FREQ: Frequency-Domain Enhanced Cross-Domain Sequential Recommendation via Item Popularity Dynamics

Xin Li, Yong Zhang*, Jing Wang, Shuo Li
Central China Normal University, Wuhan, Hubei, China
3148913485@qq.com, 12081207@qq.com, 1316990692@qq.com, 1799628675@qq.com
*Corresponding author

***Abstract*—Sequential recommendation predicts users' next interests from historical interactions, while cross-domain recommendation transfers this capability across domains but is often hindered by data sparsity and temporal-rhythm discrepancy. To address this, we propose FREQ, a frequency-domain enhancement module built on PREPREC. FREQ constructs multi-scale item popularity windows and applies the discrete Fourier transform to obtain spectral representations, enabling better modeling of long- and short-term dynamics across domains. It further introduces residual frequency-domain injection with lightweight gating and a cross-scale consistency regularization to improve representation stability and transferability. Experiments show that FREQ typically outperforms strong baselines, with larger gains when the source and target domains exhibit greater periodic discrepancy.**

I. Introduction

Sequential recommendation (SR) predicts users' next interests from historical interactions, while cross-domain recommendation (CDR) transfers this capability from a source domain to a target domain. In practice, CDR is hindered by two major challenges: sparse interactions, which limit reliable preference learning and item representation learning [1], [2], and domain discrepancy, where differences in user behavior and item popularity degrade transfer performance [3].

Existing methods typically alleviate sparsity by transferring user/item knowledge from data-rich domains, often with auxiliary information or target-domain interactions. To reduce domain discrepancy, representative solutions include adversarial alignment and federated aggregation, such as DAA and PPCDR [4], [5], but they usually require additional target-domain supervision, unstable optimization, or high communication overhead. PREPREC learns general representations from timestamped popularity dynamics without side information or target-domain fine-tuning [6]. However, it still has three limitations: 1) it does not explicitly model cross-domain differences in behavioral rhythms; 2) it does not clearly separate long-term trends from short-term fluctuations; and 3) its multi-scale popularity signals are concatenated without explicit consistency constraints.

To address these limitations, we propose FREQ, a frequency-domain enhancement mechanism for cross-domain sequential recommendation. FREQ first constructs multi-scale popularity windows and applies the discrete Fourier transform (DFT) to obtain spectral representations, which make long- and short-term dynamics more separable and improve comparability across domains with different dominant periods. It then injects frequency-domain features into the backbone through residual connections and lightweight gating, enabling the model to capture both the magnitude and rhythm of popularity dynamics without modifying the backbone architecture. In addition, we introduce cross-scale consistency regularization in both spectral and prediction spaces to stabilize multi-scale representations.

The main contributions are as follows:

- We propose FREQ, a plug-in frequency-domain enhancement mechanism for zero-shot cross-domain sequential recommendation.
- We combine multi-scale temporal windows with spectral analysis to model long-term trends, short-term fluctuations, and cross-domain rhythm differences.
- We design residual spectral injection with lightweight gating and a cross-scale consistency regularization to improve transfer robustness and representation stability.

Experiments on multiple public cross-domain datasets show that FREQ typically outperforms strong baselines, especially when the source and target domains exhibit substantially different dominant periods. Code is available at: https://pan.baidu.com/s/10uBm5tPNfCSMr3eeAvLdEg?pwd=952t.

II. Related work

Sequential Recommendation Sequential recommendation models users' next-item preference from interaction histories. Representative methods evolve from Markov-chain-based models to deep sequential architectures such as RNNs, LSTMs, and GRUs, and more recently to self-attention-based models such as SASRec, BERT4Rec, and TiSASRec, which improve long-range dependency modeling and temporal awareness [7]–[11]. Self-supervised and contrastive methods further enhance the robustness of sequence representations [12], [13]. However, most existing SR methods are built on in-domain ID representations and local temporal dependencies, which limits their generalization in cross-domain and weakly supervised settings.

Supported by the Project of the Humanities and Social Sciences Research of the Ministry of Education (Grant No. 25YJAZH251).

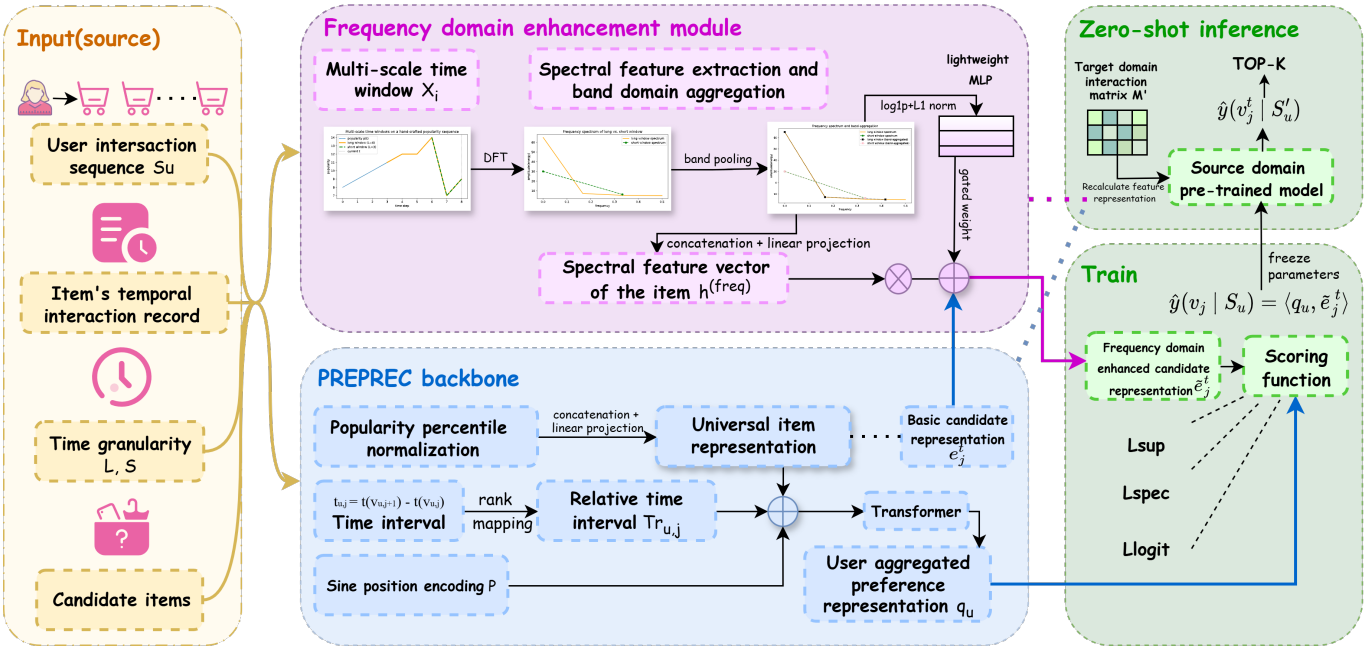

Fig. 1. The overview of FREQ.

Cross-domain Recommendation Cross-domain recommendation transfers knowledge from source domains to improve recommendation in data-sparse target domains. Existing methods mainly rely on shared features, latent-space mapping, transfer learning, deep representation learning, or multimodal fusion [14]–[19]. Although effective, these approaches focus primarily on representation alignment across domains, while paying limited attention to temporal-rhythm differences such as domain-specific interaction frequency and behavioral cycles. This limits their adaptability when the two domains exhibit asynchronous temporal patterns.

III. Method

This section introduces the FREQ framework (Fig. 1), including the PREPREC backbone, the frequency-domain module, and the zero-shot inference process.

A. Base Model Structure

PREPREC models cross-domain sequential behavior through item popularity dynamics at multiple temporal scales. It contains three components: a popularity encoding layer, a temporal–positional encoding layer, and a Transformer-based sequential modeling layer.

For popularity encoding, PREPREC aggregates the interactions of item v_j into long- and short-period popularity sequences, denoted by a_j^t and b_j^t, and converts them into percentile-based representations $\mathrm{P}(a_j^t)$ and $\mathrm{P}(b_j^t)$. After popularity encoding, the long- and short-scale representations are concatenated and linearly projected into a universal item representation:

$$\mathbf{e}_j^t = \mathbf{W}_{\mathrm{p}}\left[\mathbf{P}_j^t \,\|\, \mathbf{H}_j^t\right], \tag{1}$$

where $\mathbf{W}_{\mathrm{p}}$ is a learnable projection matrix, and $\mathbf{P}_j^t$ and $\mathbf{H}_j^t$ denote the long- and short-scale popularity representations before time t.

In the temporal–positional encoding layer, PREPREC models both relative time intervals and sequence order. For user u with interaction sequence S_u, the interval between two consecutive interactions is $t_{u,j} = t(v_{u,j+1}) - t(v_{u,j})$, and its rank within the user sequence is $r_{u,j} = \mathrm{rank}(t_{u,j})$. The rank is mapped to a sinusoidal embedding $\mathbf{T}_{r_{u,j}}$, while each position l is associated with a positional encoding $\mathbf{P}_l$.

In the sequential modeling layer, the Transformer input is $\mathbf{x}_{u,l} = \mathbf{e}_{u,l} + \mathbf{T}_{u,l} + \mathbf{P}_l$. Stacked Transformer layers produce the user representation $\mathbf{q}_u$. The prediction score for candidate item v_j is then computed by

$$\hat{y}(v_j \mid S_u) = \left\langle \mathbf{q}_u, \mathbf{e}_j^{(\mathrm{curr})} \right\rangle, \tag{2}$$

and candidate items are ranked accordingly.

B. Frequency-Domain Modeling and Residual Signal Injection Mechanism

1) Multi-Scale Time Window Design: PREPREC summarizes popularity at the current step but does not preserve its temporal evolution. To model such dynamics, FREQ treats popularity as a time-evolving signal and constructs multi-scale historical windows. Let $p_v(t)$ denote the interaction count of item v at time step t. For item v at reference time t, the window at scale s is

$$\mathbf{X}_{v,t}^{(s)} = \left[p_v(t-L_s),\, p_v(t-L_s+1),\, \ldots,\, p_v(t-1)\right], \quad s \in \{\mathrm{L}, \mathrm{S}\}, \tag{3}$$

where the long scale captures relatively stable low-frequency trends and the short scale captures high-frequency local variations. The resulting windows $\mathbf{X}_{v,t}^{(\mathrm{L})}$ and $\mathbf{X}_{v,t}^{(\mathrm{S})}$ are then fed

into the frequency-domain module. For clarity, all frequency-domain quantities are defined at the item–time level, while recommendation scores are defined at the user–time–item level.

2) Spectral Feature Extraction and Band-wise Aggregation: Multi-scale windows capture popularity dynamics in the time domain, but do not explicitly represent periodic strength. FREQ therefore applies the discrete Fourier transform (DFT) to each window. For item–time pair (v,t) and scale $s \in \{\mathrm{L},\mathrm{S}\}$, let $F_{v,t}^{(s)}(k)$ denote the DFT of $\mathbf{X}_{v,t}^{(s)}$ at frequency bin k. We use the squared magnitude $|F_{v,t}^{(s)}(k)|^2$ as spectral energy, which emphasizes periodic strength while discarding phase information.

To reduce redundancy and noise, FREQ partitions the spectrum into geometric frequency bands and computes the root mean square (RMS) energy for each band:

$$Z_{v,t,b}^{(s)} = \sqrt{\frac{1}{|B_b^{(s)}|}\sum_{k\in B_b^{(s)}} |F_{v,t}^{(s)}(k)|^2}, \tag{4}$$

where $b = 1,\ldots,B_s$ and $s \in \{\mathrm{L},\mathrm{S}\}$.

The long and short scales use different band numbers to balance low- and high-frequency resolution. After compression and normalization, we obtain spectral descriptors $\mathbf{f}_{v,t}^{(\mathrm{L})}$ and $\mathbf{f}_{v,t}^{(\mathrm{S})}$. They are concatenated and projected into a unified hidden space:

$$\mathbf{h}_{v,t}^{(\mathrm{freq})} = \mathrm{LayerNorm}\Big(\mathbf{W}_{\mathrm{f}}\left[\mathbf{f}_{v,t}^{(\mathrm{L})} \,\|\, \mathbf{f}_{v,t}^{(\mathrm{S})}\right]\Big), \tag{5}$$

where $\mathbf{h}_{v,t}^{(\mathrm{freq})}$ characterizes the multi-scale periodic energy structure of item v at time t.

3) Frequency Domain Residual Injection Strategy: To incorporate spectral information without disturbing the backbone representation, FREQ injects frequency-domain features into the candidate item representation in a residual manner:

$$\tilde{\mathbf{e}}_v^t = \mathbf{e}_v^t + \alpha_{v,t}\,\mathbf{h}_{v,t}^{(\mathrm{freq})}, \tag{6}$$

where $\mathbf{e}_v^t$ is the base representation of item v at time t, $\mathbf{h}_{v,t}^{(\mathrm{freq})}$ is its frequency-domain representation, and $\alpha_{v,t} \in [0,1]$ is an adaptive gating coefficient. Since the spectral branch characterizes item-side periodic popularity, residual injection enables effective matching with the user-side context representation $\mathbf{q}_{u,t}$.

The gate $\alpha_{v,t}$ is generated by a lightweight multilayer perceptron:

$$\alpha_{v,t} = \sigma\big(\mathbf{W}_2\,\mathrm{ReLU}(\mathbf{W}_1\mathbf{h}_{v,t}^{(\mathrm{freq})})\big), \tag{7}$$

where $\sigma(\cdot)$ is the Sigmoid function. This allows the model to adaptively control the amount of frequency-domain injection.

4) Cross-Scale Consistency Constraint: To stabilize multi-scale frequency modeling, FREQ imposes consistency constraints in both the spectral and prediction spaces. For each user–time pair (u,t), let $\Omega_{u,t} = \{v_{u,t+1}\} \cup \mathcal{N}_{u,t}$ denote the candidate set, where $v_{u,t+1}$ is the ground-truth next item and $\mathcal{N}_{u,t}$ is the sampled negative set. Each candidate item $v \in \Omega_{u,t}$ is associated with two spectral descriptors, $\mathbf{f}_{v,t}^{(\mathrm{L})}$ and $\mathbf{f}_{v,t}^{(\mathrm{S})}$.

In the spectral space, we align the directions of the two descriptors by minimizing the cosine distance:

$$\mathcal{L}_{\mathrm{spec}} = \frac{1}{N}\sum_u\sum_t\sum_{v\in\Omega_{u,t}}\Big[1-\cos\big(\mathbf{f}_{v,t}^{(\mathrm{L})},\mathbf{f}_{v,t}^{(\mathrm{S})}\big)\Big], \tag{8}$$

where $N = \sum_u\sum_t|\Omega_{u,t}|$. This term reduces cross-scale drift in spectral structure.

In the prediction space, let $\hat{y}_{u,t,v}^{(\mathrm{L})}$ and $\hat{y}_{u,t,v}^{(\mathrm{S})}$ denote the long- and short-scale scores. We center the scores within each candidate set and constrain only their relative shapes:

$$\mathcal{L}_{\mathrm{logit}} = \frac{1}{N}\sum_u\sum_t\sum_{v\in\Omega_{u,t}}\Big(\big(\hat{y}_{u,t,v}^{(\mathrm{L})}-\bar{\hat{y}}_{u,t}^{(\mathrm{L})}\big)-\big(\hat{y}_{u,t,v}^{(\mathrm{S})}-\bar{\hat{y}}_{u,t}^{(\mathrm{S})}\big)\Big)^2, \tag{9}$$

where

$$\bar{\hat{y}}_{u,t}^{(s)} = \frac{1}{|\Omega_{u,t}|}\sum_{v\in\Omega_{u,t}}\hat{y}_{u,t,v}^{(s)}, \qquad s\in\{\mathrm{L},\mathrm{S}\}.$$

This term encourages consistent relative scoring patterns across scales. Together, $\mathcal{L}_{\mathrm{spec}}$ and $\mathcal{L}_{\mathrm{logit}}$ improve the stability of cross-scale representations and outputs.

5) Training and Inference: The overall objective consists of the main recommendation loss and the cross-scale consistency regularization terms. For the main task, we adopt a binary cross-entropy objective over the ground-truth next item and sampled negative items. Specifically, for each user–time pair (u,t), let $v_{u,t+1}$ denote the ground-truth next item and let $\mathcal{N}_{u,t}$ denote the set of sampled negative items. The supervised loss is defined as

$$\mathcal{L}_{\mathrm{sup}} = -\frac{1}{M}\sum_u\sum_{t=1}^{|S_u|-1}\Big[\log\sigma\big(\hat{y}_{u,t,v_{u,t+1}}\big) + \sum_{v^-\in\mathcal{N}_{u,t}}\log\sigma\big(-\hat{y}_{u,t,v^-}\big)\Big], \tag{10}$$

where M is the number of effective training instances. This objective drives the model to distinguish the ground-truth next item from less relevant candidates under the sequential context, thereby capturing the dynamic evolution of user preferences.

On top of the supervised objective, the model jointly minimizes the cross-scale consistency regularization, which includes the spectral-structure consistency term $\mathcal{L}_{\mathrm{spec}}$ and the prediction-distribution consistency term $\mathcal{L}_{\mathrm{logit}}$. The final training objective is

$$\mathcal{L} = \mathcal{L}_{\mathrm{sup}} + \lambda_{\mathrm{spec}}\mathcal{L}_{\mathrm{spec}} + \lambda_{\mathrm{logit}}\mathcal{L}_{\mathrm{logit}}, \tag{11}$$

where λ_{spec} and λ_{logit} are weighting coefficients that control the strengths of the two regularization terms.

During inference, no additional optimization is required. Both in-domain and cross-domain tasks directly reuse the model trained on the source domain. In-domain inference follows the standard sequential recommendation pipeline. For cross-domain inference, all model parameters are kept fixed, while the frequency-domain features of target-domain items are recomputed from their popularity time series. Based

TABLE I
IN-DOMAIN RECOMMENDATION PERFORMANCE. BOLDFACE INDICATES THE BEST PERFORMANCE, AND UNDERLINING INDICATES THE SECOND-BEST PERFORMANCE.

Dataset	**Office**		**Tool**		**Movie**		**Music**		**Epinions**	
Metric	**N@10**	**R@10**	**N@10**	**R@10**	**N@10**	**R@10**	**N@10**	**R@10**	**N@10**	**R@10**
MostPop	0.272	0.450	0.274	0.459	0.361	0.586	0.327	0.519	0.296	0.438
BPR	0.289	0.457	0.216	0.363	0.477	0.747	0.434	0.646	0.397	0.568
NCF	0.266	0.446	0.239	0.388	0.505	0.784	0.437	0.652	0.396	0.570
LightGCN	0.293	0.465	0.275	0.463	0.512	0.793	0.447	0.665	0.396	0.575
Caser	0.334	0.512	0.297	0.496	0.701	0.891	0.576	0.796	0.475	0.674
SasRec	0.354	0.539	0.337	0.536	0.749	0.918	0.599	0.816	0.501	0.705
BERT4Rec	0.358	0.541	0.350	0.544	0.728	0.900	**0.602**	**0.816**	0.512	0.702
TiSasRec	0.349	0.531	0.341	0.539	0.752	**0.918**	0.523	0.809	0.499	0.701
CL4SRec	**0.358**	**0.550**	**0.352**	0.548	0.725	0.899	0.597	0.813	0.481	0.662
PREPREC	0.335	0.530	0.339	0.544	0.740	0.903	0.570	0.770	0.542	0.732
FREQ	0.343	0.544	0.347	**0.555**	**0.760**	0.914	0.599	0.795	**0.591**	**0.783**

TABLE II
CROSS-DOMAIN RECOMMENDATION PERFORMANCE.

	NDCG@10		**HR@10**		**NDCG@5**		**HR@5**		**NDCG@1**	
Source→Target	**PREPREC**	**FREQ**	**PREPREC**	**FREQ**	**PREPREC**	**FREQ**	**PREPREC**	**FREQ**	**PREPREC**	**FREQ**
Office→Tool	0.309	**0.321**	0.483	**0.509**	0.275	**0.282**	0.378	**0.389**	0.164	**0.167**
Office→Movie	0.645	**0.699**	0.856	**0.883**	0.614	**0.673**	0.761	**0.804**	0.442	**0.518**
Office→Music	0.548	**0.549**	0.752	**0.757**	0.514	**0.514**	0.645	**0.651**	**0.363**	0.358
Office→Epinions	0.511	**0.528**	0.715	**0.728**	0.478	**0.497**	0.613	**0.634**	0.324	**0.340**
Tool→Office	**0.313**	0.309	0.494	**0.495**	**0.276**	0.269	**0.378**	0.372	**0.165**	0.159
Tool→Movie	0.675	**0.720**	0.873	**0.900**	0.648	**0.696**	0.790	**0.825**	0.478	**0.541**
Tool→Music	**0.537**	0.497	**0.734**	0.723	**0.505**	0.454	**0.636**	0.592	**0.356**	0.304
Tool→Epinions	**0.492**	0.472	0.691	**0.693**	**0.459**	0.431	**0.589**	0.566	**0.311**	0.283
Movie→Office	0.310	**0.325**	0.516	**0.524**	0.266	**0.284**	0.381	**0.396**	0.145	**0.164**
Movie→Tool	0.320	**0.328**	0.525	**0.539**	0.277	**0.282**	0.392	**0.397**	0.155	**0.160**
Movie→Music	0.518	**0.538**	**0.762**	0.745	0.481	**0.506**	**0.650**	0.649	0.292	**0.342**
Movie→Epinions	0.514	**0.546**	**0.759**	0.748	0.478	**0.516**	0.648	**0.655**	0.288	**0.355**
Music→Office	0.318	**0.323**	0.508	**0.515**	0.279	**0.284**	0.387	**0.395**	0.162	**0.165**
Music→Tool	0.287	**0.295**	0.465	**0.485**	0.250	**0.254**	0.349	**0.357**	0.144	**0.144**
Music→Movie	0.534	**0.576**	0.826	**0.857**	0.481	**0.528**	0.666	**0.711**	0.284	**0.326**
Music→Epinions	0.520	**0.566**	0.715	**0.763**	0.490	**0.539**	0.662	**0.680**	0.336	**0.373**
Epinions→Office	**0.328**	0.323	**0.524**	0.521	**0.288**	0.283	**0.402**	0.396	**0.166**	0.162
Epinions→Tool	**0.330**	0.322	**0.538**	0.527	**0.286**	0.278	**0.402**	0.392	**0.163**	0.158
Epinions→Movie	0.579	**0.597**	0.862	**0.877**	0.535	**0.558**	0.730	**0.755**	0.321	**0.336**
Epinions→Music	0.574	**0.588**	0.766	**0.780**	0.543	**0.559**	0.671	**0.690**	0.396	**0.406**

on these signals, the model constructs the corresponding residual-enhanced candidate representations $\tilde{\mathbf{e}}_v^t$, which are then matched with the user-side context representation for forward prediction. This procedure does not rely on interaction-based training in the target domain, thereby enabling zero-shot cross-domain transfer.

Overall, FREQ learns complementary representations in the time and frequency domains without modifying the backbone sequential modeling architecture.

IV. EXPERIMENTS

We conduct experiments to answer the following questions:

- **RQ1**: Does FREQ consistently improve performance in both in-domain and cross-domain recommendation?
- **RQ2**: Which side benefits more from residual frequency-domain injection?
- **RQ3**: Which temporal normalization scheme is more suitable for zero-shot cross-domain inference?
- **RQ4**: How do source–target spectral differences relate to the cross-domain gains of FREQ?
- **RQ5**: Are all components of FREQ effective?

A. Experimental setup

Datasets. We evaluate on five real-world datasets: Amazon (Office, Tool) [20], Douban (Movie, Music) [21], and Epinions [22], [23]. Following prior work [9], [10], we adopt leave-one-out splitting for each user: the most recent interaction for testing, the second most recent for validation, and the rest for training.

Baselines. We compare FREQ with four groups of baselines: non-sequential recommenders (MostPop, BPR [24]), collaborative filtering methods (NCF, LightGCN [25], [26]), sequential recommenders (SASRec, BERT4Rec, TiSASRec, CL4SRec, Caser [9]–[11], [13], [27]), and the zero-shot cross-domain baseline PREPREC [6].

Implementation details. All models are implemented in PyTorch under the same data split and evaluation protocol. FREQ is built on the PREPREC backbone and shares the same optimizer, learning rate, and early-stopping strategy; only frequency-related hyperparameters are tuned separately. All cross-domain experiments follow a strict zero-shot setting. Each experiment is repeated five times with different random

TABLE III
CROSS-DOMAIN PERFORMANCE CHANGES WITH DIFFERENT FREQUENCY-DOMAIN RESIDUAL INJECTION STRATEGIES. NOTE: Δ REPRESENTS THE RELATIVE CHANGE (IN %) COMPARED TO THE REFERENCE SETTING, WHERE 0 INDICATES THE REFERENCE SCHEME ITSELF.

Source → Target	History-side Injection	Candidate-side Injection	Both-side Injection
Office→Movie	$\Delta\text{NDCG} = -0.8$ $\Delta\text{HR} = +0.3$	$\Delta\text{NDCG} = +0.8$ $\Delta\text{HR} = +0.7$	0
Office→Epinions	$\Delta\text{HR} = +0.1$	$\Delta\text{NDCG} = +0.6$ $\Delta\text{HR} = +0.2$	0
Movie→Epinions	0	0	$\Delta\text{NDCG} = -0.2$ $\Delta\text{HR} = -0.6$

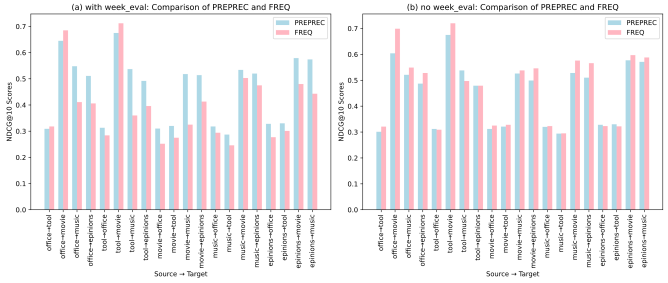

Fig. 2. Cross-Domain evaluation results under different time statistics strategies.

seeds, and average results are reported.

Evaluation metrics. We use NDCG@k and HR@k with $k \in \{10, 5, 1\}$. For each test instance, the candidate set contains the ground-truth item and 100 randomly sampled unobserved items.

B. Overall Performance (RQ1)

As shown in Table I, FREQ consistently improves over PREPREC in the in-domain setting, especially on Movie and Epinions, while remaining competitive with strong sequential baselines on Office, Tool, and Music. These results show that frequency-domain enhancement does not harm backbone performance and can further improve ranking quality in in-domain recommendation.

Table II shows that FREQ also improves over PREPREC in most cross-domain tasks. Smaller gains or slight drops mainly occur when Tool is the source or target domain, likely because Tool is the sparsest domain and exhibits weaker transferable temporal rhythms. Overall, the results indicate that the frequency-domain residual signals learned by FREQ are generally effective and transferable across domains.

C. Residual Injection Strategy (RQ2)

To compare different injection locations, we evaluate three heterogeneous source–target domain pairs, as shown in Table III. FREQ performs best and most stably when the residual signal is injected only on the candidate side. This is likely because item-level popularity sequences are smoother and exhibit clearer periodic patterns, whereas user histories are shorter, sparser, and more variable. As a result, injecting spectral residuals into user representations may disturb sequence modeling, while candidate-side injection more directly improves matching between user preference and item temporal rhythm.

D. Time Statistics Strategy (RQ3)

We compare FREQ and PREPREC under two target-side time-statistics settings (Fig. 2): *with week_eval*, which re-estimates weekly popularity and time embeddings in the target domain following PREPREC, and *no week_eval*, which directly uses raw target-domain counts with source-domain time representations.

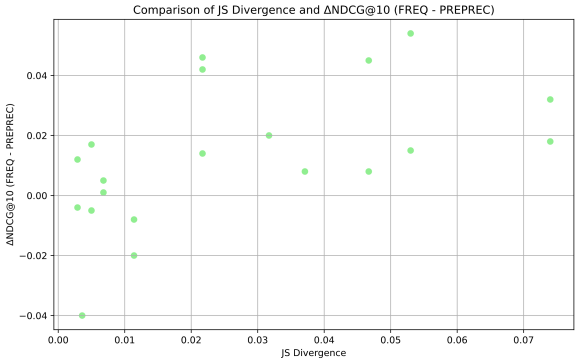

Fig. 3. Relationship between inter-domain spectral discrepancy and the relative gains of FREQ.

The results show that FREQ performs better under *no week_eval*, whereas PREPREC benefits more from *with week_eval*. This difference is likely due to their distinct dependencies on temporal statistics: PREPREC relies mainly on current-time popularity statistics, while FREQ models the full temporal trajectory, for which additional weekly normalization may distort the target-domain spectral structure. Therefore, in RQ1 we adopt *with week_eval* for PREPREC and *no week_eval* for FREQ.

E. Inter-Domain Spectral Discrepancy and Cross-Domain Gains (RQ4)

To examine when FREQ is most beneficial, we compute the Jensen–Shannon divergence (JS) between the normalized popularity spectral energy distributions of the source and target domains over a recent time window (Fig. 3). A larger JS indicates greater discrepancy in periodic structure.

We observe a clear trend: larger inter-domain JS generally corresponds to larger ΔNDCG. This suggests that when source and target domains differ more in temporal rhythm, PREPREC becomes less transferable, whereas FREQ is more robust because its frequency-domain residuals provide additional alignment of periodic structures across domains.

F. Ablation Study (RQ5)

Table IV shows that removing any non-residual core component degrades performance to different extents. Removing

TABLE IV
THE IMPACT OF DIFFERENT COMPONENTS ON MODEL PERFORMANCE (USING OFFICE → MOVIE AS AN EXAMPLE, COMPARED WITH THE FULL FREQ MODEL)

Variant Configuration	ΔNDCG@10	ΔHR@10
w/o Long Window	-0.023	-0.009
w/o Short Window	+0.004	-0.008
w/o Gating	-0.001	+0.004
w/o Consistency Regularization	-0.012	-0.003

the long window causes the largest drop, indicating that low-frequency trend information is critical for cross-domain transfer. Removing the short window has a smaller effect, suggesting that it mainly complements the model by capturing high-frequency local variations. Removing the gating module leads to only minor changes, implying that the candidate-side residual signal itself provides most of the gain, while the gate mainly improves robustness. Removing the consistency regularization consistently hurts performance, confirming its importance in stabilizing multi-scale spectral representations.

Overall, FREQ benefits from the combination of multi-scale spectral modeling and consistency regularization, which together improve the stability and transferability of frequency-domain representations.

V. CONCLUSION

This paper proposes FREQ, a frequency-domain enhancement module for zero-shot cross-domain sequential recommendation. By combining multi-scale temporal windows, spectral modeling, residual frequency-domain injection, lightweight gating, and cross-scale consistency regularization, FREQ improves the transferability of time-driven popularity representations. Experimental results show that FREQ typically outperforms strong baselines, with larger gains when the source and target domains exhibit greater periodic discrepancy.

Future work will explore more adaptive temporal windows and finer-grained frequency decomposition for irregular and highly heterogeneous cross-domain scenarios.

ACKNOWLEDGMENT

This work was generously supported by the Project of the Humanities and Social Sciences Research of the Ministry of Education under grant number 25YJAZH251. We also would like to thank the anonymous reviewers for their valuable feedback.

REFERENCES

[1] Y. Liu, S. Wang, X. Li, and F. Sun, "A meta-adversarial framework for cross-domain cold-start recommendation," *Data Sci. Eng.*, vol. 9, no. 2, pp. 238–249, 2024.

[2] C. Zhao, C. Li, and C. Fu, "Cross-domain recommendation via preference propagation GraphNet," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manag. (CIKM)*, pp. 2165–2168, 2019.

[3] C. Park, T. Kim, T. Choi, J. Hong, Y. Yu, M. Cho, *et al.*, "Cracking the code of negative transfer: A cooperative game theoretic approach for cross-domain sequential recommendation," in *Proc. 32nd ACM Int. Conf. Inf. Knowl. Manag. (CIKM)*, pp. 2024–2033, 2023.

[4] H. Su, J. Li, Z. Du, L. Zhu, K. Lu, and H. T. Shen, "Cross-domain recommendation via dual adversarial adaptation," *ACM Trans. Inf. Syst.*, vol. 42, no. 3, pp. 83:1–83:26, 2024.

[5] C. Tian, Y. Xie, X. Chen, Y. Li, and X. Zhao, "Privacy-preserving cross-domain recommendation with federated graph learning," *ACM Trans. Inf. Syst.*, vol. 42, no. 5, pp. 135:1–135:29, 2024.

[6] J. Wang, P. Rathi, and H. Sundaram, "A pre-trained sequential recommendation framework: Popularity dynamics for zero-shot transfer," in *Proc. 18th ACM Conf. Recommender Syst. (RecSys)*, pp. 433–443, 2024.

[7] M. L. P. Toledo and M. N. De Rezende, "Comparison of LSTM, GRU and hybrid architectures for usage of deep learning on recommendation systems," in *Proc. 4th Int. Conf. Adv. Artif. Intell.*, pp. 16–22, 2020.

[8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems* 30 (NeurIPS), 2017.

[9] W.-C. Kang and J. McAuley, "Self-attentive sequential recommendation," in *Proc. 2018 IEEE Int. Conf. Data Min. (ICDM)*, pp. 197–206, 2018.

[10] F. Sun, J. Liu, J. Wu, C. Pei, X. Lin, W. Ou, and P. Jiang, "BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer," in *Proc. 28th ACM Int. Conf. Inf. Knowl. Manag. (CIKM)*, pp. 1441–1450, 2019.

[11] J. Li, Y. Wang, and J. McAuley, "Time interval aware self-attention for sequential recommendation," in *Proc. 13th Int. Conf. Web Search Data Min. (WSDM)*, pp. 322–330, 2020.

[12] H. Du, H. Yuan, P. Zhao, D. Wang, V. S. Sheng, Y. Liu, *et al.*, "Feature-aware contrastive learning with bidirectional transformers for sequential recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 36, no. 12, pp. 8192–8205, 2024.

[13] X. Xie, F. Sun, Z. Liu, S. Wu, J. Gao, B. Ding, *et al.*, "Contrastive learning for sequential recommendation," arXiv:2010.14395, 2021.

[14] T. Zang, Y. Zhu, H. Liu, R. Zhang, and J. Yu, "A survey on cross-domain recommendation: Taxonomies, methods, and future directions," *ACM Trans. Inf. Syst.*, vol. 41, no. 2, pp. 1–39, 2022.

[15] L. Li, Q. Do, and W. Liu, "Cross-domain recommendation via coupled factorization machines," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, pp. 9965–9966, 2019.

[16] A. Samra, E. Frolov, A. Vasilev, A. Grigorevskiy, and A. Vakhrushev, "Cross-domain latent factors sharing via implicit matrix factorization," in *Proc. 18th ACM Conf. Recommender Syst. (RecSys)*, pp. 309–317, 2024.

[17] A. Ahmed, K. Saleem, O. Khalid, and U. Rashid, "On deep neural network for trust aware cross domain recommendations in e-commerce," *Expert Systems with Applications*, vol. 174, Art. no. 114757, 2021.

[18] D. Zhou, X. Cai, and W. Pan, "Contrastive text-enhanced transformer for cross-domain sequential recommendation," in *Proc. 31st ACM SIGKDD Conf. Knowl. Discov. Data Min. (KDD)*, pp. 4110–4119, 2025.

[19] Z. Li, H. Jiang, and S. Zhao, "MCD-EARS: A multimodal cross-domain expertise-aware recommender system for healthcare applications," *Information Sciences*, vol. 728, Art. no. 122821, 2026.

[20] J. Ni, J. Li, and J. McAuley, "Justifying recommendations using distantly-labeled reviews and fine-grained aspects," in *Proc. 2019 Conf. Empir. Methods Nat. Lang. Process. and 9th Int. Joint Conf. Nat. Lang. Process. (EMNLP-IJCNLP)*, pp. 188–197, 2019.

[21] W. Song, Z. Xiao, Y. Wang, L. Charlin, M. Zhang, and J. Tang, "Session-based social recommendation via dynamic graph attention networks," in *Proc. 12th ACM Int. Conf. Web Search Data Min. (WSDM)*, pp. 555–563, 2019.

[22] J. Tang, H. Gao, H. Liu, and A. Das Sarma, "eTrust: Understanding trust evolution in an online world," in *Proc. 18th ACM SIGKDD Conf. Knowl. Discov. Data Min. (KDD)*, pp. 253–261, 2012.

[23] J. Tang, H. Gao, and H. Liu, "mTrust: Discerning multi-faceted trust in a connected world," in *Proc. 5th ACM Int. Conf. Web Search Data Min. (WSDM)*, pp. 93–102, 2012.

[24] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: Bayesian personalized ranking from implicit feedback," arXiv:1205.2618, 2012.

[25] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," arXiv:1708.05031, 2017.

[26] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "LightGCN: Simplifying and powering graph convolution network for recommendation," arXiv:2002.02126, 2020.

[27] J. Tang and K. Wang, "Personalized top-N sequential recommendation via convolutional sequence embedding," arXiv:1809.07426, 2018.