

# E2E-GRec: An End-to-End Joint Training Framework for Graph Neural Networks and Recommender Systems

Rui Xue  
 rxue@ncsu.edu  
 North Carolina State University  
 Raleigh , US

Guangmou Pan  
 TikTok Inc.  
 San Jose, US

Shichao Zhu  
 TikTok Inc.  
 San Jose, US

Yang Song  
 TikTok Inc.  
 San Jose, US

Liang Qin  
 TikTok Inc.  
 Chengdu, CN

Tianfu Wu  
 North Carolina State University  
 Raleigh, US

## Abstract

Graph Neural Networks (GNNs) have emerged as powerful tools for modeling graph-structured data and have been widely used in recommender systems, such as for capturing complex user-item and item-item relations. However, most industrial deployments adopt a two-stage pipeline: GNNs are first pre-trained offline to generate node embeddings, which are then used as static features for downstream recommender systems. This decoupled paradigm leads to two key limitations: (1) high computational overhead, since large-scale GNN inference must be repeatedly executed to refresh embeddings; and (2) lack of joint optimization, as the gradient from the recommender system cannot directly influence the GNN learning process, causing the GNN to be suboptimally informative for the recommendation task. In this paper, we propose **E2E-GRec**, a novel **end-to-end** training framework that unifies GNN training with the recommender system. Our framework is characterized by three key components: (i) efficient subgraph sampling from a large-scale cross-domain heterogeneous graph to ensure training scalability and efficiency; (ii) a Graph Feature Auto-Encoder (GFAE) serving as an auxiliary self-supervised task to guide the GNN to learn structurally meaningful embeddings; and (iii) a two-level feature fusion mechanism combined with Gradnorm-based dynamic loss balancing, which stabilizes graph-aware multi-task end-to-end training. Extensive offline evaluations, online A/B tests (e.g., a +0.133% relative improvement in stay duration, a 0.3171% reduction in the average number of videos a user skips) on large-scale production data, together with theoretical analysis, demonstrate that E2E-GRec consistently surpasses traditional approaches, yielding significant gains across multiple recommendation metrics.

## ACM Reference Format:

Rui Xue, Shichao Zhu, Liang Qin, Guangmou Pan, Yang Song, and Tianfu Wu. 2018. E2E-GRec: An End-to-End Joint Training Framework for Graph Neural Networks and Recommender Systems. In *Proceedings of Make sure*

\*This work was done while the author was an intern at TikTok Inc.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX, Woodstock, NY

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-XXXX-X/2018/06  
<https://doi.org/XXXXXXX.XXXXXXX>

*to enter the correct conference title from your rights confirmation email (Conference acronym 'XX). ACM, New York, NY, USA, 11 pages. <https://doi.org/XXXXXXX.XXXXXXX>*

## 1 Introduction

Recommender systems have become an indispensable component of modern digital platforms, driving personalized content delivery across e-commerce, social media, and video-streaming services [1, 17, 37]. As users are increasingly overwhelmed by massive and continuously expanding content catalogs, the ability to accurately model user preferences directly impacts user engagement, retention, and overall platform revenue [7, 40].

Graph Neural Networks (GNNs) have emerged as a powerful paradigm for capturing the complex relational structures inherent in graph-structured data [9, 11, 14, 21, 24, 27, 28, 31, 32] and have achieved remarkable success across multiple domains, including biological networks, transportation infrastructures, and recommender systems [8, 18, 19, 25, 29, 30, 36]. By iteratively aggregating information from neighboring nodes, GNNs can effectively model various relations to learn expressive, high-order representations that capture both local neighborhood and global patterns. These learned representations go far beyond traditional pairwise similarities, enabling GNNs to uncover latent collaborative signals through multi-hop message passing. This capability has made GNNs particularly attractive for recommendation scenarios involving intricate user-item and item-item dependencies, where traditional models often fail to fully capture the underlying relational semantics [23, 35].

However, despite the recent success, most existing industrial applications of GNNs in recommender systems still adopt a two-stage training paradigm: GNNs are first trained offline to generate static item or user embeddings, which are then fed into a downstream recommendation model. Although this decoupled design simplifies implementation and enables each component to be optimized independently, it introduces two fundamental drawbacks: First, *high computational overhead*. Since GNN embeddings are generated offline, the inference must be repeatedly executed to refresh embeddings for all nodes as the underlying graph evolves. In dynamic environments such as TikTok's recommendation system, where new content and user interactions emerge continuously, this process necessitates frequent recomputation of billions of node embeddings, resulting in substantial infrastructure costs and increased latency. Second and more critically, *lack of joint optimization*. In

the two-stage paradigm, and the recommendation system cannot provide direct gradient feedback to refine the GNN’s learned representations. This disconnect limits the expressive power of the overall system and often results in suboptimal performance.

To address these challenges, we propose **E2E-GRec**, a novel end-to-end joint training framework that unifies GNN representation learning and recommendation modeling. Unlike traditional approaches, E2E-GRec first samples a subgraph from a heterogeneous cross-domain graph using importance sampling. Then, to mitigate task misalignment, we introduce a Graph Feature Auto-Encoder (GFAE) as a complementary self-supervised learning (SSL) objective that reconstructs node features, thereby providing a direct and stable learning signal to guide the learning of GNN. More importantly, to enable effective gradient flow from the recommendation model to the GNN in end-to-end training, we first employ two-level feature fusion that strengthens the interaction between graph representations and recommendation features. We then adopt Gradnorm to adaptively balance the graph learning and recommendation gradients over shared parameters, preventing task dominance and ensuring stable convergence. Our framework is built upon four key contributions:

- **Efficient subgraph construction.** Task-specific subgraphs are efficiently extracted from large-scale cross-domain graphs via importance sampling, enabling scalable and adaptive graph construction tailored to recommendation tasks.
- **GNN Multi-Task learning.** A Graph Feature Auto-Encoder (GFAE) is introduced as an auxiliary SSL objective, guiding the GNN to learn structurally meaningful embeddings that are subsequently jointly optimized with the recommendation objectives.
- **Two levels feature fusion and dynamic loss balancing.** GNN representations are integrated with recommendation features at both bottom and upper levels through complementary gating and attention based fusion strategies. Gradnorm is further employed to dynamically balance the SSL and recommendation losses, mitigating dominance and ensuring stable convergence.
- **Comprehensive validation.** Extensive theoretical analysis, offline experiments, online A/B testing, and ablation studies collectively validate the effectiveness of the framework. Significant improvements observed on real-world production dataflow highlight its strong practical value.

## 2 Related Work

### 2.1 Recommender Systems

Recommender systems are the cornerstone of modern industrial platforms such as e-commerce, online advertising, and content delivery [7, 37]. A typical recommendation pipeline consists of three stages: candidate generation, ranking, and re-ranking. The candidate generation stage retrieves a small subset of potentially relevant items from billions of candidates using collaborative filtering, approximate nearest neighbor search, or large-scale retrieval models [10, 34]. The ranking stage then scores these candidates with more sophisticated models to optimize engagement metrics such as click-through rate (CTR) or dwell time. Finally, the re-ranking stage refines the results by incorporating business objectives such as diversity, fairness, and monetization [26].

Learning-to-Rank (LTR) approaches have become fundamental in the ranking stage of industrial recommendation systems, where the primary goal is to optimize the ordering of candidate items presented to users. Within this component, traditional pointwise approaches such as logistic regression and gradient boosted decision trees (GBDT) have been widely adopted in industry, with frameworks like XGBoost and LightGBM showing strong performance in production systems at LinkedIn and Microsoft [5, 13]. Pairwise methods including RankNet [2] and LambdaMART [3] learn relative item preferences, while listwise methods such as ListNet [4] and SoftRank [20] directly optimize ranking metrics (e.g., NDCG), achieving superior results in large-scale commercial systems.

### 2.2 Graph Neural Networks in Industrial Recommender Systems

Graph Neural Networks (GNNs) have emerged as powerful tools to capture higher-order connectivity and context in recommendation scenarios, where user-item interactions form complex graphs. Early industrial adoptions include PinSage [35] deployed at Pinterest, which scales GraphSAGE with random-walk sampling and neighborhood importance weighting to generate web-scale item embeddings. Similarly, LightGCN [12] simplify GCNs by focusing on user-item bipartite structures, achieving state-of-the-art performance in large-scale recommendation benchmarks.

In practice, large companies have developed proprietary GNN-based recommendation frameworks. AliGraph [41] at Alibaba and ByteGNN [38] at ByteDance introduce distributed training and heterogeneous graph modeling to support billion-scale data, ensuring both offline training efficiency and online serving latency. Other works such as Euler (Alibaba) and DistDGL (Amazon) [39] focus on distributed GNN training frameworks to meet industrial throughput requirements. However, most industrial GNN-based recommendation models are not trained in a fully end-to-end manner. In such pipelines, the ranking gradients cannot directly influence the GNN parameters, leading to a mismatch between the objectives used for graph representation learning and the final goals optimized in ranking or re-ranking. This objective misalignment often results in suboptimal performance.

Although a few studies have attempted to enable end-to-end training by cascading GNNs with ranking models [15], these approaches typically treat the GNN outputs as augmented features that are fused with the ranking network. Consequently, the downstream task remains mismatched—while the GNN component focuses on capturing structural semantics, the ranking model instead aims to optimize task-specific metrics such as click-through rate (CTR) or relevance ranking scores.

## 3 Preliminary

*Notation.* Let  $G = (V, E)$  be a graph with  $|V| = n$  nodes and  $|E| = m$  edges. We denote by  $X \in \mathbb{R}^{n \times d}$  the node-feature matrix, where the  $i$ -th row  $x_i^\top$  is the  $d$ -dimensional feature of node  $i$ . The adjacency matrix is  $A \in \mathbb{R}^{n \times n}$ , with  $A_{ij} > 0$  iff  $(i, j) \in E$ ;  $D = \text{diag}(A\mathbf{1})$  is the degree matrix. We use  $\tilde{A} = A + I$  and  $\tilde{D} = D + I$  for self-loop augmentation, and write the symmetric normalized adjacency matrix  $\hat{A} = \tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}$  and the normalized Laplacian  $L = D - A /$

$\hat{L} = I - \hat{A}$ . For a node  $u$ , its neighborhood is  $\mathcal{N}(u) = \{v : (u, v) \in E\}$ , and the  $L$ -hop neighborhood is  $\mathcal{N}^{(L)}(u)$  defined recursively. When neighbor sampling is used,  $\mathcal{S}_\ell(u) \subseteq \mathcal{N}(u)$  denotes the sampled neighbor multiset at layer  $\ell$ . A  $L$ -layer GNN encoder is  $f_\theta : (X, A) \mapsto H^{(L)} \in \mathbb{R}^{n \times d_L}$ , with hidden representations  $H^{(\ell)} = [h_1^{(\ell)}; \dots; h_n^{(\ell)}]$  and  $H^{(0)} = X$ . A generic message-passing layer is

$$h_u^{(\ell)} = \psi^{(\ell)} \left( h_u^{(\ell-1)}, \text{Agg}_{v \in \mathcal{N}(u)}(h_v^{(\ell-1)}, A_{uv}) \right), \quad (1)$$

where  $\text{Agg}$  is an aggregation operator (e.g., mean/sum/max/attention), and  $\psi^{(\ell)}$  the update function.

### 3.1 GNN for Recommendations

In this section, we introduce two classical GNN architectures that are employed in our implementation owing to their well-established advantages in both performance and efficiency. These models have been extensively validated in prior research and are widely adopted in large-scale recommendation systems.

**GraphSAGE[11].** GraphSAGE performs inductive neighborhood aggregation with learnable transformations and (optionally) sampled neighbors. The initial node representation is set as  $h^{(0)} = X$  when raw features are available; otherwise we initialize  $h^{(0)}$  via a trainable embedding lookup. At layer  $l$ , for a node  $u$  we sample a multiset  $\mathcal{S}_\ell(u) \subseteq \mathcal{N}(u)$  of size  $m_\ell$  and compute

$$\bar{h}_{\mathcal{N}(u)}^{(l)} = \frac{1}{|\mathcal{S}_\ell(u)|} \sum_{v \in \mathcal{S}_\ell(u)} h_v^{(l)}, \quad (2)$$

$$h_u^{(l+1)} = \sigma \left( W^{(l)} \cdot \text{CONCAT} \left( h_u^{(l)}, \bar{h}_{\mathcal{N}(u)}^{(l)} \right) \right), \quad (3)$$

where  $W^{(l)}$  is a trainable matrix and  $\sigma(\cdot)$  a nonlinearity. Other aggregators (sum/max/LSTM/attention) can be substituted for the mean operators.

**LightGCN [12].** LightGCN simplifies the design of graph convolution by discarding feature transformation and nonlinear activations, and focuses only on neighborhood aggregation.  $h^{(0)}$  is initialized by a trainable embedding lookup table. At layer  $l$ , the embeddings of a user  $u$  and an item  $i$  are updated by averaging over their neighbors with degree normalization:

$$h_u^{(l+1)} = \sum_{i \in \mathcal{N}(u)} \frac{1}{\sqrt{|\mathcal{N}(u)|} \sqrt{|\mathcal{N}(i)|}} h_i^{(l)}, \quad (4)$$

$$h_i^{(l+1)} = \sum_{u \in \mathcal{N}(i)} \frac{1}{\sqrt{|\mathcal{N}(i)|} \sqrt{|\mathcal{N}(u)|}} h_u^{(l)}. \quad (5)$$

In matrix form, the propagation can be compactly written as

$$H^{(l+1)} = \hat{A} H^{(l)}, \quad l = 0, \dots, L-1, \quad (6)$$

$$Y = \frac{1}{L+1} \sum_{l=0}^L H^{(l)}, \quad (7)$$

where  $Y$  is the final representation obtained by averaging embeddings across all layers.

### 3.2 End-to-end Training

In most industrial applications of GNNs for recommender systems, the GNN component is typically trained *offline*, and the resulting node embeddings are then used as augmented features for the downstream ranking model [16, 22, 35]. However, in such decoupled architectures, the gradients from the ranking model cannot be back-propagated into the GNN, resulting in completely decoupled optimization between the GNN encoder and the recommendation model. This gradient isolation leads to several drawbacks: (i) the GNN is optimized solely for structural reconstruction rather than task-specific ranking discrimination; (ii) the feature space learned by the GNN may not align well with the recommendation model's objectives; and (iii) updating the GNN embeddings requires costly offline retraining whenever the user-item distribution changes.

Here, we present a theorem with a formal proof regarding gradients to theoretically demonstrate why end-to-end learning is critical for aligning the GNN and recommendation objectives. The proof can be found in Appendix A.

**Theorem 1** (Gradient Coupling in E2E-GRec). *Let  $h_\theta(\cdot; \mathcal{G})$  denote the GNN embeddings with parameters  $\theta$ ,  $z_i = [h_\theta(x_i; \mathcal{G}) \| b_i]$ ,  $s_\psi$  be the recommendation scorer and*

$$J(\theta, \psi, \alpha) = \frac{1}{n} \sum_{i=1}^n \left[ \alpha_1 L_{\text{rec}}(y_i, s_\psi(z_i)) + \alpha_2 L_{\text{gnn}}(\theta) \right] \quad (8)$$

*Assume  $\frac{\partial s_\psi(z)}{\partial \theta_\theta} \neq 0$ . Then:*

- (i)  $\nabla_\theta J$  contains a nonzero contribution from  $L_{\text{rec}}$  ( $\text{Rec} \rightarrow \text{GNN}$ );
- (ii)  $\frac{\partial}{\partial \theta} (\nabla_\psi J) \neq 0$ , i.e., the recommendation model gradient depends on  $\theta$  ( $\text{GNN} \rightarrow \text{Rec}$ ).

*In contrast, cascaded pipelines satisfy  $\nabla_\theta J_{\text{rec}} = 0$  and  $\frac{\partial}{\partial \theta} (\nabla_\psi J) = 0$ , preventing capture of higher-order graph-rec interactions.*

## 4 Methodology

Motivated by this potential drawback, we introduce our **E2E-GRec** : End-to-End Joint Training Framework for Graph Neural Networks and Recommender Systems in this section. We begin by describing our subgraph construction process from the cross-domain graph in Section 4.1. Next, we discuss the potential challenges of integrating GNNs with recommendation systems. We then propose our multi-task learning GNNs designed to address these issues in Section 4.2. Finally, we present the core component of our framework—the effective integration of GNNs and recommendation systems through joint optimization and feature fusion in Section 4.3.

### 4.1 Subgraph Construction from Cross Domain Graph

To comprehensively capture complex user interests and content associations across diverse business scenarios, we first construct a large-scale weighted cross-domain heterogeneous graph (**CDG**) [42] as the foundation for our subgraph sampling process. Formally, the CDG integrates multi-type nodes and edges across multiple business domains, including video, search, live streaming, and e-commerce. The graph incorporates two categories of relational connections. The first is explicit connections, which are directly derived from observable user behaviors—such as likes, shares, and follows in the video (item) scenario. The second is implicit connections,

mined through algorithms that uncover semantic relationships between nodes to capture higher-order behavioral co-occurrences. A key advantage of this heterogeneous graph design lies in its flexibility. Unlike traditional graph learning approaches that rely on manually defined meta-paths requiring domain expertise—and are therefore difficult to adapt to rapidly evolving business environments—our CDG can flexibly incorporate scenario-specific relation types or meta-paths that align with different business objectives and evolving recommendation contexts, without being constrained by fixed or predefined connection rules.

In this work, we focus specifically on the implicit item-to-item (i2i) relation construction as the core foundation. This is because, first, i2i graphs have consistently demonstrated effectiveness, low noise, and high quality in our prior experiments on recommendation systems (e.g., retrieval stage). In addition, given the large-scale industrial setting, we construct homogeneous item-to-item relations, which is more beneficial for maintaining overall model, data, and system simplicity while ensuring training and serving efficiency. Nevertheless, as emphasized, different business scenarios and stages can flexibly introduce additional meta-paths or relation types to build graphs according to their specific needs.

Specifically, we built this i2i relations based on the Swing [33] algorithm, which models substitute relationships by exploring user–item–user–item “swing” structures in the interaction graph. For items  $i$  and  $j$ , Swing traverses all user pairs  $(u, v)$  who have clicked both items, with each user pair contributing a score of  $\frac{1}{\alpha + |I_u \cap I_v|}$ , where  $|I_u \cap I_v|$  represents the number of common items clicked by both users. This design cleverly penalizes user pairs with overly overlapping click patterns, as they might just be highly active users browsing randomly. Additionally, the algorithm introduces a user weighting factor  $w_u = \frac{1}{\sqrt{|I_u|}}$  to reduce the influence of highly active users. Compared with traditional similarity measures, Swing leverages richer structural information and effectively filters out noisy data and captures more reliable product similarity relationships, constructing high-quality product substitute relationship graphs.

Due to the large-scale nature of the CDG, leveraging the complete graph is computationally prohibitive for online serving. Hence, we sample i2i subgraphs from the CDG, which is refreshed daily from real-time data pipeline to capture the evolving item relationships. To address this, we adopt an importance sampling strategy that builds multi-hop subgraphs for both training and inference. Specifically, given a source item  $u$  which is from input sequence and sampling depth  $L$ , the probability of sampling a neighbor  $v \in \mathcal{N}(u)$  at layer  $\ell$  is defined as

$$P_{u \rightarrow v}^{(\ell)} = \frac{(\omega_{uv})^\beta}{\sum_{k \in \mathcal{N}(u)} (\omega_{uk})^\beta}, \quad (9)$$

where  $\omega_{uv}$  denotes the edge weight (e.g., Swing score) between items  $u$  and  $v$ ,  $\mathcal{N}(u)$  is the neighbor set of  $u$  in the CDG, and  $\beta \geq 0$  is a temperature parameter that controls the sampling concentration. The number of sampled neighbors is further constrained by a hyperparameter  $k$ , ensuring tractable time complexity. Note that the subgraph is constructed in real time, which strikes a balance between graph completeness and computational efficiency, enabling scalable and responsive recommendations.

## 4.2 GNN Multi-Task Co-Optimization

Given a subgraph, the most straightforward way to achieve end-to-end training is to place a GNN before the recommendation model, forming a supervised cascaded architecture [15]. Specifically, the subgraph is first fed into the GNN, and the resulting node embeddings are concatenated with other features as inputs to the downstream recommendation model. While this approach is simple and easy to implement, it suffers from several significant drawbacks:

- **Long backpropagation path yields superficial gradient signal:** In this setup, the GNN receives its learning signal solely from the final recommendation objective. This supervision must propagate backward through the entire recommendation module, leading to weak guidance for the GNN module. As the gradient passes through multiple nonlinear transformations, it tends to vanish or become unstable, resulting in a shallow and noisy signal by the time it reaches the GNN parameters. Formally, the gradient reaching GNN parameters is computed as:

$$\frac{\partial \mathcal{L}_{\text{rec}}}{\partial \theta_{\text{GNN}}} = \underbrace{\frac{\partial \mathcal{L}_{\text{rec}}}{\partial o}}_{\text{loss-to-output}} \cdot \underbrace{\frac{\partial o}{\partial z}}_{J_g(z) = \prod_{l=1}^L \frac{\partial f_l}{\partial f_{l-1}}} \cdot \underbrace{\frac{\partial z}{\partial \theta_{\text{GNN}}}}_{\text{GNN-internal Jacobian}} \quad (10)$$

where  $o = g(z) = f_L \circ f_{L-1} \circ \dots \circ f_1(z)$  denotes the recommendation model composed of  $L$  nonlinear layers. This product of Jacobians can either vanish (if singular values  $< 1$ ) or explode (if singular values  $> 1$ ), making training unstable.

- **Potential Task Misalignment between GNN and Recommendation:** The second critical issue is the inherent task misalignment between GNN and recommendation systems. The primary objective of a recommendation system is to present users with an optimal set or ordering of items that maximizes user satisfaction or engagement. This task is inherently discriminative and relative—it focuses on comparing candidate items to identify and rank those most relevant to a user’s preferences. In contrast, GNN aims to learn meaningful graph topology representations and node embeddings that capture structural properties such as local neighborhood patterns and graph properties, which is different from recommendation systems.

According to the above analysis, we need to design an effective auxiliary graph learning task to train the GNN, rather than simply cascading it with the recommendation module. However, since explicit labels for training the GNN on a specific graph learning task are unavailable, we propose a self-supervised learning (SSL) task specifically formulated as a generative objective that reconstructs node features. This auxiliary SSL objective provides a stable and informative learning signal for the GNN, allowing it to capture the intrinsic structural semantics of the graph. We first introduce the technical details of our design, followed by a theoretical analysis that demonstrates the advantages over cascaded supervised learning frameworks.

**Graph Feature Auto-Encoder (GFAE)** Graph self-supervised learning (SSL) methods can be broadly categorized into contrastive and generative approaches. Contrastive methods aim to bring representations of similar nodes (or augmented graph views) closer together while pushing apart those of dissimilar ones. In contrast, generative approaches train models to reconstruct original node features or predict the existence or weights of edges.

Since our goal is to enable end-to-end learning while maintaining efficiency for online serving, we adopt the generative paradigm via a graph auto-encoder framework. This choice is motivated by the fact that contrastive learning typically requires complex training strategies—such as the careful construction of negative samples and reliance on high-quality data augmentations [31]—which can hinder scalability and deployment efficiency.

Specifically, we employ a variant of the **Graph Auto-Encoder (GAE)**. Instead of reconstructing the adjacency matrix, we adopt a mean squared error (MSE) loss to reconstruct the input node features.

This design is motivated by the need to obtain more informative node representations that can be effectively concatenated with recommendation features for joint optimization. (See Section 4.3). Formally, given an input graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with node features  $X \in \mathbb{R}^{n \times d}$ , our Graph feature Auto-Encoder consists of an encoder  $f_\theta(\cdot)$  and a decoder  $g_\phi(\cdot)$  with the reconstruction loss as:

$$\mathcal{L}_{\text{ssl}} = \|X - \hat{X}\|_F^2 = \|X - g_\phi(f_\theta(X, A))\|_F^2. \quad (11)$$

In traditional settings,  $f_\theta$  is often implemented as a GNN (e.g., LightGCN or GraphSAGE), and the decoder  $g_\phi$  reconstructs the node features  $\hat{X}$  through an MLP. To further reduce computational complexity, we adopt a lightweight decoder design and set  $g_\phi(\cdot)$  to be the identity mapping:

$$f_\theta(X, A) = \text{GNN}(X, A), \quad g_\phi(Z) = Z, \quad (12)$$

Despite its simplicity, this configuration empirically achieves strong performance in our experiments, indicating that the encoder (GNN) itself learns sufficiently expressive node embeddings to recover the original features without requiring an additional decoder.

Next, we provide a theoretical result to demonstrate why introducing a separate SSL task for the GNN is more effective than simply using the GNN output as an augmented feature and training the downstream recommendation model in a cascaded manner:

**Theorem 2** (SSL vs. Cascaded: Objective Misalignment). *Let  $X \in \mathbb{R}^{n \times d}$  be node features and  $Z = f_\theta(X) \in \mathbb{R}^{n \times k}$  ( $k \geq d$ ) the encoder output. Assume a linear decoder  $W_d \in \mathbb{R}^{k \times d}$  with full column rank  $\text{rank}(W_d) = d$ . If the SSL reconstruction objective attains zero loss,*

$$\|X - ZW_d\|_F^2 = 0, \quad (13)$$

then  $\text{col}(X) \subseteq \text{col}(Z)$  (i.e.,  $Z$  preserves the full feature subspace of  $X$ ). For simplicity, consider the cascaded recommendation head with BPR scores  $s_i = w^\top z_i$  for some  $w \in \mathbb{R}^k$  and

$$\mathcal{L}_{\text{BPR}}(Z) = \sum_{(i,j) \in \mathcal{P}} \ell(w^\top(z_i - z_j)), \quad (14)$$

where  $\ell'(\cdot)$  exists. If  $\text{col}(X) \not\subseteq \text{span}(w)$ , then the two objectives are misaligned: there exists a nontrivial subspace

$$\mathcal{U} = \text{col}(X) \cap \ker(w^\top) \neq \{0\} \quad (15)$$

such that SSL requires  $Z$  to preserve information along  $\mathcal{U}$ , while BPR is invariant to any perturbations  $\{\delta z_i\}$  with  $\delta z_i \in \ker(w^\top) = \{v \in \mathbb{R}^k : w^\top v = 0\}$  for all  $i$  (hence supplies no gradient on  $\mathcal{U}$ ). Therefore BPR constrains only the one-dimensional subspace  $\text{span}(w)$  and leaves the  $(k-1)$ -dimensional orthogonal subspace  $\ker(w^\top)$  unconstrained, establishing the misalignment.

Our theorem theoretically demonstrates that the proposed joint training framework yields more informative embeddings. The proof can be found in Appendix B.

### 4.3 End-to-End Joint Training for GNN and Recommendation Systems

As introduced above, the key challenge lies in effectively integrating GNN components with recommendation systems. To achieve this, we combine the two modules from two complementary perspectives: *loss combination* and *feature fusion*. Note that our proposed framework can be applied to various stages of recommendation systems. In this work, we specifically focus on the Learning-to-Rank (LTR) component, which serves as one of the fundamental modules in the ranking stage of industrial recommendation pipelines. An overview of the proposed framework is presented in Figure 1.

**4.3.1 Multi-Task Loss Combination.** To enable end-to-end optimization between the GNN encoder and the LTR model, we jointly optimize two complementary objectives: the self-supervised reconstruction loss  $\mathcal{L}_{\text{SSL}}$  for GNN representation learning and the supervised loss  $\mathcal{L}_{\text{LTR}}$  for downstream prediction.

**Overall Loss.** The total objective is a weighted combination of the two tasks:

$$\mathcal{L}_{\text{total}} = w_{\text{SSL}}(t) \mathcal{L}_{\text{SSL}} + w_{\text{LTR}}(t) \mathcal{L}_{\text{LTR}}, \quad (16)$$

where  $w_{\text{SSL}}(t)$  and  $w_{\text{LTR}}(t)$  are dynamic task weights at training step  $t$ . Because of the dynamic update process, the scale of the losses changes over time. Naively fixing  $w_{\text{SSL}}$  and  $w_{\text{LTR}}$  across all time steps may let one loss dominate the other. To avoid this, we apply Gradnorm [6] for adaptive gradient balancing based on shared parameter space.

**Gradient-Based Task Weighting.** For each task  $i \in \{\text{SSL, LTR}\}$ , we define its gradient norm with respect to the shared parameters  $\theta_s$  (the input source node feature  $h^{(0)}$  in our case):

$$G_i(t) = \|\nabla_{\theta_s}(w_i(t) \mathcal{L}_i(t))\|_2. \quad (17)$$

The goal of GradNorm is to ensure all tasks train at a similar rate by adjusting  $w_i(t)$  to maintain proportional gradient magnitudes. The relative training rate of task  $i$  is computed as:

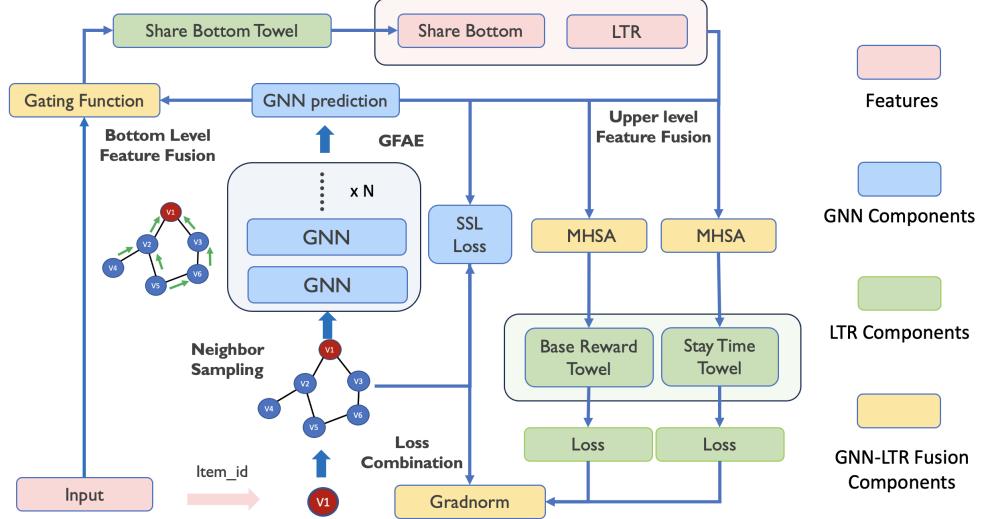
$$r_i(t) = \frac{\mathcal{L}_i(t)/\mathcal{L}_i(0)}{\frac{1}{T} \sum_{j=1}^T \mathcal{L}_j(t)/\mathcal{L}_j(0)}, \quad (18)$$

where  $\mathcal{L}_i(0)$  is the initial loss and  $T$  is the total number of tasks. Then, Gradnorm minimizes the following objective to align the gradient magnitudes:

$$\mathcal{L}_{\text{Gradnorm}} = \sum_{i \in \{\text{SSL, LTR}\}} |G_i(t) - \bar{G}(t) \cdot r_i(t)^\gamma|, \quad (19)$$

where  $\bar{G}(t)$  is the mean gradient norm across tasks, and  $\gamma$  controls how strongly faster tasks are slowed down. In our experiment,  $\gamma$  is typically set to 1, which has been found empirically effective for maintaining stable multi-task learning and balanced convergence rates. This mechanism adaptively increases  $w_i(t)$  for slower-learning tasks and decreases it for dominating ones, keeping both  $\mathcal{L}_{\text{SSL}}$  and  $\mathcal{L}_{\text{LTR}}$  at balanced learning rates.

**A Dual-Objective Optimization.** The complete training is a dual-objective optimization process where model parameters and task weights are updated in parallel based on different objectives. At



**Figure 1: Overview of E2E-GRec . Colors indicate distinct functional blocks. (1) Subgraph sampling: Given a source item ID, we sample  $k$ -hop neighbors to form a subgraph (Sec. 4.1); (2) GNN SSL: GNN trained via a graph autoencoder (Sec. 4.2); (3) GNN-LTR fusion: GNN representations are integrated and jointly optimized with LTR in an end-to-end manner. (Sec. 4.3)**

each training step, we perform two distinct updates: (1) *Model Parameter Update*: The parameters of the neural network,  $\theta = \{\theta_s, \theta_{\text{GNN}}, \theta_{\text{LTR}}\}$ , are updated by minimizing the main task loss,  $\min_{\theta} \mathcal{L}_{\text{total}}(t)$ . (2) *Task Weight Update*: The dynamic task weights,  $w_i = \{w_{\text{SSL}}, w_{\text{LTR}}\}$ , are updated by minimizing the Gradnorm meta-loss  $\min_{w_i > 0} \mathcal{L}_{\text{GradNorm}}(t)$ . Here  $w_i$  are normalized after each update to keep  $\sum_i w_i = T$ , following [6].

This joint optimization framework ensures that the GNN continuously refines its representations under supervision from both structural (SSL) and LTR signals. By separating the optimization goals, the model learns task-aligned embeddings in a more stable and balanced manner.

**4.3.2 Two-level Feature Fusion.** Besides the loss function, the GNN-derived embeddings play a crucial role, as they capture high-order collaborative filtering signals over the graph structure. To effectively inject this graph information into the recommendation model and enhance its capacity, we fuse the GNN embeddings with other feature sources (e.g., user profile features, LTR). As illustrated in Figure 1, we integrate the GNN-derived features at two hierarchical levels—the *bottom* and the *upper* levels. At the bottom level, we combine the GNN output with other input features and feed the fused representation into a shared bottom tower. The output of this tower serves as the fundamental feature foundation for the two downstream Learning-to-Rank (LTR) branches: the base-reward tower and the stay-time tower. At the upper level, we re-inject the GNN features together with the task-specific LTR feature along with the output from the shared bottom tower. This design provides an extra gradient flow that travels through a shorter path with less attenuation and enhances the model’s high-order perceptive capacity, enabling more expressive interactions between graph-derived and task-specific features across different semantic levels.

However, simple concatenation is insufficient to capture the complex interrelations between these features. To address this limitation, we explore two complementary fusion strategies: (1) **Concatenation with gating**, which provides simple yet effective feature-wise control, and (2) **Attention-based token fusion**, which treats each feature source as a token and performs multi-head attention across them. Let the GNN output be  $F_{\text{gnn}} \in \mathbb{R}^{N \times d}$  and the remaining feature groups  $\{F_{\text{shared}}, F_{\text{ltr}}\}$  each in  $\mathbb{R}^{N \times d}$ , where  $F_{\text{shared}} = \text{Tower}_{\text{share}}([F_{\text{in}}, F_{\text{gnn}}])$  is the output of the share bottom tower.

**1. Concatenation with Gating.** We first concatenate all feature groups together  $F_{\text{cat}} = \text{CONCAT}(F_{\text{shared}}, F_{\text{ltr}}, F_{\text{gnn}})$ . To adaptively control the contribution of different feature sources, we introduce a gating function that learns task-relevant feature weights:

$$g = \sigma(\mathbf{W}_g F_{\text{cat}} + \mathbf{b}_g), \quad F_{\text{fused}}^{(\text{gate})} = g \odot F_{\text{cat}} \quad (20)$$

where  $\mathbf{W}_g \in \mathbb{R}^{d \times d}$  and  $\mathbf{b}_g \in \mathbb{R}^d$  are learnable parameters,  $\sigma(\cdot)$  is the sigmoid activation, and  $\odot$  denotes element-wise multiplication. This gating mechanism allows the model to dynamically emphasize or suppress specific features based on their relevance to the task.

**2. Attention-based Token Fusion.** To capture inter-feature dependencies and further refine the fused representation, we apply Multi-Head Self-Attention(**MHSA**) over the stacked feature sequence  $F_{\text{stack}} = \text{STACK}(F_{\text{shared}}, F_{\text{ltr}}, F_{\text{gnn}})$ . Specifically, we treat each feature as an individual token:

$$F_{\text{fused}}^{(\text{attn})} = \text{Mean}_{[:, :, :]} (\text{MHSA}(F_{\text{stack}}) + F_{\text{stack}}), \quad F_{\text{stack}} \in \mathbb{R}^{N \times T \times d}, \quad (21)$$

where the residual connection ensures stable gradient flow. The refined features  $F_{\text{fused}}^{(\text{attn})}$  (or  $F_{\text{fused}}^{(\text{gate})}$ ) are then fed into downstream base-reward/stay-time towers:

$$F_{\text{reward}} = \text{Tower}_{\text{reward}}(F_{\text{fused}}), \quad F_{\text{stay}} = \text{Tower}_{\text{stay}}(F_{\text{fused}}). \quad (22)$$

$$F_{\text{final}} = \beta_r \cdot F_{\text{reward}} + \beta_s \cdot F_{\text{stay}}, \quad (23)$$

Here,  $F_{\text{reward}}$  and  $F_{\text{stay}}$  denote the outputs of the base-reward and stay-time towers, respectively, and their weighted sum forms the final LTR prediction  $F_{\text{final}}$ , where  $\beta_r$  and  $\beta_s$  is a dynamic weighting coefficient that balances the contribution between the base-reward prediction and the stay-time signal. The GNN parameters are optimized through a multi-task loss (also see Section 4.2):

$$\nabla_{\theta} \mathcal{L}_{\text{total}} = w_{\text{SSL}} \nabla_{\theta} \mathcal{L}_{\text{SSL}} + w_r J_{\theta \rightarrow Z}^{\top} \frac{\partial \mathcal{L}_{\text{reward}}}{\partial F_{\text{gnn}}} + w_s J_{\theta \rightarrow Z}^{\top} \frac{\partial \mathcal{L}_{\text{stay}}}{\partial F_{\text{gnn}}}, \quad (24)$$

In our experiments, the attention-based token fusion achieved the best overall performance, meanwhile, the gating-based fusion serves as an efficient alternative with latency constraints. Considering that the upper-level fusion directly precedes the task-specific towers and thus receives more immediate gradient signals, we adopt attention-based fusion at the upper level to maximize feature interaction and gating-based fusion at the bottom level provides efficient feature modulation while reducing computational overhead. Nevertheless, both fusion modules are fully interchangeable, allowing flexible adaptation to different deployment scenarios or efficiency requirements (such as online serving).

## 5 Experiments

In this section, we present comprehensive experiments, including offline AUC evaluations across days in Section 5.1, ablation studies on key components in Section 5.2, and online A/B testing results conducted within practical recommendation systems in Section 5.3, to demonstrate the effectiveness of our proposed E2E-GRec .

### 5.1 Offline Results

**Settings** As mentioned earlier, although our framework can be integrated into various stages of a recommendation system, we focus on the LTR stage in this work. The experimental configurations are summarized as follows:

- **Subgraph Sampling.** We consider two sampling strategies: (a) sampling one-hop neighbors, where 100 neighbors are sampled for each source node to construct the subgraph; and (b) sampling two-hop neighbors, where 25 and 15 neighbors are sampled for the first and second hops, respectively. During experiments, we found that the one-hop subgraph already provides significant offline improvements while maintaining high efficiency. Therefore, we adopt the one-hop configuration as the default setting in all experiments.
- **GNN Architecture.** For the GNN component, we primarily adopt LIGHTGCN [12] (as introduced in Section 3.1), which has demonstrated strong effectiveness in large-scale recommendation systems. We also conduct an ablation study on different GNN backbones, as presented in Section 5.2. The number of GNN layers  $k$  is set to match the number of sampled hops (i.e.,  $k = 1$  for the one-hop subgraph described above).

The hyperparameter searching space is provided in Appendix C.

**Evaluation** We first convert the continuous *staytime* metric into a binary label  $y_i^{(0)}$  based on a predefined threshold  $\tau$ :

$$y_i^{(0)} = \begin{cases} 1, & \text{if } \text{staytime}_i > \tau, \\ 0, & \text{otherwise.} \end{cases} \quad (25)$$

To enhance the supervision signal, we perform a label refinement step that aggregates multiple user interaction indicators (each typically binary, 0 or 1) into a unified interaction score:

$$\text{interact}_i = \sum_k \text{pos\_action}_{i,k} - \text{neg\_action}_{i,k}. \quad (26)$$

If a user exhibits any positive interaction (e.g., like, comment, share), the interaction score increases, while negative feedback decreases it. The final label is then obtained by combining the staytime-based label with this interaction score and clipping the result into [0, 1]:

$$y_i = \text{clip}(y_i^{(0)} + \text{interact}_i, 0, 1). \quad (27)$$

We apply the binary cross-entropy (BCE) loss for the two major towers (base-reward and stay-time) and AUC as the primary evaluation metric. Instead of absolute AUC values, we report the *relative improvements* over the baseline LTR model, where the improvement is computed as the percentage gain with respect to the baseline performance. Since our offline training is conducted in a **streaming** manner, to better demonstrate the stability and generalization of our model, we report the AUC improvement across six consecutive days after the model has converged, ensuring the fair comparison. We evaluate two fusion strategies for the upper-level feature integration: (1) simple gating fusion (gate) and (2) attention-based fusion (attn) in the result Table 1 (see Section 4.3).

**Table 1: Relative AUC improvements (%) across different models and dates.**

| Model           | Average | Day1   | Day2   | Day3   | Day4   | Day5   | Day6   |
|-----------------|---------|--------|--------|--------|--------|--------|--------|
| E2E-GRec (gate) | +1.40%  | +1.37% | +1.33% | +1.40% | +1.43% | +1.45% | +1.44% |
| E2E-GRec (attn) | +1.65%  | +1.71% | +1.56% | +1.59% | +1.59% | +1.60% | +1.58% |

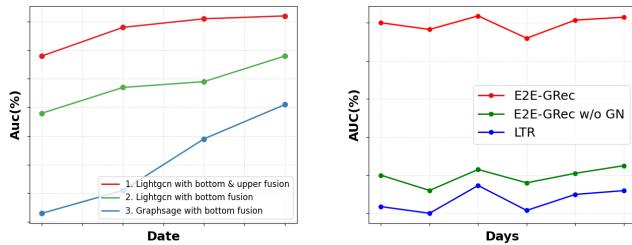
From the results, we can observe that our proposed E2E-GRec consistently outperforms the baseline LTR model. This improvement can be attributed to the following factors:

- (1) **Higher-order collaborative signals.** By introducing graph-based message passing, our model captures higher-order collaborative relationships among items, allowing the ranking tower to exploit neighborhood information beyond individual interactions.
- (2) **End-to-end optimization.** The end-to-end training framework creates a powerful synergy between the GNN and the downstream recommendation task. This structure allows the gradients from the LTR loss to flow all the way back through the network and directly influence the GNN's parameter updates. This ensures that the GNN also learns task-specific graph representations optimized for LTR task, leading to more discriminative and task-relevant embeddings.
- (3) **Self-supervised learning.** The SSL task guides the GNN to learn and generate more informative and higher-quality node

representations by focusing on the intrinsic structural properties of the graph. This approach mitigates potential task misalignment (see Theorem 2) and ultimately provides a more powerful set of embeddings for the downstream recommendation system.

## 5.2 Ablation Study

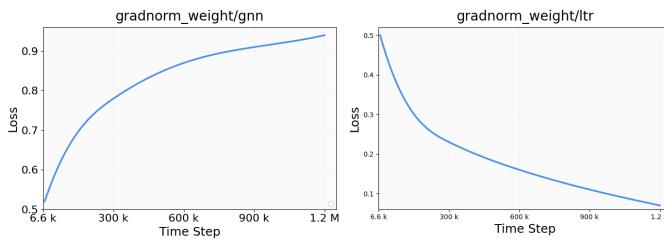
**5.2.1 GNN backbones.** We first provide an ablation study on different GNN backbones in Figure 2. From the results, we can see that LightGCN significantly outperforms GraphSAGE. This is because LightGCN removes unnecessary nonlinear transformations and feature mixing, allowing it to focus purely on high-order collaborative signals through linear neighborhood aggregation. Such a design not only reduces potential over-smoothing issue but also preserves the original embedding space structure, leading to more stable and better representations in recommendation tasks.



**Figure 2: Ablation on GNN backbones and fusion.**

**Figure 3: Ablation on the effect of Gradnorm.**

**5.2.2 Gradnorm & SSL.** In this section, we provide an ablation study on the influence of Gradnorm. We simply replace Gradnorm with a fixed coefficient in front of our SSL loss when performing the loss combination, while keeping all other settings the same. The results are shown in Figure 3. From the results, we can observe that: (1) Without Gradnorm, the GNN module can still enhance the basic recommendation system by incorporating neighbor information. (2) With Gradnorm, the performance is significantly improved. This is because it adjusts the gradient magnitudes of the SSL and LTR losses according to their learning speeds, ensuring that neither task dominates the optimization. We also provide Figure 4 to further illustrate this phenomenon, showing how Gradnorm dynamically balances the loss weights during training. As training progresses, Gradnorm adaptively adjusts the relative importance of each objective, ensuring stable optimization and preventing any single task from dominating the learning process.



**Figure 4: Dynamic Weight assigned by Gradnorm**

**5.2.3 Fusion Strategy.** As mentioned in Section 4.3, gating and attn serve as two different fusion strategies that trade off between performance and efficiency. To validate this, we provide an ablation study in Table 1. From the results, we can observe that attention-based fusion consistently outperforms gating-based fusion, which confirms our intuition that the attention mechanism can more effectively capture inter-relations between different feature representations from the LTR and GNN modules.

Furthermore, from Figure 2, we can observe that applying both upper and bottom fusion yields a much larger improvement compared to using bottom fusion alone. Compared with bottom-only fusion, the additional upper fusion path introduces an extra gradient contribution that is shorter and less attenuated, thereby generally providing more effective supervision and leading more stable convergence.

## 5.3 Online Results

To validate our method's real-world effectiveness, we deployed it in the online A/B test on the production large-scale recommendation system. We compared our model against the platform's highly optimized production baseline, which incorporates both rule-based strategies and state-of-the-art models. The evaluation focused on key user engagement and retention metrics: *StayDuration (SD)*: The average time a user spends on the platform per session, a primary indicator of overall engagement. *Last-7-Active Days (LT-7)*: The number of days a user was active in the last seven days, measuring short-term user retention and habit formation. Skip-related Metrics: *Skip/Play*: The ratio of skipped videos to total videos played. A lower value signifies higher content relevance. *Skip/User*: The average number of videos a user skips, indicating overall user satisfaction. *PlayTimeRate/Play (PTR)*: The ratio of a video's actual playtime to its total duration, measuring how engaging a specific piece of content is.

Our model demonstrated statistically significant improvements across all key metrics, as shown in Table 2. Notably, we observed a **+0.133%** relative increase in *StayDuration (SD)* and **+0.0262%** relative increase in *Last-7-Active Days (LT-7)*, showing users spent more time on the platform. At the same time, a **-0.1735%** reduction in *Skip/Play* and a **-0.3171%** reduction in *Skip/User* indicate that our model recommended more relevant content that users were less likely to skip. These results confirm that our method E2E-GRec delivers a positive impact on both immediate user engagement and longer-term retention. We attribute this success to the model's ability to leverage higher-order collaborative signals from the graph structure within an end-to-end training framework, leading to more effective and holistic optimization of the recommendation pipeline.

**Table 2: Relative Improvement (%) of Key Metrics in Online A/B Testing**

| Metric   | SD     | LT-7    | Skip/Play | Skip/User | PTR     |
|----------|--------|---------|-----------|-----------|---------|
| Lift (%) | +0.133 | +0.0262 | -0.1735   | -0.3171   | +0.1247 |

## 6 Conclusion

In this paper, we presented E2E-GRec, a novel end-to-end joint training framework that unifies GNNs with industrial recommender

systems. Built upon subgraphs sampled from a large-scale, cross-domain graph, our framework first introduces a Graph Feature Auto-Encoder as an auxiliary self-supervised task to enhance the quality of the learned GNN embeddings. The subsequent two-level feature fusion enables the effective injection of graph information into the recommendation model. Furthermore, the Gradnorm-based dynamic loss balancing ensures stable convergence and prevents task dominance during end-to-end training. Extensive offline evaluations, online A/B tests, and ablation studies on practical production data, together with theoretical analysis, demonstrate the superior effectiveness of our proposed approach.

## References

- [1] Charu C Aggarwal. 2016. An introduction to recommender systems. In *Recommender systems: The textbook*. Springer, 1–28.
- [2] Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. 2005. Learning to rank using gradient descent. In *Proceedings of the 22nd international conference on Machine learning*. 89–96.
- [3] Christopher JC Burges. 2010. From ranknet to lambdamart to lambdamart: An overview. *Learning* 11, 23–581 (2010). 81.
- [4] Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. 2007. Learning to rank: from pairwise approach to listwise approach. In *Proceedings of the 24th international conference on Machine learning*. 129–136.
- [5] Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. 785–794.
- [6] Zhao Chen, Vijay Badrinarayanan, Chen-Yu Lee, and Andrew Rabinovich. 2018. Gradnorm: Gradient normalization for adaptive loss balancing in deep multitask networks. In *International conference on machine learning*. PMLR, 794–803.
- [7] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*. 191–198.
- [8] Alex Fout, Jonathon Byrd, Basir Shariat, and Asa Ben-Hur. 2017. Protein interface prediction using graph convolutional networks. *Advances in neural information processing systems* 30 (2017).
- [9] Johannes Gasteiger, Aleksandar Bojchevski, and Stephan Günnemann. 2019. Combining Neural Networks with Personalized PageRank for Classification on Graphs. In *International Conference on Learning Representations*. <https://openreview.net/forum?id=H1gl-2A9Ym>
- [10] Mihajlo Grbovic and Haibin Cheng. 2018. Real-time personalization using embeddings for search ranking at airbnb. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 311–320.
- [11] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *Advances in neural information processing systems* 30 (2017).
- [12] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, Yongdong Zhang, and Meng Wang. 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.
- [13] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems* 30 (2017).
- [14] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).
- [15] Feng Li, Zhenrui Chen, Pengjie Wang, Yi Ren, Di Zhang, and Xiaoyu Zhu. 2019. Graph intention network for click-through rate prediction in sponsored search. In *Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval*. 961–964.
- [16] Kelong Mao, Jieming Zhu, Xi Xiao, Biao Lu, Zhaowei Wang, and Xiuqiang He. 2021. UltraGCN: ultra simplification of graph convolutional networks for recommendation. In *Proceedings of the 30th ACM international conference on information & knowledge management*. 1253–1262.
- [17] Francesco Ricci, Lior Rokach, and Bracha Shapira. 2010. Introduction to recommender systems handbook. In *Recommender systems handbook*. Springer, 1–35.
- [18] Aravind Sankar, Yozen Liu, Jun Yu, and Neil Shah. 2021. Graph neural networks for friend ranking in large-scale social platforms. In *Proceedings of the Web Conference 2021*. 2535–2546.
- [19] Xianfeng Tang, Yozen Liu, Neil Shah, Xiaolin Shi, Prasenjit Mitra, and Suhang Wang. 2020. Knowing your fate: Friendship, action and temporal explanations for user engagement prediction on social apps. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*. 2269–2279.
- [20] Michael Taylor, John Guiver, Stephen Robertson, and Tom Minka. 2008. Sofrank: optimizing non-smooth rank metrics. In *Proceedings of the 2008 International Conference on Web Search and Data Mining*. 77–86.
- [21] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).
- [22] Jizhe Wang, Pipei Huang, Huan Zhao, Zhibo Zhang, Binjiang Zhao, and Dik Lun Lee. 2018. Billion-scale commodity embedding for e-commerce recommendation in alibaba. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 839–848.
- [23] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. 2019. Neural graph collaborative filtering. In *Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval*. 165–174.
- [24] Felix Wu, Amauri Souza, Tianyi Zhang, Christopher Fifty, Tao Yu, and Kilian Weinberger. 2019. Simplifying graph convolutional networks. In *International conference on machine learning*. PMLR, 6861–6871.
- [25] Shiwén Wu, Fei Sun, Wentao Zhang, Xu Xie, and Bin Cui. 2022. Graph neural networks in recommender systems: a survey. *Comput. Surveys* 55, 5 (2022), 1–37.
- [26] Long Xia, Jun Xu, Yanyan Lan, Jiafeng Guo, Wei Zeng, and Xueqi Cheng. 2017. Adapting Markov decision process for search result diversification. In *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval*. 535–544.
- [27] Rui Xue. 2025. VISAGNN: Versatile Staleness-Aware Efficient Training on Large-Scale Graphs. *arXiv preprint arXiv:2111.12434* (2025).
- [28] Rui Xue, Haoyu Han, MohamadAli Torkamani, Jian Pei, and Xiaorui Liu. 2023. Lazygnn: Large-scale graph neural networks via lazy propagation. In *International Conference on Machine Learning*. PMLR, 38926–38937.
- [29] Rui Xue, Haoyu Han, Tong Zhao, Neil Shah, Jiliang Tang, and Xiaorui Liu. 2023. Large-Scale Graph Neural Networks: The Past and New Frontiers. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 5835–5836.
- [30] Rui Xue, Xipeng Shen, Ruozhou Yu, and Xiaorui Liu. 2023. Efficient large language models fine-tuning on graphs. *arXiv preprint arXiv:2312.04737* (2023).
- [31] Rui Xue and Tianfu Wu. 2025. H<sup>3</sup> GNNs: Harmonizing Heterophily and Homophily in GNNs via Joint Structural Node Encoding and Self-Supervised Learning. *arXiv preprint arXiv:2504.11699* (2025).
- [32] Rui Xue, Tong Zhao, Neil Shah, and Xiaorui Liu. 2024. Haste Makes Waste: A Simple Approach for Scaling Graph Neural Networks. *arXiv preprint arXiv:2410.05416* (2024).
- [33] Xiaoyong Yang, Yadong Zhu, Yi Zhang, Xiaobo Wang, and Quan Yuan. 2020. Large scale product graph construction for recommendation in e-commerce. *arXiv preprint arXiv:2010.05525* (2020).
- [34] Xinyang Ji, Ji Yang, Lichan Hong, Derek Zhiyuan Cheng, Lukasz Heldt, Aditee Kumthekar, Zhi Zhao, Li Wei, and Ed Chi. 2019. Sampling-bias-corrected neural modeling for large corpus item recommendations. In *Proceedings of the 13th ACM conference on recommender systems*. 269–277.
- [35] Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L Hamilton, and Jure Leskovec. 2018. Graph convolutional neural networks for web-scale recommender systems. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 974–983.
- [36] Jiahao Zhang, Rui Xue, Wenqi Fan, Xin Xu, Qing Li, Jian Pei, and Xiaorui Liu. 2024. Linear-Time Graph Neural Networks for Scalable Recommendations. In *Proceedings of the ACM on Web Conference 2024*. 3533–3544.
- [37] Shuai Zhang, Lina Yao, Aixin Sun, and Yi Tay. 2019. Deep learning based recommender system: A survey and new perspectives. *Comput. Surveys* 52, 1 (2019), 1–38.
- [38] Changguang Zheng, Hongzhi Chen, Yuxuan Cheng, Zhezheng Song, Yifan Wu, Changji Li, James Cheng, Hao Yang, and Shuai Zhang. 2022. ByteGNN: efficient graph neural network training at large scale. *Proceedings of the VLDB Endowment* 15, 6 (2022), 1228–1242.
- [39] Da Zheng, Chao Ma, Minjie Wang, Jinjing Zhou, Qidong Su, Xiang Song, Quan Gan, Zheng Zhang, and George Karypis. 2020. DistDGL: Distributed graph neural network training for billion-scale graphs. In *2020 IEEE/ACM 10th Workshop on Irregular Applications: Architectures and Algorithms (IA3)*. IEEE, 36–44.
- [40] Guorui Zhou, Xiaoqiang Zhu, Chenru Song, Ying Fan, Han Zhu, Xiao Ma, Yanghui Yan, Junqi Jin, Han Li, and Kun Gai. 2018. Deep interest network for click-through rate prediction. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 1059–1068.
- [41] Rong Zhu, Kun Zhao, Hongxia Yang, Wei Lin, Chang Zhou, Baole Ai, Yong Li, and Jingren Zhou. 2019. Aligraph: A comprehensive graph neural network platform. *arXiv preprint arXiv:1902.08730* (2019).
- [42] Shichao Zhu, Mufan Li, Guangmou Pan, and Xixun Lin. 2025. TTGL: large-scale multi-scenario universal graph learning at TikTok. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining* V. 2. 5249–5259.

## A Proof of Theorem 1

**Theorem 1** (Gradient Coupling in E2E-GRec). Let  $h_\theta(\cdot; \mathcal{G})$  denote the GNN embeddings with parameters  $\theta$ ,  $z_i = [h_\theta(x_i; \mathcal{G}) \| b_i]$ ,  $s_\psi$  be the recommendation scorer and

$$J(\theta, \psi, \alpha) = \frac{1}{n} \sum_{i=1}^n \left[ \alpha_1 L_{\text{rec}}(y_i, s_\psi(z_i)) + \alpha_2 L_{\text{gnn}}(\theta) \right] \quad (28)$$

Assume  $\frac{\partial s_\psi(z)}{\partial \theta} \neq 0$ . Then:

- (i)  $\nabla_\theta J$  contains a nonzero contribution from  $L_{\text{rec}}$  ( $\text{Rec} \rightarrow \text{GNN}$ );
- (ii)  $\frac{\partial}{\partial \theta} (\nabla_\psi J) \neq 0$ , i.e., the recommendation model gradient depends on  $\theta$  ( $\text{GNN} \rightarrow \text{Rec}$ ).

In contrast, cascaded pipelines satisfy  $\nabla_\theta J_{\text{rec}} = 0$  and  $\frac{\partial}{\partial \theta} (\nabla_\psi J) = 0$ , preventing capture of higher-order graph-rec interactions.

We show that, unlike cascaded pipelines that pretrain a GNN and then freeze its embeddings as augmented features for an recommendation model, our end-to-end (E2E) co-training causes the gradients of the GNN and the recommendation model to *mutually influence* each other. This coupling occurs (i) through feature fusion—where the recommendation model consumes the GNN embeddings via concatenation—and (ii) through loss fusion—where multiple objectives are combined via GradNorm.

*Setting.* Let  $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^n$  denote training instances (e.g., user-item pairs or query-document pairs) on a graph  $\mathcal{G}$ . A GNN with parameters  $\theta$  produces node/item/user embeddings  $h_\theta(\cdot; \mathcal{G})$ . The recommendation scorer with parameters  $\psi$  consumes a fused feature  $z_i := [h_\theta(x_i; \mathcal{G}) \| b_i]$ , where  $b_i$  are non-graph (auxiliary) features and  $\|$  denotes concatenation. The recommendation score is  $s_\psi(z_i)$ . We consider a ranking loss  $L_{\text{rec}}(y_i, s_\psi(z_i))$  (pointwise/pairwise/listwise) and an auxiliary graph loss  $L_{\text{gnn}}(\theta)$  (e.g., supervised or SSL).

### 1. Cascaded Baseline (No End-to-End Fine-tuning)

*Stage A (GNN only):*

$$\min_{\theta} J_{\text{GNN}}(\theta) = \frac{1}{n} \sum_{i=1}^n L_{\text{gnn}}(\theta). \quad (29)$$

*Stage B (Recommendation model on frozen embeddings):*

$$\min_{\psi} J_{\text{rec}}(\psi) = \frac{1}{n} \sum_{i=1}^n L_{\text{rec}}(y_i, s_\psi([h_{\theta_{\text{frozen}}}(x_i; \mathcal{G}) \| b_i])). \quad (30)$$

Gradients decouple:

$$\nabla_\theta J_{\text{rec}}(\psi) = \mathbf{0} \quad \text{and} \quad \nabla_\psi J_{\text{GNN}}(\theta) = \mathbf{0}, \quad (31)$$

so neither module's optimization can influence the other during training.

### 2. E2E-GRec: Joint Feature Fusion and Loss Fusion

We jointly optimize

$$\min_{\theta, \psi, \alpha_1, \alpha_2} J(\theta, \psi, \alpha) = \frac{1}{n} \sum_{i=1}^n \left[ \alpha_1 L_{\text{rec}}(y_i, s_\psi(z_i)) + \alpha_2 L_{\text{gnn}}(\theta) \right], \quad (32)$$

$$z_i = [h_\theta(x_i; \mathcal{G}) \| b_i], \quad (33)$$

where  $\alpha_1, \alpha_2 > 0$  are task weights learned by GradNorm to balance gradient magnitudes on a shared layer.

Gradients wrt GNN parameters  $\theta$ . By the chain rule,

$$\begin{aligned} \nabla_\theta J(\theta, \psi, \alpha) &= \frac{1}{n} \sum_{i=1}^n \alpha_1 \underbrace{\frac{\partial L_{\text{recommendation}}(y_i, s_\psi(z_i))}{\partial s_\psi(z_i)}}_{\substack{\neq 0 \text{ if recommendation error} \\ \partial z_i \\ \frac{\partial h_\theta(x_i; \mathcal{G})}{\partial \theta} \\ [I \ 0]}} \underbrace{\frac{\partial s_\psi(z_i)}{\partial z_i}}_{\text{sensitivity}} \\ &\quad + \alpha_2 \nabla_\theta L_{\text{gnn}}(\theta). \end{aligned} \quad (34)$$

Hence the *Rec* loss directly backpropagates into the GNN via the nonzero Jacobian  $\frac{\partial s_\psi(z_i)}{\partial z_i}$  and the fusion map  $z_i = [h_\theta(x_i; \mathcal{G}) \| b_i]$ . If the recommendation has any non-degenerate dependence on the  $h_\theta$  coordinates (e.g., the first recommendation layer has nonzero weights on the  $h_\theta$  block), then the first term in (34) is nonzero whenever the recommendation signal is nonzero, proving **recommendation  $\Rightarrow$  GNN** influence.

Gradients wrt recommendation parameters  $\psi$ .

$$\nabla_\psi J(\theta, \psi, \alpha) = \frac{1}{n} \sum_{i=1}^n \alpha_1 \frac{\partial L_{\text{rec}}(y_i, s_\psi(z_i))}{\partial s_\psi(z_i)} \frac{\partial s_\psi(z_i)}{\partial \psi}. \quad (35)$$

Although  $\nabla_\psi J$  does not include a direct  $\frac{\partial L_{\text{gnn}}}{\partial \psi}$  term, it depends on  $z_i$ , which contains  $h_\theta(x_i; \mathcal{G})$ . Thus the *numerical value and direction* of  $\nabla_\psi J$  are functions of  $\theta$ . So changing  $\theta$  changes  $z_i$  and hence *changes the recommendation gradient itself*. This establishes **GNN  $\Rightarrow$  recommendation** influence.

GradNorm adjusts  $(\alpha_1, \alpha_2)$  to balance gradient norms on a chosen shared layer  $W$ : letting  $G_k := \|\nabla_W(\alpha_k L_k)\|$  for  $k \in \{\text{rec}, \text{gnn}\}$ , the  $\alpha$ 's are updated to drive  $G_k$  toward a common target. Because  $G_{\text{rec}}$  depends on  $(\theta, \psi)$  via  $z$  and  $s_\psi$ , and  $G_{\text{gnn}}$  depends on  $\theta$ , the *loss weights themselves* become functions of both modules' states. This dynamically re-weights  $L_{\text{rec}}$  and  $L_{\text{gnn}}$  so that each task's gradient contribution adaptively influences the other.

## B Proof of Theorem 2

**Theorem 2** (SSL vs. Cascaded Ranking Head: Objective Misalignment). Let  $X \in \mathbb{R}^{n \times d}$  be node features and  $Z = f_\theta(X) \in \mathbb{R}^{n \times k}$  ( $k \geq d$ ) the encoder output. Assume a linear decoder  $W_d \in \mathbb{R}^{k \times d}$  with full column rank  $\text{rank}(W_d) = d$ . If the SSL reconstruction objective attains zero loss,

$$\|X - ZW_d\|_F^2 = 0, \quad (36)$$

then  $\text{col}(X) \subseteq \text{col}(Z)$  (i.e.,  $Z$  preserves the full feature subspace of  $X$ ). Consider the cascaded recommendation head with BPR scores  $s_i = w^\top z_i$  for some  $w \in \mathbb{R}^k$  and

$$\mathcal{L}_{\text{BPR}}(Z) = \sum_{(i,j) \in \mathcal{P}} \ell(w^\top (z_i - z_j)), \quad (37)$$

where  $\ell'(\cdot)$  exists. If  $\text{col}(X) \not\subseteq \text{span}(w)$ , then the two objectives are misaligned: there exists a nontrivial subspace

$$\mathcal{U} = \text{col}(X) \cap \ker(w^\top) \neq \{0\} \quad (38)$$

such that SSL requires  $Z$  to preserve information along  $\mathcal{U}$ , while BPR is invariant to any perturbations  $\{\delta z_i\}$  with  $\delta z_i \in \ker(w^\top) = \{v \in \mathbb{R}^k : w^\top v = 0\}$  for all  $i$  (hence supplies no gradient on  $\mathcal{U}$ ). Therefore BPR constrains only the one-dimensional subspace  $\text{span}(w)$  and leaves the  $(k-1)$ -dimensional orthogonal subspace  $\ker(w^\top)$  unconstrained, establishing the misalignment.

**PROOF.** (*SSL side*)  $\|X - ZW_d\|_F^2 = 0$  and  $\text{rank}(W_d) = d$  imply  $X = ZW_d$ , hence  $\text{col}(X) \subseteq \text{col}(ZW_d) \subseteq \text{col}(Z)$ .

(*BPR side*) For any pair  $(i, j)$ , by chain rule  $\frac{\partial \ell(w^\top(z_i - z_j))}{\partial z_i} = \ell'(w^\top(z_i - z_j)) w$  and  $\frac{\partial \ell(w^\top(z_i - z_j))}{\partial z_j} = -\ell'(w^\top(z_i - z_j)) w$ . Summing over pairs yields  $\frac{\partial \mathcal{L}_{\text{BPR}}}{\partial z_i} = \alpha_i w \in \text{span}(w)$  for some scalar  $\alpha_i$ . Thus for any  $v \in \ker(w^\top)$ ,  $\langle \frac{\partial \mathcal{L}_{\text{BPR}}}{\partial z_i}, v \rangle = 0$ , and  $\mathcal{L}_{\text{BPR}}$  is invariant to perturbations  $z_i \mapsto z_i + \delta z_i$  with  $\delta z_i \in \ker(w^\top)$ .

(*Misalignment*) If  $\text{col}(X) \not\subseteq \text{span}(w)$ , then  $\mathcal{U} = \text{col}(X) \cap \ker(w^\top)$  is nontrivial. SSL enforces  $Z$  to retain information along  $\mathcal{U}$  (as  $\mathcal{U} \subseteq \text{col}(X) \subseteq \text{col}(Z)$ ), whereas BPR provides no constraint or gradient signal on  $\mathcal{U}$  since gradients lie in  $\text{span}(w)$ . Hence the objectives are structurally misaligned. Hence:

- **SSL objective (linear reconstruction):**

$$\min_Z \|X - ZW_d\|^2 \Rightarrow \text{col}(X) \subseteq \text{col}(Z), \text{rank}(Z) \geq d. \quad (39)$$

- **BPR objective (cascaded ranking):**

$$\min_Z \mathcal{L}_{\text{BPR}}(Z) \Rightarrow \min_Z \mathcal{L}_{\text{BPR}}(Z) \text{ only the projection } w^\top Z \text{ is constrained; } \quad (40)$$

$$\ker(w^\top) \text{ is unconstrained/invariant.} \quad (41)$$

□

## C Hyperparameter Search Space

The hyperparameter search space is defined as follows:

- Learning rate for GNN and towers: {0.01, 0.005, 0.001}.
- Dropout rate for GNN: {0.1, 0.3, 0.5, 0.7, 0.8}.
- Dropout rate for attention: {0.1, 0.3, 0.5, 0.7, 0.8}.
- Weight decay for GNN: {0,  $1 \times 10^{-3}$ ,  $5 \times 10^{-3}$ ,  $8 \times 10^{-3}$ ,  $1 \times 10^{-4}$ ,  $5 \times 10^{-4}$ ,  $8 \times 10^{-4}$ }.