



 Latest updates: <https://dl.acm.org/doi/10.1145/3701716.3715498>

SHORT-PAPER

## Graph-Sequential Alignment and Uniformity: Toward Enhanced Recommendation Systems

**YUWEI CAO**, University of Illinois at Chicago, Chicago, IL, United States

**LIANGWEI YANG**, Salesforce.com, Inc., San Francisco, CA, United States

**ZHIWEI LIU**, Salesforce.com, Inc., San Francisco, CA, United States

**YUQING LIU**, University of Illinois at Chicago, Chicago, IL, United States

**CHEN WANG**, University of Illinois at Chicago, Chicago, IL, United States

**YUEQING LIANG**, Illinois Institute of Technology, Chicago, IL, United States

[View all](#)

**Open Access Support** provided by:

**University of Illinois at Chicago**

**Salesforce.com, Inc.**

**Illinois Institute of Technology**

**Beihang University**



PDF Download  
3701716.3715498.pdf  
25 February 2026  
Total Citations: 3  
Total Downloads: 1478

Published: 23 May 2025

[Citation in BibTeX format](#)

WWW '25: The ACM Web Conference

2025

April 28 - May 2, 2025

Sydney NSW, Australia

Conference Sponsors:  
**SIGWEB**

# Graph-Sequential Alignment and Uniformity: Toward Enhanced Recommendation Systems

Yuwei Cao

University of Illinois Chicago  
Chicago, IL, USA  
ycao43@uic.edu

Yuqing Liu

University of Illinois Chicago  
Chicago, IL, USA  
yliu363@uic.edu

Hao Peng\*

Beihang University  
Beijing, China  
Hangzhou Innovation Institute of  
BUAA  
Hangzhou, Zhejiang, China  
penghao@buaa.edu.cn

Liangwei Yang

Salesforce AI Research  
Palo Alto, CA, USA  
liangwei.yang@salesforce.com

Chen Wang

University of Illinois Chicago  
Chicago, IL, USA  
cwang266@uic.edu

Zhiwei Liu

Salesforce AI Research  
Palo Alto, CA, USA  
zhiweiliu@salesforce.com

Yueqing Liang

Illinois Institute of Technology  
Chicago, IL, USA  
yliang40@hawk.iit.edu

Philip S. Yu

University of Illinois Chicago  
Chicago, IL, USA  
psyu@uic.edu

## Abstract

Graph-based and sequential methods are two popular recommendation paradigms, each excelling in its domain but lacking the ability to leverage signals from the other. To address this, we propose a novel method that integrates both approaches for enhanced performance. Our framework uses Graph Neural Network (GNN)-based and sequential recommenders as separate submodules while sharing a unified embedding space optimized jointly. To enable positive knowledge transfer, we design a loss function that enforces alignment and uniformity both within and across submodules. Experiments on three real-world datasets demonstrate that the proposed method significantly outperforms using either approach alone and achieves state-of-the-art results. Our implementations are publicly available at <https://github.com/YuweiCao-UIC/GSAU.git>.

## CCS Concepts

- Information systems → Retrieval models and ranking.

## Keywords

Recommender Systems; Sequential Learning; Graph Neural Network; Alignment and Uniformity

\*Corresponding author

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](mailto:permissions@acm.org).

WWW Companion '25, April 28-May 2, 2025, Sydney, NSW, Australia  
© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 979-8-4007-1331-6/25/04  
<https://doi.org/10.1145/3701716.3715498>

## ACM Reference Format:

Yuwei Cao, Liangwei Yang, Zhiwei Liu, Yuqing Liu, Chen Wang, Yueqing Liang, Hao Peng, and Philip S. Yu. 2025. Graph-Sequential Alignment and Uniformity: Toward Enhanced Recommendation Systems. In *Companion Proceedings of the ACM Web Conference 2025 (WWW Companion '25), April 28-May 2, 2025, Sydney, NSW, Australia*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3701716.3715498>

## 1 Introduction

In recent years, deep learning-based approaches have significantly advanced recommender systems by enabling expressive representations of users and items. Among these, Graph Neural Network (GNN) [6, 10]-based recommenders [7, 18] and sequential recommendation systems [9, 13] have emerged as two widely adopted paradigms, each demonstrating state-of-the-art performance. Despite their effectiveness, both methods exhibit inherent limitations: GNN-based recommenders excel at capturing higher-order collaborative filtering signals through convolutions on bipartite user-item graphs but fail to incorporate sequential information. On the other hand, sequential recommenders adeptly model personalized interaction histories but struggle to account for higher-order correlations between users.

Recognizing the respective strengths of GNN-based and sequential recommenders in modeling higher-order and sequential data, we propose integrating both approaches for enhanced performance. Our method is straightforward, utilizing each architecture as separate submodules while maintaining a shared embedding space optimized by both. The main challenge is ensuring effective knowledge transfer between the two submodules to achieve synergy. Specifically, GNN-based recommenders often utilize loss functions like Bayesian personalized ranking (BPR) [12] for pairwise ranking or InfoNCE [5] for contrastive representation learning, while sequential recommenders typically employ cross-entropy (CE) loss [13] or BPR loss for next-item prediction tasks. Studies [3] have shown

that directly applying these original loss functions during the co-training of different architectures, even with normalization terms, can lead to unstable training processes and erratic gradients due to conflicts in optimization dynamics. To address this issue, we design a novel loss function that enforces alignment and uniformity properties [17] both within and across the submodules. The proposed loss enables the GNN-based and sequential encoders to supplement and benefit one another, achieving significant performance improvements. Our contributions are as follows:

- We propose a simple yet effective method that co-trains GNN-based and sequential recommenders to enhance performance.
- Our custom loss function enforces alignment and uniformity, facilitating effective knowledge transfer between the GNN-based and sequential encoders.
- We demonstrate that integrating GNN-based and sequential encoders delivers significantly better performance compared to using either approach independently. Experimental results on three real-world datasets reveal that our proposed method consistently outperforms state-of-the-art graph-based and sequential recommenders by substantial margins.

## 2 Related Work

**Integrating GNN-Based and Sequential Recommenders** GNN-based [7, 18] and sequential recommenders [9, 13] excel at modeling bipartite user-item graphs and user-specific interaction sequences, respectively, but often overlook the complementary value of other data. Few works address this limitation [1]: MCLSR [19] enriches sequential recommenders by graph construction, while MRGSRec [1] combines graph and sequential embeddings via a fusion layer. In contrast, our method employs a simpler design, enabling GNN-based and sequential encoders to share embedding layers without requiring fusion modules. To ensure effective knowledge transfer, we introduce a loss function that enforces alignment and uniformity [17] within and across encoders, allowing them to augment and enhance each other's performance.

**Alignment and Uniformity in Recommendation** Recent advances in unsupervised contrastive representation learning [17] highlight two crucial properties for high-quality representations: alignment, ensuring positive pairs are close, and uniformity, distributing instances evenly across a hypersphere. Contrastive losses that optimize these properties have been applied in collaborative filtering (CF) [11, 16], GNN-based [15, 21, 22], and sequential recommendation systems [4], yielding improved performance over traditional losses like BPR [12], InfoNCE [5], and CE [13]. Unlike prior studies, this work investigates the use of these properties for the joint optimization of GNN-based and sequential encoders.

## 3 Methodology

We introduce the architecture of our proposed model and detail its loss function that enforces alignment and uniformity. The method is named as graph-sequential alignment and uniformity (GSAU). An overview of GSAU is provided in Figure 1.

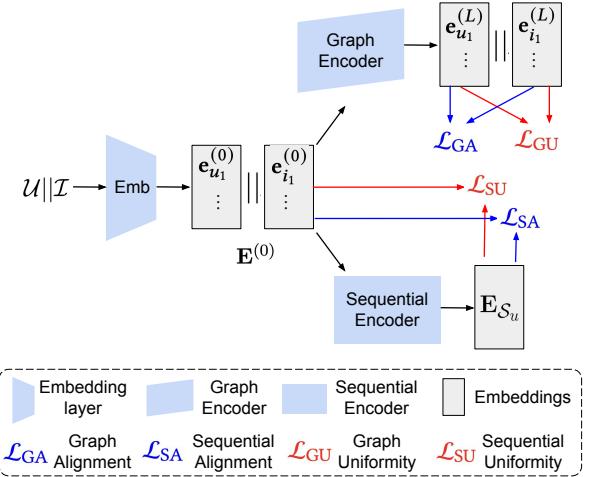


Figure 1: Overview of the proposed GSAU model.

### 3.1 Model Architecture

We adopt a simple approach that retains the original architectures of the GNN-based and sequential encoders, leveraging a shared embedding layer to enable knowledge sharing between them.

**3.1.1 Shared Embedding Layer.** Let  $\mathcal{U} = \{u_n\}_{n=1}^N$  and  $\mathcal{I} = \{i_m\}_{m=1}^M$  denote the user and item set, respectively. An embedding layer, denoted as  $\text{Emb}(\cdot)$ , is a network that maps each user and item into a low-dimensional representation  $\text{Emb}(u), \text{Emb}(i) \in \mathbb{R}^d$ , where  $d$  is the dimension of the latent space. The embedding layer maps all user and items to their representations, denoted as  $E^{(0)} = (e_u^{(0)}, \dots, e_i^{(0)}, \dots) = (\text{Emb}(u_1), \dots, \text{Emb}(i_1), \dots) \in \mathbb{R}^{(N+M) \times d}$ .

**3.1.2 GNN Encoder.** The core principle of a GNN encoder is to learn node representations through propagating features across the graph [7]. This is achieved through iteratively aggregating information from neighboring nodes, as formulated by:

$$\mathbf{e}_u^{(l+1)} = \text{GraphAGG}^{(l+1)}(\mathbf{e}_u^{(l)}, \{\mathbf{e}_i^{(l)} \mid (u, i) \in \mathcal{R}\}), \quad (1)$$

$$\mathbf{e}_i^{(l+1)} = \text{GraphAGG}^{(l+1)}(\mathbf{e}_i^{(l)}, \{\mathbf{e}_u^{(l)} \mid (u, i) \in \mathcal{R}\}), \quad (2)$$

where  $\mathcal{R} = \{(u, i) \mid u \text{ interacted with } i\}$  is a set of observed user-item interactions.  $l = 1, \dots, L$  is the layer number.  $\mathbf{e}_u^{(l)}$  and  $\mathbf{e}_i^{(l)}$  are the embeddings of user  $u$  and item  $i$  at the  $l$ -th layer.  $\text{GraphAGG}(\cdot)$ , the graph aggregator, is a summarizing function that aggregates the representations of the target node and its neighbors. Notable examples of aggregators include the weighted sum aggregators used in GIN [20] and LightGCN [7], the LSTM aggregator from GraphSAGE [6], and the attention-based aggregator from GAT [14].

**3.1.3 Sequential Encoder.** A sequential recommender identifies patterns in users' historical interactions. Let  $\mathcal{S}_u = [i_1^u, \dots, i_t^u, \dots, i_{|\mathcal{S}_u|}^u]$  denote the interacted items in chronological order for  $u \in \mathcal{U}$ , where  $i_t^u \in \mathcal{I}$  is the item that  $u$  has interacted with at time step  $t$ . Let  $\mathbf{E}_{\mathcal{S}_u}^{(0)} \in \mathbb{R}^{|\mathcal{S}_u| \times d}$  denote the initial embedding of  $\mathcal{S}_u$  given by the embedding layer. The embedding process of a sequential encoder

Method	Beauty				Sports				Toys			
	R@10	R@50	N@10	N@50	R@10	R@50	N@10	N@50	R@10	R@50	N@10	N@50
LightGCN	0.0435	0.1157	0.0216	0.0372	0.0299	0.0818	0.0153	0.0265	0.0425	0.1016	0.0226	0.0355
DirectAU	0.0492	0.1270	0.0252	0.0421	0.0318	0.0892	0.0165	0.0289	0.0501	0.1160	0.0264	0.0406
GraphAU	0.0489	0.1300	0.0245	0.0421	0.0323	0.0929	0.0168	0.0299	0.0472	0.1154	0.0248	0.0396
BERT4Rec	0.0390	0.1072	0.0189	0.0335	0.0174	0.0553	0.0086	0.0167	0.0299	0.0813	0.0157	0.0267
BERT4Rec (u)	0.0487	0.1207	0.0258	0.0414	0.0195	0.0604	0.0095	0.0182	0.0471	0.1090	0.0255	0.0387
SASRec	0.0826	0.1752	0.0414	0.0616	0.0467	0.1115	0.0218	0.0359	0.0891	0.1767	0.0434	0.0625
SASRec (u)	<u>0.0833</u>	<u>0.1806</u>	<u>0.0443</u>	<u>0.0653</u>	0.0470	0.1137	0.0239	0.0383	0.0878	<u>0.1799</u>	<u>0.0453</u>	<u>0.0653</u>
MRGSRec <sup>1</sup>	0.0681	/	0.0349	/	0.0434	/	0.0227	/	/	/	/	/
GSAU	0.0631	0.1622	0.0303	0.0518	<u>0.0491</u>	<u>0.1188</u>	<u>0.0256</u>	<u>0.0406</u>	0.0655	0.1653	0.0306	0.0524
GSAU (rec)	<b>0.0876</b>	<b>0.1974</b>	<b>0.0453</b>	<b>0.0694</b>	<b>0.0533</b>	<b>0.1285</b>	<b>0.0272</b>	<b>0.0434</b>	<b>0.0937</b>	<b>0.1966</b>	<b>0.0474</b>	<b>0.0698</b>
$\Delta(\%)$	5.16	9.30	2.26	6.28	13.40	13.02	13.81	13.32	5.16	9.28	4.64	6.89

Table 1: Overall performance. The best (second-best) results are in bold (underlined). <sup>1</sup> marks results from [1].

Method	R@10	R@50	N@10	N@50
GSAU (rec)	<b>0.0876</b>	<b>0.1974</b>	<b>0.0453</b>	<b>0.0694</b>
w/o graph	0.0833	0.1806	0.0443	0.0653
w/o sequential	0.0492	0.1270	0.0252	0.0421
w/o rec	0.0631	0.1622	0.0303	0.0518
w/o u-i uniform	0.0739	0.1743	0.0391	0.0609
replace sequential	0.0526	0.1397	0.0256	0.0444

Table 2: Ablation study on Beauty.

can be outlined as follows:

$$\mathbf{E}_{\mathcal{S}_u} = \text{SequentialAGG}(\text{Mask}(\mathbf{E}_{\mathcal{S}_u}^{(0)}) + \mathbf{P}_{\mathcal{S}_u}), \quad (3)$$

where  $\mathbf{P}_{\mathcal{S}_u} \in \mathbb{R}^{|\mathcal{S}_u| \times d}$  is the positional embedding [13] of  $\mathcal{S}_u$  that enables the usage of the sequential information.  $\text{Mask}(\cdot)$  constructs the cloze task [13] in which a proportion of  $\mathcal{S}_u$ , denoted as  $\mathcal{S}_u^m$ , are masked out and are to be inferred from the remaining part of  $\mathcal{S}_u$ , i.e.,  $\mathcal{S}_u^{\setminus m}$ . In the case of single-direction attention-based recommenders [9], the masked subset typically includes the right-most item.  $\text{SequentialAGG}(\cdot)$  denotes the sequential aggregator, examples include the GRU layers used in GRU4Rec [8], as well as the Transformer layers used in SASRec [9] and BERT4Rec [13].

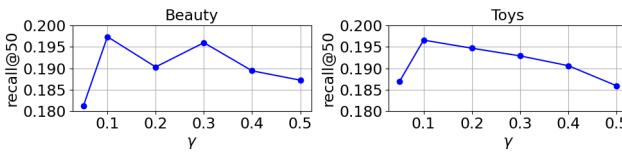


Figure 2: Parameter sensitivity to  $\gamma$ .

### 3.2 Alignment and Uniformity Loss

Effective positive knowledge transfer between the GNN and sequential encoders is crucial for obtaining high-quality representations. However, using heterogeneous loss functions (e.g., BPR loss for

the GNN encoder and CE loss for the sequential encoder) has been shown [3] to cause instability during model training and degrade representation quality. To mitigate this, we propose loss terms that enforce alignment and uniformity [17] between the encoders. Given the distribution of the positive pairs  $p_{\text{pos}}(\cdot, \cdot)$ , the alignment loss follows:

$$\mathcal{L}_A \triangleq \mathbb{E}_{(\mathbf{e}_x, \mathbf{e}_{x^+}) \sim p_{\text{pos}}} \|f(\mathbf{e}_x) - f(\mathbf{e}_{x^+})\|^2, \quad (4)$$

where  $f(\cdot)$  denotes the  $l_2$  norm of the input embedding. On the other hand, given the distribution of all instances  $p_{\text{data}}$ , the uniformity loss is formalized as:

$$\mathcal{L}_U \triangleq \log \mathbb{E}_{(\mathbf{e}_x, \mathbf{e}_{x'}) \stackrel{\text{i.i.d.}}{\sim} p_{\text{data}}} \exp(-2\|f(\mathbf{e}_x) - f(\mathbf{e}_{x'})\|^2). \quad (5)$$

As suggested by their mathematical formulations, the alignment and uniformity losses aim to achieve two objectives: bringing positive instances closer together and dispersing random instances uniformly across a hypersphere.

For the GNN encoder, both items and users are treated as instances. Specifically, the alignment objective encourages users to align closely with the items they have interacted with, while the uniformity objective ensures that the representations of all item and user instances are evenly distributed. Accordingly, we adopt graph alignment loss  $\mathcal{L}_{GA}$  that samples positive pairs from:

$$P_{g\_pos} = \{(\mathbf{e}_u^{(L)}, \mathbf{e}_i^{(L)}) \mid (u, i) \in \mathcal{R}\}, \quad (6)$$

while the random instances for calculating the graph uniformity loss,  $\mathcal{L}_{GU}$ , are sampled from:

$$P_{g\_data} = \{\mathbf{e}_i^{(L)}, \mathbf{e}_{i'}^{(L)} \mid i, i' \in \mathcal{I}\} \cup \{\mathbf{e}_u^{(L)}, \mathbf{e}_{u'}^{(L)} \mid u, u' \in \mathcal{U}\}. \quad (7)$$

For the sequential encoder, we treat items and users' historical interaction sequences as instances. We adopt sequential alignment loss  $\mathcal{L}_{SA}$  that samples positive pairs from:

$$P_{s\_pos} = \{(\mathbf{e}_i^{(0)}, \mathbf{E}_{\mathcal{S}_u}) \mid i \in \mathcal{S}_u^m\}, \quad (8)$$

while the random instances for calculating the sequential uniformity loss,  $\mathcal{L}_{\text{SU}}$ , are sampled from:

$$\begin{aligned} P_{\text{s\_data}} = & \{\mathbf{e}_i^{(0)}, \mathbf{e}_{i'}^{(0)} \mid i, i' \in \mathcal{I}\} \\ & \cup \{\mathbf{E}_{\mathcal{S}_u}, \mathbf{E}_{\mathcal{S}_{u'}} \mid u, u' \in \mathcal{U}\} \\ & \cup \{\mathbf{e}_i^{(0)}, \mathbf{E}_{\mathcal{S}_u} \mid i \notin \mathcal{S}_u\}. \end{aligned} \quad (9)$$

Note that unlike items and users, which have their own dedicated embedding layers (Section 3.1.1), the embeddings for users' historical interaction sequences are derived from item embeddings. Consequently, these interaction sequences act as 'hyper-items', mapped within the item embedding space. To further unify this space, we introduce a third subset in Equation (9). This subset encourages sequences to diverge from negative items, complementing the first two subsets that promote dispersion of items and sequences independently. The overall loss function is as follows:

$$\mathcal{L} = \mathcal{L}_{\text{GA}} + \mathcal{L}_{\text{SA}} + \gamma(\mathcal{L}_{\text{GU}} + \mathcal{L}_{\text{SU}}), \quad (10)$$

The weight  $\gamma$  determines the desired balance between alignment and uniformity. All sampling is conducted at the batch level.

## 4 Experiment

### 4.1 Experimental Setup

**Datasets and Evaluation Metrics.** We conduct experiments on three real-world datasets: Amazon Toys & Games, Beauty, and Sports & Outdoors. Following [1, 2, 4], we retain the 5-core data and adopt a leave-one-out evaluation strategy. Specifically, for each user's historical interaction sequence (with interactions sorted in ascending order by timestamp), the last interaction is used for testing, the second-to-last for validation, and the remaining interactions for training. We report top-k recall and Normalized Discounted Cumulative Gain (NDCG), where k is set to 10/50.

**Models and Implementation Details.** We evaluate **GSAU** and its variant, **GSAU (rec)**, which replaces the sequential alignment loss,  $\mathcal{L}_{\text{SA}}$ , with recommendation loss (i.e., CE). Comparisons are made with graph-based models like **LightGCN** [7], **DirectAU** [16] (applying alignment and uniformity loss on LightGCN), and **GraphAU** [21] (enhancing DirectAU with higher-order alignment). We also compare with sequential recommenders, including **BERT4Rec** [13], **SASRec** [9], and their uniformity-enhanced variants [4], **BERT4Rec (u)** and **SASRec (u)**. Additionally, we evaluate against **MRGSRec** [1], a graph-sequential method. For GSAU and GSAU (rec), we use LightGCN as the graph encoder, SASRec as the sequential encoder, and tune  $\gamma$  in {0.05, 0.1, 0.2, 0.3, 0.4, 0.5}. For a fair comparison, all methods are implemented using the RecBole framework [23]. We adopt the Adam optimizer, with a learning rate of 1e-3, embedding size of 64, and a maximum of 300 epochs, applying early stopping if validation NDCG@20 drops for 10 consecutive epochs.

### 4.2 Overall Performance

As shown in Table 1, GSAU (rec) outperforms the best baseline, SASRec (u), by notable margins across datasets and metrics ( $\Delta \%$ ). This demonstrates that integrating graph and sequential models delivers superior performance compared to using either model independently, with GSAU (rec) offering an effective approach to achieve this synergy. Comparing GSAU with GSAU (rec), replacing

the sequential alignment loss with CE loss improves performance. This is likely because CE loss not only aligns sequences with their positive items but also disperses them from all negatives, enhancing uniformity of the item embedding space alongside alignment.

### 4.3 Ablation and Hyperparameter Sensitivity Analysis

Table 2 shows the ablation study results. The degraded performance of *w/o graph* (removing the graph encoder) and *w/o sequential* (removing the sequential encoder) highlights the importance of both encoders. Similarly, the poorer results of *w/o rec* (using alignment loss instead of CE loss in the sequential encoder) and *w/o u-i uniform* (removing the third subset in Equation 9) demonstrate the importance of dispersing sequences from the negative items and fully unifying the item embedding space. The results of *replace sequential* (using BERT4Rec instead of SASRec as the sequential encoder) confirm the value of a strong sequential encoder and show that the proposed method remains effective regardless of the choice of the sequential encoder, outperforming BERT4Rec and BERT4Rec (u) in Table 1. Figure 2 shows that a  $\gamma$  value between 0.1 and 0.4 strikes a balance between alignment and uniformity losses, yielding good performance, with  $\gamma = 0.1$  achieving the best results. We also observe notable performance deterioration (not shown here) when  $\gamma$  is too large ( $> 0.5$ ) or too small ( $< 0.05$ ).

## 5 Conclusion

We propose a simple and effective method that combines GNN-based and sequential recommenders, leveraging a custom loss function to enforce alignment and uniformity for efficient knowledge transfer. Experiments show significant performance gains over using GNN-based or sequential recommenders alone.

## Acknowledgments

This work is supported in part by NSF under grants III-2106758, and POSE-2346158. Hao Peng is supported by the National Key R&D Program of China through grant 2022YFB3104703, the NSFC through grants 62322202 and 62441612, and the "Pioneer" and "Leading Goose" R&D Program of Zhejiang through grant 2025C02044.

## References

- [1] Vladimir Baikarov and Evgeny Frolov. 2024. End-to-End Graph-Sequential Representation Learning for Accurate Recommendations. In *Companion Proceedings of the ACM on Web Conference 2024*. 501–504.
- [2] Yuwei Cao, Nikhil Mehta, Xinyang Yi, Raghuandan Keshavan, Lukasz Heldt, Lichan Hong, Ed H Chi, and Maheswaran Sathiamoorthy. 2024. Aligning Large Language Models with Recommendation Knowledge. *arXiv preprint arXiv:2404.00245* (2024).
- [3] Yuwei Cao, Liangwei Yang, Chen Wang, Zhiwei Liu, Hao Peng, Chenyu You, and Philip S Yu. 2023. Multi-task Item-attribute Graph Pre-training for Strict Cold-start Item Recommendation. *arXiv preprint arXiv:2306.14462* (2023).
- [4] Hyunsoo Chung, AI Omnis, and Jungtaek Kim. 2023. Leveraging Uniformity of Normalized Embeddings for Sequential Recommendation. *History* 4, 5 (2023).
- [5] Michael Gutmann and Aapo Hyvärinen. 2010. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. In *Proc. AISTATS*.
- [6] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. In *Proceedings of the NeurIPS*. 1024–1034.
- [7] Xiangnan He, Kuan Deng, et al. 2020. Lightgc: Simplifying and powering graph convolution network for recommendation. In *Proc. SIGIR*.
- [8] B Hidasi. 2015. Session-based Recommendations with Recurrent Neural Networks. *arXiv preprint arXiv:1511.06939* (2015).
- [9] Wang-Cheng Kang and Julian McAuley. 2018. Self-attentive sequential recommendation. In *Proc. ICDM*.

- [10] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).
- [11] Seongmin Park, Mincheol Yoon, Jae-woong Lee, Hogun Park, and Jongwuk Lee. 2023. Toward a Better Understanding of Loss Functions for Collaborative Filtering. In *Proc. CIKM*. 2034–2043.
- [12] Steffen Rendle, Christoph Freudenthaler, et al. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proc. UAI*.
- [13] Fei Sun, Jun Liu, et al. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proc. CIKM*.
- [14] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2018. Graph attention networks. In *Proceedings of the ICLR*.
- [15] Chen Wang, Liangwei Yang, Zhiwei Liu, Xiaolong Liu, Mingdai Yang, Yueqing Liang, and Philip S. Yu. 2024. Collaborative Alignment for Recommendation (CIKM '24). Association for Computing Machinery, New York, NY, USA.
- [16] Chenyang Wang, Yuanqing Yu, et al. 2022. Towards Representation Alignment and Uniformity in Collaborative Filtering. In *Proc. SIGKDD*.
- [17] Tongzhou Wang and Phillip Isola. 2020. Understanding contrastive representation learning through alignment and uniformity on the hypersphere. In *Proc. ICML*.
- [18] 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.
- [19] Ziyang Wang, Huoyu Liu, Wei Wei, Yue Hu, Xian-Ling Mao, Shaojian He, Rui Fang, and Dangyang Chen. 2022. Multi-level contrastive learning framework for sequential recommendation. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*. 2098–2107.
- [20] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. 2018. How powerful are graph neural networks? *arXiv preprint arXiv:1810.00826* (2018).
- [21] Liangwei Yang, Zhiwei Liu, Chen Wang, Mingdai Yang, Xiaolong Liu, Jing Ma, and Philip S. Yu. 2023. Graph-based alignment and uniformity for recommendation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*. 4395–4399.
- [22] Mingdai Yang, Zhiwei Liu, Liangwei Yang, Xiaolong Liu, Chen Wang, Hao Peng, and Philip S. Yu. 2024. Unified Pretraining for Recommendation via Task Hypergraphs. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining, WSDM 2024, Merida, Mexico, March 4-8, 2024*, Luz Angelica Caudillo-Mata, Silvia Lattanzi, Andrés Muñoz Medina, Leman Akoglu, Aristides Gionis, and Sergei Vassilvitskii (Eds.). ACM, 891–900. <https://doi.org/10.1145/3616855.3635811>
- [23] Wayne Xin Zhao, Shanlei Mu, Yupeng Hou, Zihan Lin, Yushuo Chen, Xingyu Pan, Kaiyuan Li, Yujie Lu, Hui Wang, Changxin Tian, et al. 2021. Recbole: Towards a unified, comprehensive and efficient framework for recommendation algorithms. In *proceedings of the 30th acm international conference on information & knowledge management*. 4653–4664.