognitive homophily, and that of designing both individual- and community-level representations.

Disentanglement Perspective. We further assess the impact of different disentanglement objectives by removing each one individually: *w/o CL* removes the cross-view contrastive loss, *w/o SLP* removes the social link prediction loss, and *w/o CD* removes the cognitive diagnosis loss. As shown in Table 3, removing any of the objectives leads to performance degradation across all metrics, confirming the complementary role of each disentanglement constraint. Notably, removing cognitive diagnosis leads to the largest degradation on MOOCCubeX, while removing social link prediction leads to the largest degradation on MOOPer. The results validate the effectiveness of disentanglement effects.

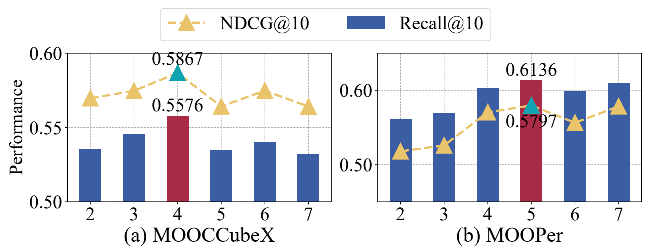


Figure 4: Varying the hop length K.

Hyperparameter Analysis

Impact of Hop Length. The hop length K defines the neighborhood range in the graph. Although larger K enables the model to capture higher-order structural signals, conventional GNN-based models typically limit K to 2 or 3 due to over-smoothing, where node representations become indistinguishable (Chen et al. 2020). In contrast, EdGCL leverages Graph Transformer to aggregate community-level representations from multi-hop sequences, effectively alleviating over-smoothing. As shown in Figure 4, we observe that EdGCL achieves the best performance for $K = 4$ on MOOCCubeX and $K = 5$ on MOOPer.

Impact of Loss Weights. Figure 5 illustrates the results of varying the loss weights of three auxiliary tasks: α for cognitive diagnosis, β for social link prediction, and γ for cross-view contrastive learning. We perform a grid search for each weight. As shown in the figure, moderate weight values lead

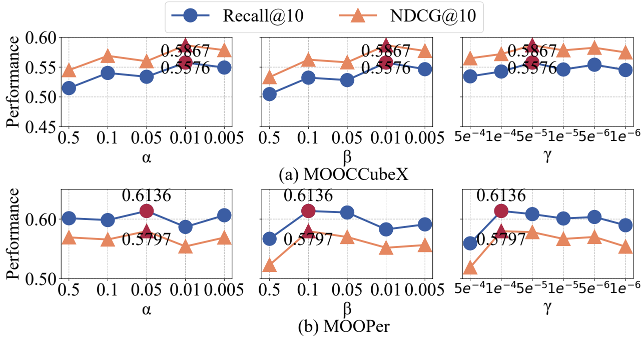


Figure 5: Varying the loss weights α, β, and γ.

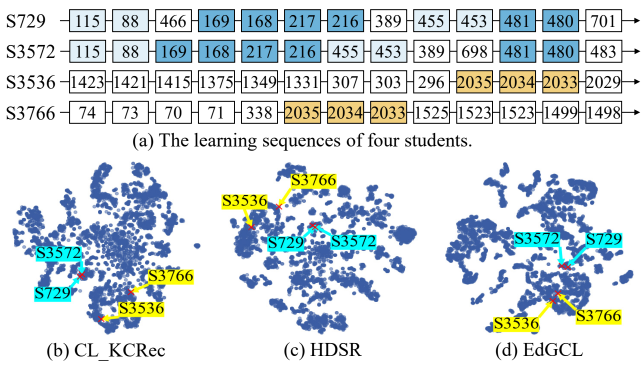


Figure 6: The t-SNE visualization of student representations.

to the best results. Specifically, EdGCL achieves the best performance with $\alpha = 0.01$, $\beta = 0.01$, and $\gamma = 5e^{-5}$ on MOOCCubeX, and with $\alpha = 0.05$, $\beta = 0.1$, and $\gamma = 1e^{-4}$ on MOOPer. These results highlight the importance of balancing the influence of each objective.

Case Study

To demonstrate the effectiveness of EdGCL in capturing homophily influence in both learning and social behaviors, we examine two pairs of students with social ties. As shown in Figure 6, the pair $(S729, S3572)$ exhibits high overlap in learning behaviors, while $(S3536, S3766)$ shows low overlap. But, $S3536$ and $S3766$ frequently interact with each other in the discussion forum. We visualize the student representations learned by the two best baselines CL_KCRec and HDSR as well as EdGCL using t-SNE. All models position $S729$ and $S3572$ very close, indicating their ability to capture the cognitive homophily in learning behaviors. However, only EdGCL positions $S3536$ and $S3766$ in a close distance, while the other models fail to position them close. This indicates that EdGCL effectively captures both social and cognitive homophily in learning and social behaviors.

Related Work

Graph-Based Educational Recommender Systems

Educational Recommender systems have attracted much attention (Da Silva et al. 2023), among which graph-based methods have been extensively studied (Wang et al. 2023; Zhang et al. 2023; Gong et al. 2023). For example, Gong et al. (2020) construct a heterogeneous information network (HIN) consisting of entities such as students, videos and knowledge concepts, and use the meta-paths of the HIN to guide the propagation of students' preferences. Wang et al. (2022) propose a hyperedge-based graph neural network to model students' relationships and use a modified graph attention network to learn course representations for recommendation. Guan et al. (2023) construct a knowledge graph for exercise recommendation while providing recommendation reasons. Wang et al. (2023) propose a graph-enhanced multi-objective method for computerized adaptive testing, ensuring concept diversity and question exposure control. Gu et al. (2024) utilize a relation-updated graph convolutional network and a stacked multi-channel graph neural network to model the explicit and implicit relations in the HIN for recommendation. The above methods have not considered the social homophily among students.

Social Recommendation

Social Recommendation improves traditional recommender systems by exploiting social relations among users (Tang, Hu, and Liu 2013). Recent studies have widely leveraged graph neural networks for social recommendation (Fan et al. 2019; Yang et al. 2024a; Sharma et al. 2024). For instance, Huang et al. (2021) jointly inject the inter-dependent knowledge across items and users into recommendations, enabling global graph structure awareness. Wang, Xia, and Huang (2023) learn graph neural relations from both interaction and social views, and alleviate the noisy effects of transferring social knowledge into interaction modeling. Xia et al. (2023) design a disentangled graph neural network to maintain factorized representations for heterogeneous types of user and item connections. Yang et al. (2024b) exploit social structure with hyperbolic social embedding pre-training and treat the embeddings as additional features to learn user preference. Hu et al. (2025) introduce a hierarchical denoising method to tackle both intra-domain noise resulting from social trust relationships and inter-domain noise stemming from the entanglement of the latent factors over heterogeneous relations. To the best of our knowledge, we are the first to explicitly model social relations for educational recommender systems.

Conclusion

In this paper, we introduce EdGCL, a novel graph-based educational recommendation framework that explicitly disentangles social and cognitive homophily to enhance personalized learning experiences in online education platforms. EdGCL adopts graph transformers to enhance robust representation learning via attentive multi-hop aggregation and designs a cross-view contrastive disentanglement mechanism to push social-view representations away from cognitive-view representations, respectively. Experimental results on two large-scale real-world educational datasets demonstrate that EdGCL significantly outperforms state-of-the-art baselines across various evaluation metrics. In the future, we plan to extend the current study by investigating learning communities formed by social relations.

Acknowledgments

This work is supported by the National Key R&D program of China (2023YFC3341200) and the National Natural Science Foundation of China (Grant No. 62277017, 62137001, and 62507015).

References

Bai, Y.; Li, X.; Liu, Z.; Huang, Y.; Tian, M.; and Luo, W. 2025. Rethinking and improving student learning and forgetting processes for attention based knowledge tracing models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 27822–27830.

Bao, K.; Zhang, J.; Zhang, Y.; Wang, W.; Feng, F.; and He, X. 2023. Tallrec: An effective and efficient tuning framework to align large language model with recommendation. In *Proceedings of the 17th ACM conference on recommender systems*, 1007–1014.

Chen, C.-M. 2008. Intelligent web-based learning system with personalized learning path guidance. *Computers & Education*, 51(2): 787–814.

Chen, D.; Lin, Y.; Li, W.; Li, P.; Zhou, J.; and Sun, X. 2020. Measuring and relieving the over-smoothing problem for graph neural networks from the topological view. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 3438–3445.

Chen, J.; Gao, K.; Li, G.; and He, K. 2023. NAGphormer: A Tokenized Graph Transformer for Node Classification in Large Graphs. In *Proceedings of the International Conference on Learning Representations*.

Da Silva, F. L.; Slodkowski, B. K.; Da Silva, K. K. A.; and Cazella, S. C. 2023. A systematic literature review on educational recommender systems for teaching and learning: research trends, limitations and opportunities. *Education and information technologies*, 28(3): 3289–3328.

Fan, W.; Ma, Y.; Li, Q.; He, Y.; Zhao, E.; Tang, J.; and Yin, D. 2019. Graph neural networks for social recommendation. In *The world wide web conference*, 417–426.

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.

Gu, H.; Duan, Z.; Xie, P.; and Zhou, D. 2024. Modeling balanced explicit and implicit relations with contrastive learning for knowledge concept recommendation in MOOCs. In *Proceedings of the ACM Web Conference 2024*, 3712–3722.

Guan, Q.; Xiao, F.; Cheng, X.; Fang, L.; Chen, Z.; Chen, G.; and Luo, W. 2023. Kg4ex: An explainable knowledge graph-based approach for exercise recommendation. In *Proceedings of the 32nd ACM international conference on information and knowledge management*, 597–607.

Guo, D.; Yang, D.; Zhang, H.; Song, J.; Zhang, R.; Xu, R.; Zhu, Q.; Ma, S.; Wang, P.; Bi, X.; et al. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*.

Hjelm, R. D.; Fedorov, A.; Lavoie-Marchildon, S.; Grewal, K.; Bachman, P.; Trischler, A.; and Bengio, Y. 2018. Learning deep representations by mutual information estimation and maximization. In *International Conference on Learning Representations*.

Hu, Z.; Nakagawa, S.; Zhuang, Y.; Deng, J.; Cai, S.; Zhou, T.; and Ren, F. 2025. Hierarchical Denoising for Robust Social Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 37(2): 739–753.

Huang, C.; Xu, H.; Xu, Y.; Dai, P.; Xia, L.; Lu, M.; Bo, L.; Xing, H.; Lai, X.; and Ye, Y. 2021. Knowledge-aware coupled graph neural network for social recommendation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, 4115–4122.

Jing, M.; Zhu, Y.; Zang, T.; and Wang, K. 2023. Contrastive self-supervised learning in recommender systems: A survey. *ACM Transactions on Information Systems*, 42(2): 1–39.

Li, Z.; Xia, L.; and Huang, C. 2024. Recdiff: Diffusion model for social recommendation. In *Proceedings of the 33rd ACM international conference on information and knowledge management*, 1346–1355.

Liu, K.; Zhao, X.; Tang, J.; Zeng, W.; Liao, J.; Tian, F.; Zheng, Q.; Huang, J.; and Dai, A. 2021. MOOPer: a large-scale dataset of practice-oriented online learning. In *China Conference on Knowledge Graph and Semantic Computing*, 281–287. Springer.

Liu, S.; Yu, X.; Ma, H.; Wang, Z.; Qin, C.; and Zhang, X. 2023. Homogeneous cohort-aware group cognitive diagnosis: A multi-grained modeling perspective. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 4094–4098.

Sharma, K.; Lee, Y.-C.; Nambi, S.; Salian, A.; Shah, S.; Kim, S.-W.; and Kumar, S. 2024. A survey of graph neural networks for social recommender systems. *ACM Computing Surveys*, 56(10): 1–34.

Shen, J.; Qian, H.; Zhang, W.; and Zhou, A. 2024. Symbolic cognitive diagnosis via hybrid optimization for intelligent education systems. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, 14928–14936.

Shen, S.; Liu, Q.; Chen, E.; Huang, Z.; Huang, W.; Yin, Y.; Su, Y.; and Wang, S. 2021. Learning process-consistent knowledge tracing. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*, 1452–1460.

Tang, J.; Hu, X.; and Liu, H. 2013. Social recommendation: a review. *Social network analysis and mining*, 3(4): 1113–1133.

Van Merrienboer, J. J.; and Ayres, P. 2005. Research on cognitive load theory and its design implications for e-learning. *Educational Technology Research and Development*, 53(3): 5–13.

Wang, H.; Long, T.; Yin, L.; Zhang, W.; Xia, W.; Hong, Q.; Xia, D.; Tang, R.; and Yu, Y. 2023. GMOCAT: A graph-enhanced multi-objective method for computerized adaptive testing. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2279–2289.

Wang, T.; Xia, L.; and Huang, C. 2023. Denoised Self-Augmented Learning for Social Recommendation. In *IJCAI International Joint Conference on Artificial Intelligence*.

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.

Xia, L.; Shao, Y.; Huang, C.; Xu, Y.; Xu, H.; and Pei, J. 2023. Disentangled graph social recommendation. In *2023 IEEE 39th international conference on data engineering (ICDE)*, 2332–2344. IEEE.

Yang, Y.; Huang, C.; Xia, L.; and Huang, C. 2023. Knowledge graph self-supervised rationalization for recommendation. In *Proceedings of the 29th ACM SIGKDD conference on knowledge discovery and data mining*, 3046–3056.

Yang, Y.; Huang, C.; Xia, L.; and Li, C. 2022. Knowledge graph contrastive learning for recommendation. In *Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval*, 1434–1443.

Yang, Y.; Wu, L.; Wang, Z.; He, Z.; Hong, R.; and Wang, M. 2024a. Graph bottlenecked social recommendation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 3853–3862.

Yang, Y.; Wu, L.; Zhang, K.; Hong, R.; Zhou, H.; Zhang, Z.; Zhou, J.; and Wang, M. 2024b. Hyperbolic Graph Learning for Social Recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 36(12): 8488–8501.

Ying, C.; Cai, T.; Luo, S.; Zheng, S.; Ke, G.; He, D.; Shen, Y.; and Liu, T.-Y. 2021. Do transformers really perform badly for graph representation? *Advances in neural information processing systems*, 34: 28877–28888.

Yu, J.; Wang, Y.; Zhong, Q.; Luo, G.; Mao, Y.; Sun, K.; Feng, W.; Xu, W.; Cao, S.; Zeng, K.; et al. 2021. MOOCCubeX: a large knowledge-centered repository for adaptive learning in MOOCs. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 4643–4652.

Yu, M.; Ding, Z.; Yu, J.; Zhang, W.; Yang, M.; and Zhao, M. 2023. Graph contrastive learning with adaptive augmentation for knowledge concept recommendation. In *2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 1281–1286. IEEE.

Yu, X.; Qin, C.; Shen, D.; Yang, S.; Ma, H.; Zhu, H.; and Zhang, X. 2024. Rigl: A unified reciprocal approach for tracing the independent and group learning processes. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 4047–4058.

Zhang, H.; Shen, S.; Xu, B.; Huang, Z.; Wu, J.; Sha, J.; and Wang, S. 2024. Item-difficulty-aware learning path recommendation: From a real walking perspective. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 4167–4178.

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.