

Exploiting Inter-Session Information with Frequency-enhanced Dual-Path Networks for Sequential Recommendation

Peng He¹, Yanglei Gan¹, Tingting Dai¹, Run Lin¹, Xuexin Li², Yao Liu^{1*}, Qiao Liu^{1*}

¹University of Electronic Science and Technology of China, Chengdu, China

²Hebei Agricultural University, Baoding, China

{hepenglk, yangleigan, runlin, ttdai}@std.uestc.edu.cn, {liuyao, qliu}@uestc.edu.cn, lxuexin9@gmail.com

Abstract

Sequential recommendation (SR) aims to predict a user’s next item preference by modeling historical interaction sequences. Recent advances often integrate frequency-domain modules to compensate for self-attention’s low-pass nature by restoring the high-frequency signals critical for personalized recommendations. Nevertheless, existing frequency-aware solutions process each session in isolation and optimize exclusively with time-domain objectives. Consequently, they overlook cross-session spectral dependencies and fail to enforce alignment between predicted and actual spectral signatures, leaving valuable frequency information under-exploited. To this end, we propose **FreqRec**, a **F**requency-Enhanced Dual-Path Network for sequential **R**ecommendation that jointly captures inter-session and intra-session behaviors via a learnable Frequency-domain Multi-layer Perceptron. Moreover, FreqRec is optimized under a composite objective that combines cross entropy with a frequency-domain consistency loss, explicitly aligning predicted and true spectral signatures. Extensive experiments on three benchmarks show that FreqRec surpasses strong baselines and remains robust under data sparsity and noisy-log conditions.

Code — <https://github.com/AONE-NLP/FreqRec>

Introduction

Sequential recommendation (SR) (Fang et al. 2020; Wang et al. 2019) seeks to infer users’ future item preferences by modeling their historical interaction sequences. Such methods have become indispensable in a range of applications, including personalized content delivery (Liang, Lai, and Ku 2006) and social media feeds (Vombatkere et al. 2024). Over time, these recommenders have undergone significant advances from traditional approaches like Markov chain (Rendle, Freudenthaler, and Schmidt-Thieme 2010) to convolutional neural networks (CNNs) (Tang and Wang 2018; Niu et al. 2024) and recurrent neural networks (RNNs) (Donkers, Loepf, and Ziegler 2017; Liu et al. 2018), and more recently, self-attention mechanisms inspired by the Transformer family. These Transformer-based models (Kang and McAuley 2018; Sun et al. 2019; Lin et al. 2024) offer both improved training efficiency and stronger predictive accuracy.

*Corresponding Author.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

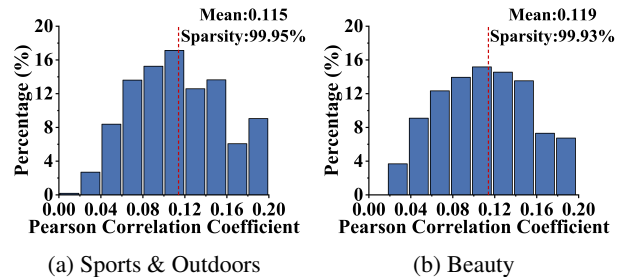


Figure 1: The distribution of Pearson correlation coefficient between sessions sharing at least one item on Sports & Outdoors and Beauty datasets, respectively.

Despite these advances, transformer-based approaches are particularly vulnerable to the noisy signals inherent in logged user data (Agichtein, Brill, and Dumais 2006), which can lead models to overfit and degrade generalization (Ying 2019; Liu et al. 2023a). Moreover, the global receptive field of self-attention acts similarly to a low-pass filter that attenuates high-frequency components (Du et al. 2023b). Consequently, abrupt shifts or periodic consumption in user behavior can be smoothed out (Wang et al. 2022; Dovonon, Bronstein, and Kusner 2024), erasing fine-grained dynamics that are crucial for personalized recommendations.

Recent work has explored two complementary strategies to address these limitations in the frequency domain. One line of work (Zhou et al. 2022; Du et al. 2023a; Liu et al. 2023a; Kim et al. 2025) applies explicit spectral filtering to suppress spurious signals and highlight essential sequence features. Another line (Du et al. 2023b; Shin et al. 2024) integrates the Fourier transform modules alongside self-attention to recover periodic patterns and complementary high-frequency information. Despite their state-of-the-art performance, frequency-based methods introduce new challenges while answering the aforementioned questions:

- **Neglect of Inter-Session Dependencies.** Contemporary frequency-domain models, such as FMLPRec (Du et al. 2023b), BSARec (Shin et al. 2024) treat each session in isolation, forfeiting the rich interaction patterns that span multiple sessions. Given that individual sessions are inherently short (Qiu et al. 2020), ignoring these external signals amplifies data sparsity and undermines predictive

accuracy. Figure 1 shows strong positive correlations between sessions that share at least one item on benchmark datasets, underscoring the importance of inter-session patterns. Graph-based solutions attempt to compensate by constructing inter-session (Guo et al. 2022; Qiao et al. 2023) or global session graphs (Wang et al. 2020, 2023), but their efficacy depends on meticulously hand-crafted edge rules that are expensive to tune and fragile under sparsity and noise (Liu et al. 2023b).

- **Frequency Features Underserved by Time-Domain Objectives.** Most sequential recommendation frameworks optimize losses defined purely in the time domain, such as (Binary) Cross-Entropy (Sun et al. 2019; Shin et al. 2024), Contrastive Loss (Chen et al. 2022), and Bayesian Personalized Ranking Loss (Rendle et al. 2009; Liu et al. 2025). Although effective for aligning predicted and observed target items, these objectives provide no direct incentive to capture frequency-domain structure. Consequently, periodic or high-frequency behavioral signals remain underutilized, limiting the model’s capacity to represent the full spectrum of user dynamics.

Considering the above issues, we present **FreqRec**, a **Frequency-Enhanced Dual-Path Network** for sequential **Recommendation** that jointly models cohort-level and user-specific behaviors. Specifically, FreqRec encodes the interaction history with self-attention to preserve long-range contextual dependencies, while channels the sequence through two complementary frequency paths: a Global Spectral Aggregator that distills cohort-level rhythms across sessions, and a Local Spectral Refiner that sharpens user-specific nuances. Both paths employ a unified Frequency-Domain Multi-layer Perceptrons (MLPs), whose complex-valued filters are learned end-to-end, eliminating the need for manual frequency cut-offs. The resulting representations are fused in either parallel or serial mode and reconciled with the contextual branch via a gated residual update. To ensure FreqRec learns meaningful spectral features, we augment the standard cross-entropy loss with a frequency-domain consistency loss that explicitly aligns the predicted and ground-truth spectral signatures. Our contribution are three-fold:

- We propose a **Frequency-enhanced dual-path network** for sequential **Recommendation (FreqRec)**, that jointly captures inter-session and intra-session behaviors by integrating a learnable, complex-valued Fourier Transform.
- We introduce a frequency-domain consistency loss that explicitly aligns the model’s predicted and ground-truth spectral coefficients, enforcing the recovery of high-frequency interaction patterns often smoothed out by standard self-attention mechanisms.
- Comprehensive experiments conducted on three real-world benchmarks, demonstrating that FreqRec outperforms state-of-the-art baselines, and remains robust under varying data sparsity and noisy-log scenarios.

Related Works

Sequential Recommendation

Sequential recommendation (SR) aims to learn from users’ historical interaction sequences to deliver personalized recommendations (Gan et al. 2025). Mainstream SR models predominantly adopt deep learning frameworks: Convolutional Neural Networks (CNNs) (Tang and Wang 2018; Niu et al. 2024) and Recurrent Neural Networks (RNNs) (Li et al. 2018; Liu et al. 2018) were initially widely employed. For instance, Caser (Tang and Wang 2018) pioneered the treatment of items in recommendation sequences, utilizing convolutional operations to capture item-item interactions. GRU4Rec (Hidasi et al. 2015) successfully implemented a GRU network to construct a SR framework for modeling long-term dependencies between users-items. Transformers (Vaswani 2017), owing to their exceptional long-term contextual modeling capabilities and parallel computation efficiency, have gained widespread adoption in SR. Models such as SASRec (Kang and McAuley 2018) and BERT4Rec (Sun et al. 2019) use self-attention to adaptively weight historical interactions, thereby more effectively capturing shifts in user intent. However, the global receptive field of self-attention acts as a low-pass filter, smoothing out abrupt changes in interests in user behavior (Du et al. 2023b).

To recover these crucial spectral cues, a growing body of work integrates frequency-domain processing within SR models. One line of research designs explicit spectral filters to attenuate noise and amplify informative frequency bands. For instance, Zhou et al. (Zhou et al. 2022) propose a learned band-pass filter that adapts its frequency response to sequence statistics, suppressing spurious interactions while retaining periodic patterns. SLIME4Rec (Du et al. 2023a) employs contrastive learning on spectral representations to improve robustness against noisy logs, while DIFF (Kim et al. 2025) leverages DFT-based filtering, combined with side information, to remove short-term fluctuations and enhance recommendation accuracy. A second strand embeds Fourier transform operations directly into the network architecture to recover high-frequency dynamics that self-attention may smooth over. FEARec (Du et al. 2023b) and BSARec (Shin et al. 2024) utilize DFT to capture high-frequency information, counteracting the inherent low-pass filtering properties of self-attention mechanisms. Although these methods achieve strong empirical gains, they process each session in isolation and optimize solely with time-domain losses. These gaps motivate our development of a dual-path architecture that simultaneously models inter-session and intra-session spectral dependencies while incorporating a frequency-domain consistency objective.

DFT for Sequential Modeling

The Discrete Fourier Transform (DFT) is a cornerstone of digital signal processing, decomposing time-domain signals into their constituent frequency components (Duhamel and Vetterli 1990; Frigo and Johnson 2005). Leveraging its inherent ability to efficiently capture global periodicities and complex patterns, DFT-based modules have recently been adopted to effectively enhance sequential representations.

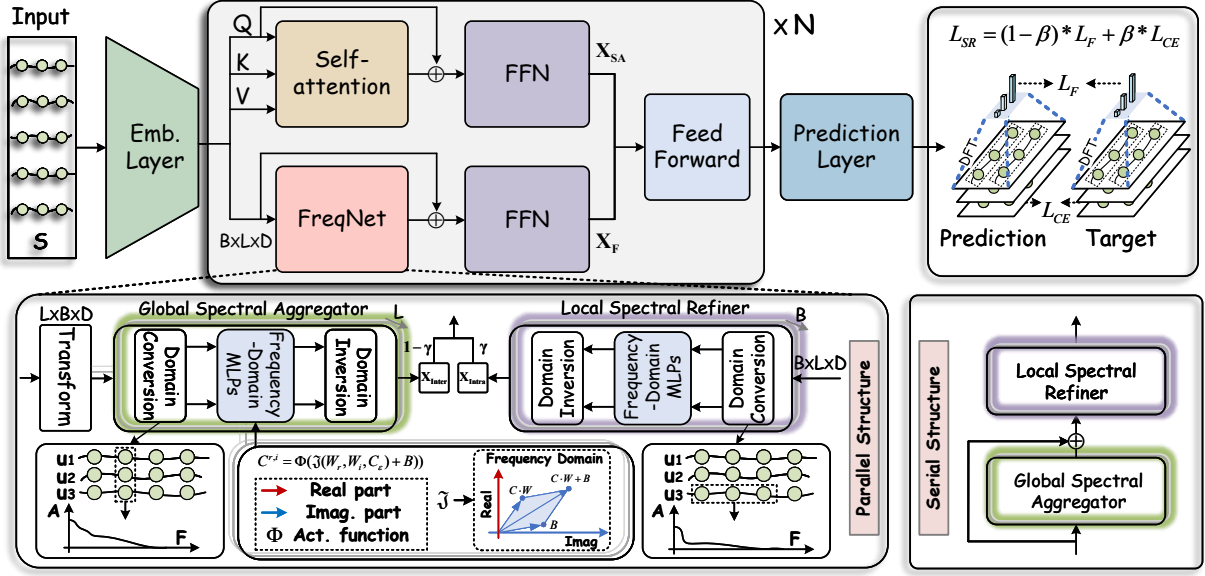


Figure 2: The overall framework of FreqRec. User sequences are embedded and then processed by two parallel paths: a self-attention branch and the *FreqNet* branch. *FreqNet* contains a *Global Spectral Aggregator* that performs batch-axis DFT \rightarrow FreqMLPs \rightarrow IDFT to distill cohort-level signals, while a *Local Spectral Refiner* applies the same pipeline along the temporal axis to refine user-specific cues. The two spectral modules can be fused in either parallel or serial form. Training minimizes a hybrid objective that couples standard cross-entropy prediction with *frequency-domain consistency loss*.

This trend is evident across diverse domains, including time-series forecasting (Yi et al. 2023; Yang et al. 2024), and even natural language processing (Tamkin, Jurafsky, and Goodman 2020; Lee-Thorp et al. 2022). Building on this trend, our work further extends DFT’s applicability by jointly modeling inter- and intra-session spectral dynamics.

Preliminaries

Sequential Recommendation. Let \mathcal{U} and \mathcal{I} denote the sets of users and items, respectively. For each user $u \in \mathcal{U}$, we observe a chronologically ordered interaction sequence $\mathcal{S}^u = [i_1, i_2, \dots, i_{|\mathcal{S}^u|}]$, where $i_j \in \mathcal{I}$ signifies the j -th item which u has interacted and $|\mathcal{S}^u|$ is the sequence length. The goal of SR task is to predict the next item i^* that user u is most probable to consume, given the historical sequence \mathcal{S}^u :

$$i_{|\mathcal{S}^u|+1}^u = \arg \max_{i^* \in \mathcal{I}} p(i^* | \mathcal{S}^u), \quad (1)$$

where $p(i^* | \mathcal{S}^u)$ stands for the conditional probability of user u will interact with item i^* next.

Discrete Fourier Transform (DFT). The Discrete Fourier Transform provides an exact, bijective mapping between a sequence in the time domain and its representation in the frequency domain, thereby enabling precise analysis and manipulation of periodic and noisy components. Let $\mathbf{T} = [T_0, T_1, \dots, T_{L-1}] \in \mathbb{R}^L$ denote a real-valued sequence of length L . The forward transform, denoted $\mathcal{F} : \mathbb{R}^L \rightarrow \mathbb{C}^L$, computes complex-valued coefficients C_k that encode the amplitude and phase information at each frequency k . Its inverse, $\mathcal{F}^{-1} : \mathbb{C}^L \rightarrow \mathbb{R}^L$, reconstructs the original sequence exactly from those coefficients.

- **Domain Conversion (\mathcal{F}).** The sequence T_n is mapped to its complex frequency coefficients C_k via the DFT:

$$C_k = \sum_{n=0}^{L-1} T_n \cdot e^{-j2\pi kn/L}, k = 0, 1, \dots, L-1 \quad (2)$$

where n and k are the discrete time and frequency indices ($0 \leq n, k < L$), and j is the imaginary unit. Using Euler’s formula, this is separated into its real and imaginary parts:

$$\begin{aligned} C_k &= \mathbf{Real}(C_k) + j\mathbf{Imag}(C_k), \\ \mathbf{Real}(C_k) &= \sum_{n=0}^{L-1} T_n \cos\left(\frac{2\pi kn}{L}\right), \\ \mathbf{Imag}(C_k) &= -\sum_{n=0}^{L-1} T_n \sin\left(\frac{2\pi kn}{L}\right). \end{aligned} \quad (3)$$

- **Domain Inversion (\mathcal{F}^{-1}).** We recover the original time-domain sequence T_n via the inverse discrete transform:

$$T_n = \frac{1}{L} \sum_{k=0}^{L-1} C_k \cdot e^{j2\pi kn/L} \quad (4)$$

For brevity, we refer to the forward and inverse mappings simply as \mathcal{F} and \mathcal{F}^{-1} .

Methodology

Historical Sequence Encoding

Suppose each user has a historical interaction sequence $\mathcal{S}^u = [i_1, i_2, \dots, i_{n-1}]$. We first map \mathcal{S}^u into item embedding $\mathcal{E}_u = \mathcal{M}[\mathcal{S}^u] \in \mathbb{R}^{L \times D}$ via $\mathcal{M} \in \mathbb{R}^{\mathcal{I} \times D}$. We then add

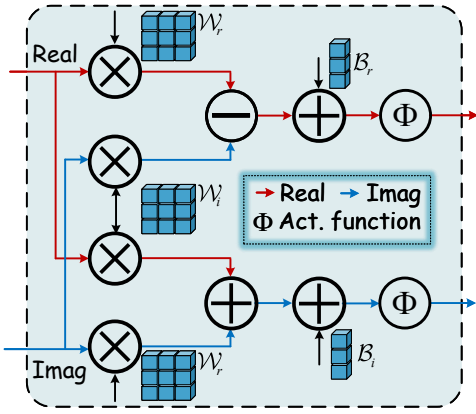


Figure 3: Demonstration of Frequency-Domain MLP.

a positional embedding $P \in \mathbb{R}^{L \times D}$ and apply layer normalization (LN) to obtain the sequence embedding:

$$\mathcal{E} = \text{Dropout}(\text{LN}(\mathcal{E}_u + P)). \quad (5)$$

Dual-Path Representation Learning

Contextual-enhanced Representation. We apply self-attention to \mathcal{E} to obtain its contextual representation:

$$\begin{aligned} \mathbf{X}_{\text{SA}} &= \text{Self-Attention}(\mathcal{E}), \\ \mathbf{X}_{\text{SA}} &= \text{FFN}(\mathbf{X}_{\text{SA}} + \mathcal{E}), \end{aligned} \quad (6)$$

where $\mathcal{E} \in \mathbb{R}^{B \times L \times D}$. FFN denotes Feed Forward Network.

Frequency-enhanced Network. As shown in Figure 2, given the sequence embedding \mathcal{E} , we first project it into the frequency domain using the DFT:

$$\mathcal{C}_{\mathcal{E}} = \mathcal{F}(\mathcal{E}) \rightarrow \mathcal{C}_{\mathcal{E}} = \text{Real}(\mathcal{C}_{\mathcal{E}}) + j\text{Imag}(\mathcal{C}_{\mathcal{E}}). \quad (7)$$

Instead of hand-crafting a low-pass or band-pass filter, we propose a **Frequency-Domain MLP (FreqMLP)** that learns which frequencies to amplify or suppress. As illustrated in Figure 3, FreqMLP jointly processes the real and imaginary channels, enabling cross-component interactions:

$$\begin{aligned} \mathcal{C}_{\mathcal{E}}^r &= \Phi(\mathcal{W}_r \cdot \text{Real}(\mathcal{C}_{\mathcal{E}}) - \mathcal{W}_i \cdot \text{Imag}(\mathcal{C}_{\mathcal{E}}) + \mathcal{B}_r), \\ \mathcal{C}_{\mathcal{E}}^i &= \Phi(\mathcal{W}_i \cdot \text{Real}(\mathcal{C}_{\mathcal{E}}) + \mathcal{W}_r \cdot \text{Imag}(\mathcal{C}_{\mathcal{E}}) + \mathcal{B}_i), \\ \mathcal{C}'_{\mathcal{E}} &= \mathcal{C}_{\mathcal{E}}^r + j\mathcal{C}_{\mathcal{E}}^i, \end{aligned} \quad (8)$$

where $\mathcal{W}_r, \mathcal{W}_i \in \mathbb{R}^{D \times D}$, $\mathcal{B}_r, \mathcal{B}_i \in \mathbb{R}^D$ and Φ is a non-linear activation. We recover a refined time-domain representation via the inverse DFT, denoted as $\mathbf{X}_{\text{Freq}} = \mathcal{F}^{-1}(\mathcal{C}'_{\mathcal{E}})$. For brevity, we denote the entire pipeline as:

$$\mathbf{X}_{\text{Freq}} = \mathcal{F}^{-1}(\text{FreqMLP}(\mathcal{F}(\mathcal{E}))). \quad (9)$$

Global Spectral Aggregator (GSA). User logs often share clear population-wide rhythms, that are invaluable when a single user's history is sparse. Hence, we treat a mini-batch of historical embedding $\mathcal{E}_{\text{Inter}} \in \mathbb{R}^{L \times B \times D}$ as a signal over the batch axis, perform a batch-axis DFT, pass the complex coefficients through a dedicated **FreqMLP** to suppress high-frequency noise and boost low-frequency cohort trends:

$$\mathbf{X}_{\text{Inter}} = \mathcal{F}^{-1}(\text{FreqMLP}(\mathcal{F}(\mathcal{E}_{\text{Inter}}))). \quad (10)$$

Local Spectral Refiner (LSR). Complementing the GSA stage, the LSR captures each user's fine-grained and often transient interests by applying the same frequency filtering along the temporal dimension L , given the sequence embedding of users $\mathcal{E}_{\text{Intra}} \in \mathbb{R}^{B \times L \times D}$:

$$\mathbf{X}_{\text{Intra}} = \mathcal{F}^{-1}(\text{FreqMLP}(\mathcal{F}(\mathcal{E}_{\text{Intra}}))), \quad (11)$$

where the input $\mathcal{E}_{\text{Intra}}$ depends on the fusion strategy:

$$\mathcal{E}_{\text{Intra}} = \begin{cases} \mathcal{E}, & \text{Parallel Fusion} \\ \mathcal{E} + \mathbf{X}_{\text{Inter}}. & \text{Serial Fusion} \end{cases} \quad (12)$$

Fusion and Output. Once we have both the cohort representation $\mathbf{X}_{\text{Inter}}$ and the personalized representation $\mathbf{X}_{\text{Intra}}$, we integrate them into a single feature map, which can proceed in the following way:

- Parallel Fusion** processes inter- and intra-session streams independently and combines their respective outputs with a weighted sum, which preserves the integrity of the raw sequence features and allows each module to specialize without mutual interference:

$$\mathbf{X}_{\text{F}} = \text{FFN}((1 - \gamma) * \mathbf{X}_{\text{Inter}} + \gamma * \mathbf{X}_{\text{Intra}}, \mathcal{E}). \quad (13)$$

- Serial Fusion** feeds $\mathbf{X}_{\text{Inter}}$ directly into the LSR stage, ensuring that individual preference modeling is directly informed by cohort-level patterns:

$$\mathbf{X}_{\text{F}} = \text{FFN}(\mathbf{X}_{\text{Intra}}, \mathcal{E}). \quad (14)$$

Finally, we integrate the contextual-enhanced representation via a gated residual update:

$$\begin{aligned} \mathbf{X}_{\text{out}} &= (1 - \alpha) * \mathbf{X}_{\text{SA}} + \alpha * \mathbf{X}_{\text{F}}, \\ \mathbf{X}_{\text{out}} &= \text{LN}(\text{Dropout}(\text{GELU}(\mathbf{X}_{\text{out}})) + \mathbf{X}_{\text{out}}), \end{aligned} \quad (15)$$

where $\mathbf{X}_{\text{out}} \in \mathbb{R}^{B \times L \times D}$, and α is the hyper-parameter. The final recommendation scores are obtained by computing the similarity between \mathbf{X}_{out} and the item embedding matrix \mathcal{M} :

$$\bar{y}_i = \mathbf{X}_{\text{out}} * \mathcal{M}^{\top}. \quad (16)$$

Training Objective

Our training objective is composed of two parts: the Cross-Entropy (CE) and the Frequency Domain consistency loss.

Cross-Entropy Loss. In the experiment, we select CE loss as the first train loss \mathcal{L}_{CE} . CE loss regards predicting future items as a categorization task across the item set. The process of CE loss is expressed as:

$$\mathcal{L}_{\text{CE}} = -\log \frac{\exp(\bar{y}_t)}{\sum_{i \in |\mathcal{I}|} \exp(\bar{y}_i)}, \quad (17)$$

where \bar{y}_t is the ground truth item.

Frequency Domain Loss. We transform the prediction P and target T into the frequency domain. We define $\mathcal{L}_{\text{L1}}(P, T) = \frac{1}{N} \sum_{i=1}^N |P_i - T_i|$, $\mathcal{L}_{\text{L2}}(P, T) = \frac{1}{N} \sum_{i=1}^N (P_i - T_i)^2$, and $\mathcal{L}_{\text{mix}} = \frac{1}{2}(\mathcal{L}_{\text{L1}} + \mathcal{L}_{\text{L2}})$. The distance function $\mathcal{F}_{\mathcal{D}}$ is selected from:

$$\mathcal{F}_{\mathcal{D}}(P, T) \in \{\mathcal{L}_{\text{L1}}(P, T), \mathcal{L}_{\text{L2}}(P, T), \mathcal{L}_{\text{mix}}(P, T)\} \quad (18)$$

Dataset	Beauty				Sports & Outdoors				Toys & Games			
	Metric	HR@10	HR@20	NDCG@10	NDCG@20	HR@10	HR@20	NDCG@10	NDCG@20	HR@10	HR@20	NDCG@10
GRU4Rec (2015)	0.0304	0.0527	0.0147	0.0203	0.0187	0.0303	0.0101	0.0131	0.0211	0.0348	0.0106	0.0140
Caser (2018)	0.0225	0.0403	0.0108	0.0153	0.0163	0.0260	0.0080	0.0104	0.0161	0.0268	0.0079	0.0106
SASRec (2018)	0.0531	0.0823	0.0283	0.0356	0.0298	0.0459	0.0159	0.0200	0.0652	0.0929	0.0366	0.0435
BERT4Rec (2019)	0.0705	0.1073	0.0387	0.0480	0.0428	0.0649	0.0229	0.0284	0.0635	0.0939	0.0353	0.0430
DuoRec (2022)	0.0800	0.1088	0.0483	0.0555	0.0446	0.0640	0.0263	0.0312	0.0816	0.1080	0.0515	0.0582
ICLRec (2022)	0.0726	0.1055	0.0400	0.0483	0.0422	0.0632	0.0227	0.0280	0.0826	0.1137	0.0473	0.0552
CL4SRec (2022)	0.0706	0.0990	0.0417	0.0488	0.0375	0.0575	0.0201	0.0251	0.0863	0.1143	0.0522	0.0592
MAERec (2023)	0.0789	0.1094	0.0472	0.0548	0.0435	0.0645	0.0239	0.0292	0.0823	0.1108	0.0499	0.0570
SASRec _F [♣] (2024)	0.0849	0.1187	0.0421	0.0506	0.0488	0.0711	0.0231	0.0287	0.0875	0.1082	0.0571	0.0623
MSSR [♣] (2024)	0.0897	0.1281	0.0448	0.0545	0.0540	0.0785	0.0255	0.0317	0.0982	0.1201	0.0608	0.0684
FMLPRec (2022)	0.0559	0.0869	0.0291	0.0369	0.0336	0.0525	0.0183	0.0231	0.0671	0.0974	0.0365	0.0441
FEARec [♣] (2023b)	0.0927	0.1282	0.0555	0.0645	0.0539	0.0783	0.0315	0.0376	0.1005	0.1357	0.0606	0.0708
BSARec [♣] (2024)	0.0944	0.1295	0.0574	0.0672	<u>0.0581</u>	<u>0.0830</u>	0.0333	0.0387	0.1023	0.1379	0.0610	0.0713
DIFF [♣] (2025)	0.0940	<u>0.1301</u>	0.0526	0.0631	0.0580	0.0829	0.0310	0.0381	0.0969	0.1210	0.0607	0.0703
FreqRec (S)*	0.0962	0.1299	0.0589	0.0674	0.0571	0.0825	0.0334	0.0398	0.1046	0.1459	0.0655	0.0734
FreqRec (P)*	0.0989	0.1359	0.0601	0.0686	0.0583	0.0859	0.0342	0.0401	0.1044	0.1468	0.0653	0.0735
Improve.	4.77%	4.46%	4.70%	2.08%	0.34%	3.49%	2.70%	3.62%	2.25%	6.45%	7.38%	3.09%

Table 1: Overall performance comparison on the three datasets. S and P denote FreqRec’s serial and parallel structures, respectively. Best and suboptimal results are **bold** and underlined, respectively. * denotes a statistically significant improvement of FreqRec over the best competing model ($p < 0.05$). Models marked with [♣] indicate values obtained from our own re-implementation of the corresponding baseline, other results retrieved from BSARec (Shin et al. 2024). *Improve.* is FreqRec’s relative improvement over the strongest baseline. We employ **FreqRec (P)** as our standard model for the following experiments.

where $P, T \in \mathbb{R}^{B \times L \times D}$. The final frequency-domain loss $\mathcal{L}_{\mathcal{F}}$ separately computes the distance between the real and imaginary parts of P and T after the DFT \mathcal{F} :

$$\mathcal{L}_{\mathcal{F}} = \mathcal{F}_{\mathcal{D}}(\mathbf{Real}(\mathcal{F}(P)), \mathbf{Real}(\mathcal{F}(T))) + \mathcal{F}_{\mathcal{D}}(\mathbf{Imag}(\mathcal{F}(P)), \mathbf{Imag}(\mathcal{F}(T))). \quad (19)$$

Finally, our final loss function $\mathcal{L}_{\mathcal{SR}}$ is expressed as:

$$\mathcal{L}_{\mathcal{SR}} = (1 - \beta) * \mathcal{L}_{\mathcal{F}} + \beta * \mathcal{L}_{\mathcal{CE}}, \quad (20)$$

where β control the strengths of the frequency-domain loss $\mathcal{L}_{\mathcal{F}}$ and Cross-Entropy loss $\mathcal{L}_{\mathcal{CE}}$.

Experiments

Experimental Setup

Datasets. We conducted a comprehensive evaluation on three widely adopted SR benchmarks¹: *Beauty*, *Sports & Outdoors*, and *Toys & Games*. These datasets, derived from the Amazon e-commerce platform, contain product reviews from 1996 to 2014 and are widely used for the SR task. We adopted the data preprocessing approach from previous work (Shin et al. 2024; Kim et al. 2025).

Evaluation Metrics. To comprehensively evaluate the model’s recommendation accuracy, we employ the top- K evaluation method commonly used in SR, including Hit Rate ($\text{HR}@K$) and Normalized Discounted Cumulative Gain ($\text{NDCG}@K$), with K set at 10 and 20.

Baselines. To comprehensively evaluate FreqRec, we benchmark its performance against fourteen SR baselines:

- RNN or Transformer-based approaches: GRU4Rec (Hidasi et al. 2015), Caser (Tang and Wang 2018), SASRec (Kang and McAuley 2018), BERT4Rec (Sun et al. 2019),

- DuoRec (Qiu et al. 2022), ICLRec (Chen et al. 2022), CL4SRec (Xie et al. 2022), MAERec (Ye, Xia, and Huang 2023), SASRec_F (Lin et al. 2024), MSSR (Lin et al. 2024).
- Frequency-based approaches: FMLPRec (Zhou et al. 2022), FEARec (Du et al. 2023b), BSARec (Shin et al. 2024), DIFF (Kim et al. 2025).

Implementation Details Experiments were conducted on an NVIDIA RTX 4090 (24 GB) using PyTorch. We employed the Adam optimizer with 10-round early stopping, and a learning rate grid-searched from $\{5e-4, 1e-4, 1e-3\}$. For fair comparison, we followed prior work (Shin et al. 2024; Kim et al. 2025) setting $D = 64$ and $L = 50$. For FreqRec, we grid-searched $\gamma, \alpha \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$, $\beta \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, batch size $B \in \{32, 64, 128, 256, 512\}$, and $\mathcal{F}_{\mathcal{D}} \in \{\mathcal{L}_{L1}, \mathcal{L}_{L2}, \mathcal{L}_{\text{mix}}\}$.

Overall Performance

Table 1 compares FreqRec against leading baselines on *Beauty*, *Sports & Outdoors*, and *Toys & Games* benchmarks. FreqRec consistently delivers superior performance in terms of $\text{HR}@10/20$ and $\text{NDCG}@10/20$. A closer examination yields the following key insights:

- **Frequency-based approaches demonstrate strong performance.** Consistent with prior findings (Shin et al. 2024), methods leveraging frequency-based representations outperform transformer-based models in most cases. On *Beauty*, $\text{HR}@10/20$ for BSARec is 0.0944/0.1295 compared to 0.0897/0.1281 for MSSR. This gain arises because frequency-domain modules recover high-frequency and periodic signals, patterns that standard self-attention, acting as a low-pass filter, inherently attenuates.
- **FreqRec exhibits superior performance over Frequency-based models.** Across all three benchmark datasets, FreqRec demonstrates consistent improvements over the strongest baseline BSARec, with performance

¹https://cseweb.ucsd.edu/~jmcauley/datasets/amazon_v2/

Model	Beauty				Toys & Games			
	H@10	H@20	N@10	N@20	H@10	H@20	N@10	N@20
w/o SA	0.0959	0.1279	0.0587	0.0667	0.1034	0.1338	0.0644	0.0713
w/o GSA	0.0881	0.1202	0.0537	0.0618	0.0954	0.1295	0.0568	0.0672
w/o LSR	0.0888	0.1215	0.0533	0.0616	0.0961	0.1268	0.0606	0.0684
w/o GSA+LSR	0.0787	0.1097	0.0481	0.0559	0.0719	0.0956	0.0436	0.0495
w/o $\mathcal{L}_{\mathcal{F}}$	0.0969	0.1300	0.0582	0.0665	0.0990	0.1342	0.0619	0.0708
w/o \mathcal{L}_{CE}	0.0807	0.1081	0.0477	0.0548	0.0714	0.0975	0.0434	0.0499
FreqRec	0.0989	0.1359	0.0601	0.0686	0.1044	0.1468	0.0653	0.0735

Table 2: Ablation study on Beauty and Toys & Games. "H" and "N" denote HR@K and NDCG@K, respectively.

gains ranging from 0.34% to 7.38%. On the Toys & Games dataset, FreqRec achieves an NDCG@10 of 0.0655, while BSARec attains 0.0610, representing a performance improvement of 7.38%. These results provide compelling evidence for the effectiveness of the learnable inter/intra-session frequency MLPs and Fourier loss mechanism in enabling accurate recommendations.

- **Parallel modeling of inter/intra sequence interactions outperforms serial approaches.** We implement two FreqRec variants, one that captures inter- and intra-session dependencies in parallel, and another that processes them sequentially. The parallel design consistently achieves superior performance. We attribute this gap to an information bottleneck in the serial design, where the user-perspective features overwrite raw sequence signals, impeding the recovery of fine-grained dynamics. In contrast, Parallel Fusion preserves the integrity of both spectral streams, enabling the model to leverage complementary cohort and individual cues without mutual interference.

Ablation Studies

To illustrate the effectiveness of our proposed Frequency domain MLPs and the Fourier loss in FreqRec, we conduct ablation studies on Beauty and Toys & Games datasets. As demonstrated in Table 2, when we disable inter-session interaction modeling (w/o GSA.) or intra-session interaction modeling (w/o LSR.), overall performance degrades by a comparable margin. Crucially, omitting both components results in over 20% on average performance drop relative to FreqRec, confirming that both inter and intra frequency signals complement self-attention representations. Furthermore, the slightly greater performance drop observed for the w/o GSA. variant highlights the pivotal role of capturing mutual interests between sequences/users in enhancing recommendation accuracy. Beyond architectural components, the auxiliary frequency-domain loss $\mathcal{L}_{\mathcal{F}}$ also proves essential. Disabling this loss produces a noticeable performance decline on both datasets: on Toys & Games, HR@20 drops from 0.1468 to 0.1342, and NDCG@20 from 0.0735 to 0.0708. Conversely, without cross-entropy loss \mathcal{L}_{CE} yields the most severe degradation, reaffirming that point-wise prediction accuracy remains the primary supervisory signal. Taken together, these results demonstrate that FreqRec’s performance gains stem from the synergy between the joint optimization of cross-entropy and frequency-domain loss.

Model	Beauty		Toys & Games	
	HR@10	NDCG@10	HR@10	NDCG@10
SASRec	0.0531	0.0283	0.0640	0.0366
+ $\mathcal{L}_{\mathcal{F}}$	0.0592 ^{+11.49%}	0.0309 ^{+9.19%}	0.0667 ^{+4.22%}	0.0374 ^{+2.19%}
DuoRec	0.0890	0.0542	0.1027	0.0635
+ $\mathcal{L}_{\mathcal{F}}$	0.0911 ^{+2.36%}	0.0554 ^{+2.21%}	0.1030 ^{+0.29%}	0.0647 ^{+1.89%}
FMLPRec	0.0559	0.0291	0.0628	0.0356
+ $\mathcal{L}_{\mathcal{F}}$	0.0644 ^{+15.21%}	0.0349 ^{+19.93%}	0.0686 ^{+9.24%}	0.0389 ^{+9.27%}
BSARec	0.0871	0.0437	0.1065	0.0658
+ $\mathcal{L}_{\mathcal{F}}$	0.0972 ^{+11.60%}	0.0640 ^{+46.45%}	0.1077 ^{+1.13%}	0.0668 ^{+1.52%}
AVG. Improve	10.165%	19.455%	3.720%	3.718%

Table 3: Impact of the frequency-domain consistency loss (+ $\mathcal{L}_{\mathcal{F}}$). Red superscripts indicate the improvement rates.

Model	Automotive		CDs		Grocery	
	H@10	N@10	H@10	N@10	H@10	N@10
FMLPRec	0.0320	0.0204	0.0367	0.0207	0.1043	0.0705
BSARec	0.0514	0.0430	0.0537	0.0438	0.1414	0.1215
w/o GSA	0.0495	0.0411	0.0520	0.0425	0.1371	0.1177
w/o LSR	0.0507	0.0423	0.0524	0.0443	0.1379	0.1182
w/o GSA+LSR	0.0463	0.0387	0.0518	0.0402	0.1346	0.1149
w/o $\mathcal{L}_{\mathcal{F}}$	0.0501	0.0426	0.0538	0.0426	0.1402	0.1201
FreqRec	0.0521	0.0445	0.0549	0.0475	0.1434	0.1242

Table 4: Performance under noisy scenario in terms of HR@10 and NDCG@10 across the three target domains.

Effectiveness of Frequency Domain Loss

In this section, we assess the effectiveness and flexibility of our Fourier loss $\mathcal{L}_{\mathcal{F}}$ by integrating it into four representative sequential recommendation baselines on the Beauty and Toys & Games datasets. To ensure experimental fairness, we adhere strictly to each model’s original implementation and hyper-parameter settings. As shown in Table 3, our proposed Fourier loss consistently improves baseline performance in a true plug-and-play fashion, recording average improvements over both datasets in terms of HR@10 and NDCG@10 ranging from 3.70% to 19.45%. It is noteworthy that models which already exploit frequency domain features (FMLPRec and BSARec) derive the greatest benefit. FMLPRec’s HR@10 climbs by 15.21% (from 0.0559 to 0.0644) and NDCG@10 by 19.93% (from 0.0291 to 0.0349), while BSARec secures a 46.45% gain in NDCG@10 (from 0.0437 to 0.0639). These considerable improvements underscore that explicitly penalizing frequency discrepancies provides a complementary signal to cross-entropy loss, guiding DFT-based models toward more discriminative feature learning.

Effectiveness of Handling Noisy Data

We benchmark FreqRec under the noise-resilience ability following the protocol of previous work (Liu et al. 2025). Instead of the standard supervised paradigm, we train the model on the aggregate interaction logs from three Amazon categories, Automotive, CDs & Vinyl, and Grocery & Gourmet Food, and then evaluate it on each category independently, with no fine-tuning on the target domain. This evaluation setting requires the model to distinguish meaningful behavioral signals from heterogeneous

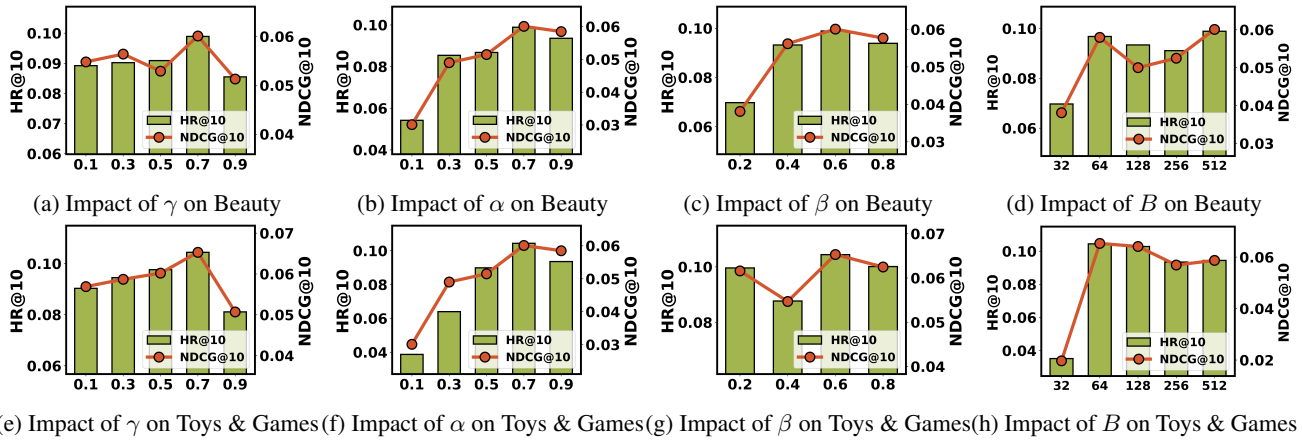


Figure 4: Sensitivity test of hyper-parameters on the Beauty and Toys & Games datasets.

and domain-specific noise in the training data.

Table 4 presents HR@10 and NDCG@10 for each target domain, comparing FreqRec against Frequency-based baselines: FMLPRec, BSARec. FreqRec achieves consistently superior results across all three target categories. Align with the observations in Table 2, removing either the GSA or the LSR module degrades performance in every domain, confirming that cohort-level denoising and user-level refinement are complementary. Eliminating both frequency-driven modules produces the most pronounced degradation, demonstrating their synergistic effect: the Inter-module suppresses pervasive noise patterns, while the intra-module restores high-frequency, user-specific signals. These results confirm that combining dual-path spectral modules with a frequency-domain consistency loss yields both robust noise suppression and strong generalization.

Performance *w.r.t* Data Sparsity

To assess the impact of data sparsity on model performance, we focus on users with below-average sequence lengths and group them by interaction count (e.g., 5-6, 7-8) based on their interaction numbers, and evaluate our FreqRec against leading baselines (i.e., SASRec, FMLPRec, BSARec) on *Sport & Outdoors* and *Beauty* datasets. As shown in Table 5, all models suffer performance drops on sparser sequences. However, FreqRec consistently demonstrates greater stability than the other models. This can be attributed to our approach, which cohort-level aggregation with user-specific refinement, compensating for the information sparsity inherent in short user histories.

Sensitivity Test on Hyper-Parameters

Figure 4 presents the hyper-parameter sensitivity analysis of FreqRec, reporting HR@10 and NDCG@10 performance on *Beauty* and *Toys & Games* datasets.

- **Sensitivity to GSA/LSR learner feature weight γ** : In the sub Figure 4 (a, c), performance peaks at $\gamma = 0.7$ for both datasets, indicating the need for balanced feature constraints between learners.

Model	Sports & Outdoors				Beauty			
	H@5	N@5	H@5	N@5	H@5	N@5	H@5	N@5
SASRec	0.0062	0.0044	0.0041	0.0028	0.0202	0.0135	0.0209	0.0131
FMLPRec	0.0061	0.0044	0.0046	0.0029	0.0254	0.0172	0.0080	0.0047
BSARec	0.0170	0.0121	0.0106	0.0093	0.0300	0.0218	0.0213	0.0149
FreqRec	0.0190	0.0135	0.0123	0.0101	0.0325	0.0232	0.0239	0.0163

Table 5: Performance of FreqRec against baselines under different sparse interaction degrees.

- **Sensitivity to Self-Attention and FreqNet α** : In the sub Figure 4 (b, f), performance peaks at $\alpha = 0.7$, showing excessive noise filtering can harm performance.
- **Sensitivity to loss balance parameter β** : In the sub Figure 4 (c, g), optimal performance occurs at $\beta = 0.6$, with performance trends varying across datasets.
- **Sensitivity to Batch Size B** : In the sub Figure 4 (d, h), smaller batches ($B = 32$) negatively impact performance. Best results for Beauty are at $B = 64$ or 512 , while Toys & Games peaks at $B = 64$, highlighting the role of batch size in noise filtering and performance.

Conclusion

In this paper, we present FreqRec, a Frequency-Enhanced Dual-Path Network for sequential recommendation that jointly models inter-session and intra-session dynamics via a learnable, complex-valued Fourier transform. By coupling a Global Spectral Aggregator for cohort-level rhythms with a Local Spectral Refiner for user-specific signals, FreqRec effectively suppresses noise and restores high-frequency behavioral patterns that standard attention mechanisms smooth out. We further introduce a frequency-domain consistency loss, which explicitly aligns the model’s predicted spectral coefficients with ground-truth signatures, thereby bridging the gap left by conventional time-domain objectives. Extensive experiments on three public benchmarks show that FreqRec delivers state-of-the-art performance, while maintaining robustness under data sparsity and noisy logs scenarios.

Acknowledgements

We would like to thank the anonymous reviewers for their valuable discussion and constructive feedback. This work is supported by National Natural Science Foundation of China (U22B2061), National Key R&D Program of China (2022YFB4300603), and Natural Science Foundation of Sichuan, China (2024NSFSC0496).

References

- Agichtein, E.; Brill, E.; and Dumais, S. 2006. Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, 19–26.
- Chen, Y.; Liu, Z.; Li, J.; McAuley, J.; and Xiong, C. 2022. Intent contrastive learning for sequential recommendation. In *Proceedings of the ACM web conference 2022*, 2172–2182.
- Donkers, T.; Loepf, B.; and Ziegler, J. 2017. Sequential user-based recurrent neural network recommendations. In *Proceedings of the eleventh ACM conference on recommender systems*, 152–160.
- Dovonon, G. J.-S.; Bronstein, M. M.; and Kusner, M. 2024. Setting the Record Straight on Transformer Oversmoothing. In *ICLR 2024 Workshop on Reliable and Responsible Foundation Models*.
- Du, X.; Yuan, H.; Zhao, P.; Fang, J.; Liu, G.; Liu, Y.; Sheng, V. S.; and Zhou, X. 2023a. Contrastive enhanced slide filter mixer for sequential recommendation. In *2023 IEEE 39th International Conference on Data Engineering (ICDE)*, 2673–2685. IEEE.
- Du, X.; Yuan, H.; Zhao, P.; Qu, J.; Zhuang, F.; Liu, G.; Liu, Y.; and Sheng, V. S. 2023b. Frequency enhanced hybrid attention network for sequential recommendation. In *Proceedings of the 46th International ACM SIGIR conference on research and development in information retrieval*, 78–88.
- Duhamel, P.; and Vetterli, M. 1990. Fast Fourier transforms: a tutorial review and a state of the art. *Signal processing*, 19(4): 259–299.
- Fang, H.; Zhang, D.; Shu, Y.; and Guo, G. 2020. Deep learning for sequential recommendation: Algorithms, influential factors, and evaluations. *ACM Transactions on Information Systems (TOIS)*, 39(1): 1–42.
- Frigo, M.; and Johnson, S. G. 2005. The design and implementation of FFTW3. *Proceedings of the IEEE*, 93(2): 216–231.
- Gan, Y.; Liu, Q.; Luo, D.; Hou, R.; Cai, Y.; Lin, R.; and Shi, X. 2025. Pareto selective error feedback suppression for popularity–diversity balanced session-based recommendation. *Engineering Applications of Artificial Intelligence*, 142: 109911.
- Guo, J.; Zhang, P.; Li, C.; Xie, X.; Zhang, Y.; and Kim, S. 2022. Evolutionary preference learning via graph nested gru ode for session-based recommendation. In *Proceedings of the 31st ACM international conference on information & knowledge management*, 624–634.
- Hidasi, B.; Karatzoglou, A.; Baltrunas, L.; and Tikk, D. 2015. Session-based Recommendations with Recurrent Neural Networks. *CoRR*, abs/1511.06939.
- Kang, W.-C.; and McAuley, J. 2018. Self-attentive sequential recommendation. In *2018 IEEE international conference on data mining (ICDM)*, 197–206. IEEE.
- Kim, H.; Choi, M.; Lee, S.; and Lee, I. B. J. 2025. DIFF: Dual Side-Information Filtering and Fusion for Sequential Recommendation. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*.
- Lee-Thorp, J.; Ainslie, J.; Eckstein, I.; and Ontanon, S. 2022. FNet: Mixing Tokens with Fourier Transforms. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4296–4313.
- Li, Z.; Zhao, H.; Liu, Q.; Huang, Z.; Mei, T.; and Chen, E. 2018. Learning from history and present: Next-item recommendation via discriminatively exploiting user behaviors. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 1734–1743.
- Liang, T.-P.; Lai, H.-J.; and Ku, Y.-C. 2006. Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. *Journal of Management Information Systems*, 23(3): 45–70.
- Lin, X.; Luo, J.; Pan, J.; Pan, W.; Ming, Z.; Liu, X.; Huang, S.; and Jiang, J. 2024. Multi-sequence attentive user representation learning for side-information integrated sequential recommendation. In *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, 414–423.
- Liu, H.; Deng, Z.; Wang, L.; Peng, J.; and Feng, S. 2023a. Distribution-based learnable filters with side information for sequential recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems*, 78–88.
- Liu, Q.; Zeng, Y.; Mokhosi, R.; and Zhang, H. 2018. STAMP: short-term attention/memory priority model for session-based recommendation. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 1831–1839.
- Liu, S.; Zhang, A.; Hu, G.; Qian, H.; and Chua, T.-S. 2025. Preference Diffusion for Recommendation. In *The Thirteenth International Conference on Learning Representations*.
- Liu, X.; Li, Z.; Gao, Y.; Yang, J.; Cao, T.; Wang, Z.; Yin, B.; and Song, Y. 2023b. Enhancing user intent capture in session-based recommendation with attribute patterns. *Advances in Neural Information Processing Systems*, 36: 30821–30839.
- Niu, Y.; Xing, X.; Jia, Z.; Liu, R.; Xin, M.; and Cui, J. 2024. Diffusion recommendation with implicit sequence influence. In *Companion Proceedings of the ACM Web Conference 2024*, 1719–1725.
- Qiao, S.; Zhou, W.; Wen, J.; Zhang, H.; and Gao, M. 2023. Bi-channel multiple sparse graph attention networks for session-based recommendation. In *Proceedings of the 32nd*

- ACM international conference on information and knowledge management, 2075–2084.
- Qiu, R.; Huang, Z.; Li, J.; and Yin, H. 2020. Exploiting cross-session information for session-based recommendation with graph neural networks. *ACM Transactions on Information Systems (TOIS)*, 38(3): 1–23.
- Qiu, R.; Huang, Z.; Yin, H.; and Wang, Z. 2022. Contrastive learning for representation degeneration problem in sequential recommendation. In *Proceedings of the fifteenth ACM international conference on web search and data mining*, 813–823.
- Rendle, S.; Freudenthaler, C.; Gantner, Z.; and Schmidt-Thieme, L. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, 452–461.
- Rendle, S.; Freudenthaler, C.; and Schmidt-Thieme, L. 2010. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th international conference on World wide web*, 811–820.
- Shin, Y.; Choi, J.; Wi, H.; and Park, N. 2024. An attentive inductive bias for sequential recommendation beyond the self-attention. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, 8984–8992.
- Sun, F.; Liu, J.; Wu, J.; Pei, C.; Lin, X.; Ou, W.; and Jiang, P. 2019. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM international conference on information and knowledge management*, 1441–1450.
- Tamkin, A.; Jurafsky, D.; and Goodman, N. 2020. Language through a prism: A spectral approach for multiscale language representations. *Advances in Neural Information Processing Systems*, 33: 5492–5504.
- Tang, J.; and Wang, K. 2018. Personalized top-n sequential recommendation via convolutional sequence embedding. In *Proceedings of the eleventh ACM international conference on web search and data mining*, 565–573.
- Vaswani, A. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*.
- Vombatkere, K.; Mousavi, S.; Zannettou, S.; Roesner, F.; and Gummadi, K. P. 2024. Tiktok and the art of personalization: investigating exploration and exploitation on social media feeds. In *Proceedings of the ACM Web Conference 2024*, 3789–3797.
- Wang, P.; Zheng, W.; Chen, T.; and Wang, Z. 2022. Anti-Oversmoothing in Deep Vision Transformers via the Fourier Domain Analysis: From Theory to Practice. In *International Conference on Learning Representations*.
- Wang, S.; Hu, L.; Wang, Y.; Cao, L.; Sheng, Q.; and Orgun, M. 2019. Sequential recommender systems: Challenges, progress and prospects. In *Twenty-Eighth International Joint Conference on Artificial Intelligence {IJCAI-19}*. International Joint Conferences on Artificial Intelligence Organization.
- Wang, T.-Y.; Chen, C.-T.; Huang, J.-C.; and Huang, S.-H. 2023. Modeling cross-session information with multi-interest graph neural networks for the next-item recommendation. *ACM Transactions on Knowledge Discovery from Data*, 17(1): 1–28.
- Wang, Z.; Wei, W.; Cong, G.; Li, X.-L.; Mao, X.-L.; and Qiu, M. 2020. Global context enhanced graph neural networks for session-based recommendation. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval*, 169–178.
- Xie, X.; Sun, F.; Liu, Z.; Wu, S.; Gao, J.; Zhang, J.; Ding, B.; and Cui, B. 2022. Contrastive learning for sequential recommendation. In *2022 IEEE 38th international conference on data engineering (ICDE)*, 1259–1273. IEEE.
- Xie, Y.; Zhou, P.; and Kim, S. 2022. Decoupled side information fusion for sequential recommendation. In *Proceedings of the 45th international ACM SIGIR conference on research and development in information retrieval*, 1611–1621.
- Yang, X.; Sun, Y.; Chen, X.; et al. 2024. Frequency-aware generative models for multivariate time series imputation. *Advances in Neural Information Processing Systems*, 37: 52595–52623.
- Ye, Y.; Xia, L.; and Huang, C. 2023. Graph masked autoencoder for sequential recommendation. In *Proceedings of the 46th international ACM SIGIR conference on research and development in information retrieval*, 321–330.
- Yi, K.; Zhang, Q.; Fan, W.; Wang, S.; Wang, P.; He, H.; An, N.; Lian, D.; Cao, L.; and Niu, Z. 2023. Frequency-domain mlps are more effective learners in time series forecasting. *Advances in Neural Information Processing Systems*, 36: 76656–76679.
- Ying, X. 2019. An overview of overfitting and its solutions. In *Journal of physics: Conference series*, volume 1168, 022022. IOP Publishing.
- Zhou, K.; Yu, H.; Zhao, W. X.; and Wen, J.-R. 2022. Filter-enhanced MLP is all you need for sequential recommendation. In *Proceedings of the ACM web conference 2022*, 2388–2399.