SeeDRec: Sememe-based Diffusion for Sequential Recommendation

Haokai Ma [1], **Ruobing Xie** [3*], **Lei Meng** [1,2*], **Yimeng Yang** [1], **Xingwu Sun** [3], **Zhanhui Kang** [3]
[1] School of Software, Shandong University, China
[2] Shandong Research Institute of Industrial Technology, China
[3] Tencent, China
mahaokai@mail.sdu.edu.cn, ruobingxie@tencent.com, lmeng@sdu.edu.cn,
y_yimeng@mail.sdu.edu.cn, sunxingwu01@gmail.com, kegokang@tencent.com

Abstract

Inspired by the power of Diffusion Models (DM) verified in various fields, some pioneering works have started to explore DM in recommendation. However, these prevailing endeavors commonly implement diffusion on item indices, leading to the increasing time complexity, the lack of transferability, and the inability to fully harness item semantic information. To tackle these challenges, we propose SeeDRec, a sememe-based diffusion framework for sequential recommendation (SR). Specifically, inspired by the notion of sememe in NLP, SeeDRec first defines a similar concept of *recommendation sememe* to represent the minimal interest unit and upgrades the specific diffusion objective from the item level to the sememe level. With the Sememe-to-Interest Diffusion Model (*S2IDM*), SeeDRec can accurately capture the user's diffused interest distribution learned from both local interest evolution and global interest generalization while maintaining low computational costs. Subsequently, an Interest-aware Prompt-enhanced (*IPE*) strategy is proposed to better guide each user's sequential behavior modeling via the learned user interest distribution. Extensive experiments on nine SR datasets and four cross-domain SR datasets verify its effectiveness and universality. The code is available in https://github.com/hulkima/SeeDRec.

1 Introduction

Recommender system (RS) has become essential for many real-world applications owing to its ability to accurately mine users' personalized interests [Wang *et al.*, 2015; Fan *et al.*, 2023; Meng *et al.*, 2020; Ma *et al.*, 2021; Ma *et al.*, 2023a]. Recently, numerous advanced techniques in NLP have been assimilated into RS, continuing to demonstrate their impact [Kang and McAuley, 2018; Sun *et al.*, 2019; Ma *et al.*, 2023b]. The sequential modeling process of user behaviors in sequential recommendation (SR) bears semblance to the language modeling task prevalent in NLP, that is, SR seeks to recommend the next-item that user may be interested in by modeling the sequential dependencies of user's temporal behaviors [de Souza Pereira Moreira *et al.*, 2021].

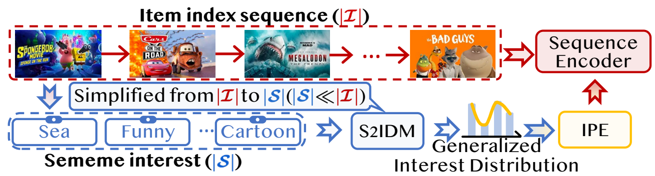

Figure 1: Illustration of the proposed SeeDRec. Our sememe-based diffusion enhances the scalability, transferability, and usage of item semantics, while keeping good performance and universality.

As motivated by the outstanding distribution generative performance of the Diffusion Model (DM) in various domains (image synthesis [Ho *et al.*, 2020], audio processing [Kong *et al.*, 2021] and semantic segmentation [Brempong *et al.*, 2022]), some pioneering studies attempt to explore the effectiveness of DM in recommendation. DiffRec [Wang *et al.*, 2023e] conducts DM on item indices to infer users' interaction probabilities in a denoising manner, CDDRec [Wang *et al.*, 2023f], DiffuRec [Li *et al.*, 2023], DreamRec [Yang *et al.*, 2023] and DDRM [Zhao *et al.*, 2024] consider injecting uncertainty into item embeddings via DM under the generative paradigm. PDRec [Ma *et al.*, 2024a] enhances DiffRec by leveraging the pre-trained DM on item indices as plugins to improve SR, which exhibits universality and analogous inference costs to the sequence encoders. However, the ID-based methods typically exhibit: (1) a linear increased complexity with the growing number of corpus, leading to the worse scalability in large-scale RS, (2) an excessive dependency on item indices, lacking the cross-domain transferability, (3) a limitation in fully utilizing the semantic correlations within items, thereby resulting in sub-optimal performance.

To tackle these challenges, an intuitive idea is to conduct DM on other objectives that are atomic, multi-domain transferable, and encapsulate the semantic correlations requisite for SR, rather than item indices. We notice the definition of sememe, which is regarded as the minimum semantic unit in NLP [Niu *et al.*, 2017]. NLP researchers believe that each word can be decomposed into a limited set of manually-defined and language-independent sememes [Qi *et al.*, 2022]. Inspired by this, we creatively propose the concept of *recommendation sememe* (abbreviated as *sememe*) to represent users' minimal interest unit in recommendation. As shown in Fig. 1, all items can be represented as the combination of several sememes, thereby reducing the quantity of DM objectives, accomplishing the transferability across diverse domains and incorporating semantic correlations among items in SR. To do this, we present a model-agnostic **Se**meme-bas**e**d **D**iffusion framework for Sequential **Rec**ommendation (**SeeDRec**). It regards each user's original interest on sememes observed by the system as a "***Seed***", nourishes it with both local and global user preferences, and cultivates the "***Blooming flower***" of generalized interest distribution via our sememe-based diffusion. Specifically, we first propose the Sememe-to-Interest Diffusion Model (*S2IDM*) to upgrade the existing DM-based recommenders from item index to sememe, mining each user's generalized interest distribution by synergistically considering temporal, frequency, and co-occurrence information interlinking sememes. The user generalized interest distribution learned by S2IDM contains both the user's local personalized behavioral preference and the global sememe correlations implied by all user behaviors. Furthermore, to incorporate the generalized interest distribution into the sequential modeling, we design an Interest-aware Prompt-enhanced (*IPE*) strategy to guide the sequential modeling towards the direction for better personalized behavior understanding. With the sememes powered by *S2IDM* and *IPE*, SeeDRec could achieve improvement on various base models as a plugin with better scalability and interest generalization. Moreover, the knowledge of user interest diffusion encoded in *S2IDM* can be transferred to other domains via our proposed recommendation sememe anchors.

We conduct extensive experiments on nine SR datasets and four cross-domain SR (CDSR) datasets to demonstrate the superiority of SeeDRec. For simplification and universality, we also use existing taxonomies and words as sememes to simulate practical scenarios. We further conduct various ablation studies, universality analyses, few-shot analyses, and model analyses to validate the effectiveness of the proposed *S2IDM* and *IPE*. The contributions of this work are as follows:

- We have verified the feasibility of conducting diffusion on the *recommendation sememe* and incorporating it in SR. To the best of our knowledge, we are the first to conduct DM on sememes to enhance SR and CDSR tasks.
- We propose *S2IDM* to mine diffused user generalized interest distribution by simultaneously considering global sememe correlations and users' local sequential behaviors.
- We creatively present the *IPE* strategy to enable the precise interest transfer from the discrete sememe interest distribution to the continuous representation to improve the personalized sequential modeling via a prompt learning paradigm.
- We conduct extensive evaluations on 13 real-world datasets to verify that the proposed SeeDRec is effective, universal, and easy-to-deploy. We also design comprehensive analyses to demonstrate the effective mechanism of SeeDRec and validate it in the more challenging few-shot scenarios.

2 Related Work

Sequential Recommendation Benefiting from the advancement of sequential modeling techniques, SR is proposed to format and encode user temporal behaviors to infer their dynamic interests [Sun *et al.*, 2022; Wang *et al.*, 2023g; Wang *et al.*, 2023d; Sun *et al.*, 2024]. Within its evolutionary trajectory, early SR methods reason users' short-term preference through the Markov Chains [Rendle *et al.*, 2010]. GRU4Rec [Hidasi *et al.*, 2016] leverages the Gate Recurrent Unit as the sequential encoder to capture users' long-term dependencies. Caser [Tang and Wang, 2018] imports the Convolutional Neural Network (CNN) to extract sequential patterns from user behaviors at different time intervals. Inspired by the success of self-attention mechanism, SASRec [Kang and McAuley, 2018] incorporates it to jointly model users' short- and long-term preferences. CL4SRec [Xie *et al.*, 2022] and MStein [Fan *et al.*, 2023] additionally utilize the mutual information on self-supervised signals to improve the sequential representations. Unlike the complex network structures and stochastic augmentations in previous SR works, SeeDRec smartly leverages the prevailing diffusion models to incorporate the generalized user interests in sequential pattern modeling. Furthermore, SeeDRec can be effortlessly deployed on diverse SR models with advanced techniques to bring significant and consistent performance improvements (see Sec. 4.5).

Diffusion Model for Recommendation Inspired by the remarkable performance in image generation [Ho *et al.*, 2020; Nichol and Dhariwal, 2021; Wang *et al.*, 2023b],scene text editing [Wang *et al.*, 2023a], and machine translation [Chen *et al.*, 2023], some researchers attempt to incorporate DM in recommendation. CODIGEM [Walker *et al.*, 2022] is the pioneer DM-based recommender which generates improved collaborative signals via the Denoising Diffusion Probabilistic Model (DDPM). DiffRec [Wang *et al.*, 2023e] reduces the scheduled noises into the reverse process to infer user interaction probabilities in a denoising manner. PDRec [Ma *et al.*, 2024a] is the state-of-the-art DM-based recommender, which takes the DM trained on all corpus as the plugin in SR to fully leverage the diffusion-based user preferences. Moreover, a series of related DM-based works (e.g., DiffuRec [Li *et al.*, 2023], DiffRec* [Du *et al.*, 2023] and DreamRec [Yang *et al.*, 2023]) employ DM on the continuous item embedding space with additional transitions and noising strategies, which shares the same modeling pipeline with classical SR methods. Disparate objectives make them inherently non-comparable with this work, and they can serve as the base SR model within our SeeDRec. In particular, we have employed PDRec as the backbone to substantiate the efficacy of SeeDRec over DM-based recommenders.

In conclusion, existing DM-based methods typically conduct DM on the discrete item indices or the continuous embedding of each item to achieve the uncertainty injection in recommendation. Therefore, the augmentation of the number of items exacerbates the multiplication of their time and space complexities, rendering them arduous to deploy in real-world million-level recommendation systems. Moreover, these DM-based recommenders solely rely on item indices, demonstrating the lack of scalability and the inability to fully leverage the semantic correlations within items. Every instance of encountering a new dataset compels them to be trained from scratch without leveraging the transferability across multiple domains, leading to the recurrent process that demands substantial computility and time consumption.

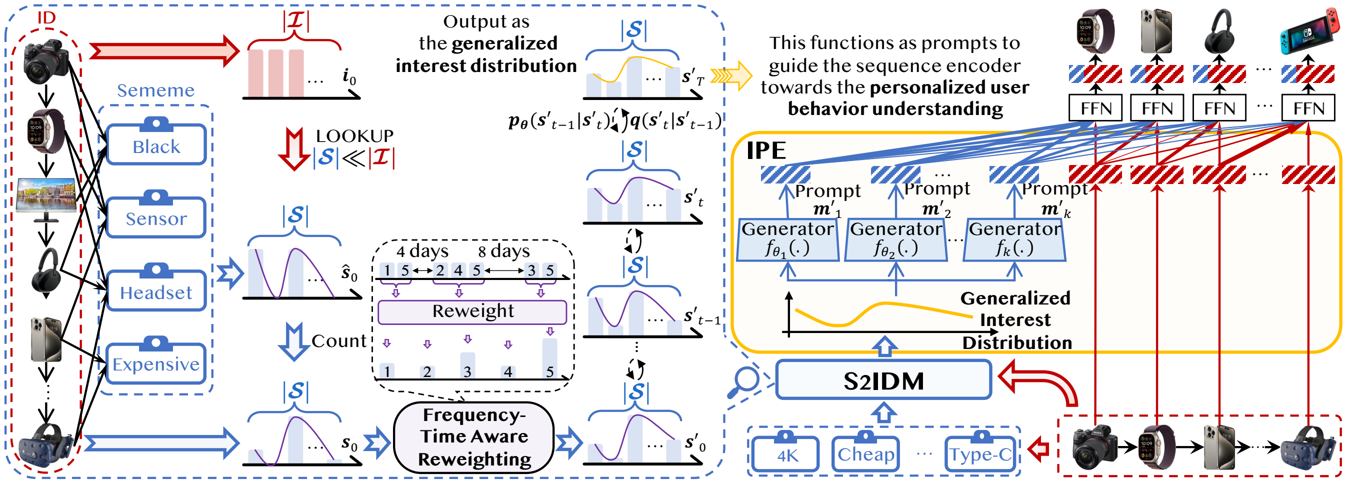

Figure 2: The overall structure of the proposed SeeDRec (based on SASRec). S2IDM provides the generalized user interest distribution based on sememes, and IPE adopts the generalized user interests to better guide sequential modeling via prompt learning.

3 Methods

3.1 Problem Statement

In this article, we focus on exploring the DM-enhanced recommendation task in multiple recommendation scenarios, with an illustration of the example of SR. We define the *behavior sequence* $S_u^{\mathcal{I}} = \{i_1, i_2, \cdots, i_p\}$ of user $u \in \mathcal{U}$, where $i_j \in \mathcal{I}$ is the j-th interacted item of user u and p denote the historical behavior length. Given $S_u^{\mathcal{I}}$, we adopt SASRec [Kang and McAuley, 2018] as the sequential encoder to predict the target item i_{p+1} that will be preferred by user u.

3.2 Overall Structure

In this section, we elaborate on the proposed model-agnostic Sememe-based Diffusion for sequential recommendation, which utilizes the DM on sememes to explore each user's atomized multi-interest units for the multiple recommendation scenarios. As illustrated in Fig. 2, the proposed SeeDRec consists of two main components, including the Sememe-to-Interest Diffusion Model (*S2IDM*) and the Interest-aware Prompt-enhanced (*IPE*) strategy. Specifically, SeeDRec first proposes *S2IDM*, which creatively conducts diffusion on the sememe level rather than the item level used in conventional works, with the aim to model users' local minimal interest units from the global interest diffusion aspect. To adeptly guide the attention of sequential modeling towards the user's personalized interests, SeeDRec designs an *IPE* strategy to convert the diffused generalized interest distribution into informative prompts via multiple personalized prompt generators. It enables the accurate extraction of the interest-centred information from each user's historical behavior sequences, thereby achieving precise user dynamic interest modeling. Furthermore, SeeDRec continues to manifest its superiority in SR and CDSR tasks where the explicit item-sememe hierarchical taxonomy information is even absent (as a more challenging setting with increased noise). Detailed experimental results and analyses can be found in Sec. 4.2 and Sec. 4.4.

3.3 Sememe-to-Interest Diffusion Model

In this section, we describe our Sememe-to-Interest Diffusion Model based on DDPM [Ho *et al.*, 2020], named *S2IDM*. The overall structure of *S2IDM* is illustrated in the left part of Fig. 2. In contrast to the direct diffusion employed by DiffRec [Wang *et al.*, 2023e] and PDRec [Ma *et al.*, 2024a] at the ID indices, *S2IDM* creatively deploys DM at the ***sememe*** level, concomitantly refining it in accordance with the intercorrelation of sememes, the sememe overlap of item and the multiple interest drift of behavioral sequences.

Forward Process

In the forward process, *S2IDM* corrupts the original sememe distribution of each user by injecting the Gaussian noises step by step in a discrete manner. To be specific, given the behavioral sequence $S_u^{\mathcal{I}}$ in the index level, *S2IDM* first extracts the prior sememe behaviors, re-weights it via the frequency and time interval of each sememe to obtain the original interest distribution $\boldsymbol{s}'_0$ and finally corrupts it in a discrete manner.

Diffusion on sememe With the behavioral sequence $S_u^{\mathcal{I}} = \{i_1, i_2, \cdots, i_p\}$ of user u, we first LOOKUP the prior sememe behaviors $\boldsymbol{s}_u = \{s_1^{i_1}, s_2^{i_1}, \cdots, s_k^{i_p}\}$ of user u as the initial input $\hat{\boldsymbol{s}}_0$ from the item-sememe dependency matrix $\mathcal{D}^{\mathcal{S}} \in \mathbb{R}^{|\mathcal{I}| \times q}$, where $s_k^{i_p} \in \mathcal{S}$ denotes the k-th sememe of item i_p and $q = |\mathcal{S}|$ denotes the total length of the sememe set $\mathcal{S}$. This enables the transfer of DM from the item domain to that of sememe, yielding a threefold advantage: (1) It entails a reduction in the space and time complexity of DM as stipulated by $|\mathcal{S}| \ll |\mathcal{I}|$, enabling the practical **scalability** of our sememe-based diffusion model. (2) By virtue of its foundational definition as the minimal unit of interest, the inherent multi-domain sharing characteristic of sememe endows it with **transferability** in CDSR tasks. (3) It brings in the **semantic interrelationship** inherent in sememes to furnish comprehensive support to DM.

Frequency-time aware reweighting In light of the pervasive one-to-many association between items and sememes, highly correlated sememes tend to iteratively manifest within user behavior sequences. The unadorned utilization of the reweight strategies in TI-DiffRec is untenable, obscuring users' authentic interests. To this end, we first count the frequency of each sememe in user behaviors to obtain the $\boldsymbol{s}_0$, then we conduct the time-interval reweighting strategy in TI-DiffRec to generate the time-interval weights $w_1^{i_k} = w_2^{i_k} = \cdots = w_j^{i_k} = w_{\min} + \frac{t_{i_k} - t_1}{t_p - t_1}(w_{\max} - w_{\min})$ for all sememes $s_1^{i_k}, s_2^{i_k}, \cdots, s_j^{i_k}$ of item i_k with the item-level timestamp sequence $T_u^{\mathcal{I}} = \{t_1, t_2, \cdots, t_p\}$. Here $s_j^{i_k}$ is the j-th sememe of item i_k in $\boldsymbol{s}_0$. Ultimately, we assign these weights to each sememe, amalgamating the weight of identical sememes to derive the original interest distribution $\boldsymbol{s}_0'$.

Discrete noising Discrete noising is one of the core phases in DM that injects uncertainty into the original interest distribution. Different from the existing DM-based recommenders, we gradually corrupt the original interest distribution $\boldsymbol{s}_0'$ on sememe via a forward transition $q(\boldsymbol{s}_t' \mid \boldsymbol{s}_{t-1}') = \mathcal{N}\left(\boldsymbol{s}_t'; \sqrt{1-\beta_t}\boldsymbol{s}_{t-1}', \beta_t \mathbf{I}\right)$, where $\beta_t \in (0,1)$ denotes the scale of the added Gaussian noise at the step t, which is generated by the linear noise schedule [Wang *et al.*, 2023e; Li *et al.*, 2023]. With the reparameterization trick [Kingma and Welling, 2013] and the inherent additivity of the independent Gaussian distribution, we can formulate that $\boldsymbol{s}_t' = \sqrt{\bar{\alpha}_t}\boldsymbol{s}_0' + \sqrt{1-\bar{\alpha}_t}\epsilon$, where $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is the added Gaussian noise, $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{t'=1}^{t} \alpha_{t'}$. We also observe that as $t \to +\infty$, $\boldsymbol{s}_t'$ undergoes a gradual convergence towards the standard Gaussian distribution.

Reverse Process

The task of the reverse process in DM is to remove the added Gaussian noise step by step to recover the interest distribution $\boldsymbol{s}_0'$ from the perturbed sememe distribution $\boldsymbol{s}_t'$. The related reverse transition $p_\theta(\boldsymbol{s}_{t-1}' \mid \boldsymbol{s}_t')$ is defined as: $p_\theta(\boldsymbol{s}_{t-1}' \mid \boldsymbol{s}_t') = \mathcal{N}\left(\boldsymbol{s}_{t-1}'; \boldsymbol{\mu}_\theta(\boldsymbol{s}_t', t); \boldsymbol{\Sigma}_\theta(\boldsymbol{s}_t', t)\right)$, where the mean $\boldsymbol{\mu}_\theta(\boldsymbol{s}_t', t)$ and variance $\boldsymbol{\Sigma}_\theta(\boldsymbol{s}_t', t)$ can be modeled by a deep neural network due to the fact that the relatively small Gaussian noise ensures the transition kernel $p_\theta(\boldsymbol{s}_{t-1}' \mid \boldsymbol{s}_t')$ follows a Gaussian distribution. The iterative process can be formulated as follows:

$$\boldsymbol{s}_t' \xrightarrow{p_\theta(\boldsymbol{s}_{t-1}' \mid \boldsymbol{s}_t')} \boldsymbol{s}_{t-1}' \xrightarrow{p_\theta(\boldsymbol{s}_{t-2}' \mid \boldsymbol{s}_{t-1}')} \cdots \xrightarrow{p_\theta(\boldsymbol{s}_0' \mid \boldsymbol{s}_1')} \boldsymbol{s}_0'. \quad (1)$$

Optimization Strategy

The training objective of *S2IDM* is to force the reverse transition $p_\theta(\boldsymbol{s}_{t-1}' \mid \boldsymbol{s}_t')$ to closely approximate the posterior distribution $q(\boldsymbol{s}_{t-1}' \mid \boldsymbol{s}_t', \boldsymbol{s}_0')$, which is achieved by minimizing the Kullback-Leibler (KL) divergence. With the simplification [Ho *et al.*, 2020] and importance sampling technique [Nichol and Dhariwal, 2021], the above D_{KL} can be rewritten as the weighted Mean Square Error (MSE) loss $\mathcal{L}_{\mathrm{D}}$ to focus on more difficult denoising tasks and alleviate the unnecessary noise:

$$\begin{aligned}
\mathcal{L}_{\mathrm{D}} &= D_{\mathrm{KL}}\left(q\left(\boldsymbol{s}_{t-1}' \mid \boldsymbol{s}_t', \boldsymbol{s}_0'\right) \| p_\theta\left(\boldsymbol{s}_{t-1}' \mid \boldsymbol{s}_t'\right)\right) \\
&= \frac{1}{2\sigma_t^2}\left\|\tilde{\mu}\left(\boldsymbol{s}_t', \boldsymbol{s}_0'\right) - \mu_\theta\left(\boldsymbol{s}_t', t\right)\right\|^2 \\
&= \cdots \\
&= \frac{g(t)}{2}\left[E_{\boldsymbol{s}_0', \boldsymbol{s}_t'}\left\|\boldsymbol{s}_0' - x_\theta\left(\boldsymbol{s}_t', t\right)\right\|^2\right]
\end{aligned} \quad (2)$$

where $g(t)$ denotes the discrepancy between the Signal-to-Noise Ratio at step t and $t-1$ with the sampling probability, and $x_\theta(.)$ is a deep neural network [Wang *et al.*, 2023e].

Diffusion Inference

Intuitively speaking, the inherent occasional noise pervades the collected user behaviors, and the user generalized interest mining at the sememe level will also inevitably introduce a certain degree of bias. Consequently, we refrain from introducing additional noise and instead treat the original sememe distribution $\boldsymbol{s}_0'$ as the inherently noise-enriched $\bar{\boldsymbol{s}}_t'$, proceeding directly with the reverse process on it as $\bar{\boldsymbol{s}}_t' \to \bar{\boldsymbol{s}}_{t-1}' \to \cdots \to \bar{\boldsymbol{s}}_0'$. This not only retains more personalized information to improve the precision of DM but also ensures that the inference process is not initiated from a fully disordered state, leading to a more robust inference process that precisely aligns with the intrinsic nature of recommendation tasks.

3.4 Interest-aware Prompt-enhanced Strategy

It is intuitive that there exists an insurmountable variation between the diffusion-based user interest and the temporal preference, which motivates us to explore how to incorporate the generalized interest distribution into the sequential modeling. In parallel, the interest distribution obtained from *S2IDM* (a) exhibits correlation with the generalized user interests, so as to provide some attention clues that are somewhat relevant to SR tasks, and (b) enables the multi-domain inter-connectivity at the sememe level, thereby leveraging its semantic relationships to uncover nuanced user preferences and yield enhanced information gain. To this end, we follow the prompt learning paradigm [Lester *et al.*, 2021; Wu *et al.*, 2024] to propose the Interest-aware Prompt-enhanced (*IPE*) strategy to convert the generalized interest distribution on discrete sememes into multiple continuous prompts. This can smooth the bias between the semantic and behavioral intentions and guide the sequential modeling towards understanding personalized behavioral patterns to enrich the user representation.

Specifically, given the generalized interest distribution $\bar{\boldsymbol{s}}_0'$, we first project it into the same feature space via multiple prompt generators $f_\theta(.)$, which is built with a two-layer fully-connected network with the Tanh activation. After obtaining the multi-interest prompts $\boldsymbol{M} = \{\boldsymbol{m}_1, \boldsymbol{m}_2, \cdots, \boldsymbol{m}_k\}$ where $\boldsymbol{m}_i = f_{\theta_i}(\bar{\boldsymbol{s}}_0')$, we place these prompt-enhanced knowledge before the original input matrix $\mathcal{D}^{\mathcal{I}}$ to inject the diffused user generalized interests into the self-attention functions in pre-trained sequential models. Here we designate the length k of the multi-interest prompts as 3 for streamlining. Beyond the rectification of diffusion objective within *S2IDM*, the lower computational complexity of SeeDRec is also evident in various facets. First is that the generalized interest distribution $\bar{\boldsymbol{s}}_0'$ used in *IPE* only needs to be generated once via the inference process of *S2IDM* before model training, yielding the asymptotic similar time complexity to its sequence encoder in online serving. Secondly, the implementation of *IPE* only requires a limited number of additional prompt generators to be trained, which also mitigates the space complexity of the existing DM-based recommenders. Note that the proposed *IPE* is an universal strategy, devoid of the tailored design in sequence modeling in SR. Thus, *IPE* can harness the possible advancement in future sequence modeling, thereby extending the lifecycle of SeeDRec (see Sec. 4.2 and Sec. 4.5).

Optimization Objective Following [Kang and McAuley, 2018], we opt the Binary Cross-Entropy Loss $\mathcal{L}$ of each user-item pair (u, i) in training set $\mathcal{R}$ as the optimization objective:

$$\mathcal{L} = -\sum_{(u,i)\in\mathcal{R}} \left[y_{u,i} \log \hat{y}_{u,i} + (1 - y_{u,i}) \log (1 - \hat{y}_{u,i}) \right]. \quad (3)$$

where $\hat{y}_{u,i} = \boldsymbol{u}^\top \boldsymbol{e}_i$ is the predicted probility between user representation $\boldsymbol{u}$ and item embedding $\boldsymbol{e}_i$, $y_{u,i} = 1$ and $y_{u,i} = 0$ denote the positive and negative samples respectively.

3.5 In-depth Model Discussion

The most related work of SeeDRec in existing DM-based recommenders is PDRec [Ma *et al.*, 2024a], which fully leverages the diffused preference on item indices to improve SR. Since the DM within PDRec is trained on item indices, its time complexity will be limited by the number of item corpus, thereby encountering the scalability challenge. Secondly, due to the ID-based diffusion, PDRec relies on overlapped users to re-train the DM from scratch in CDSR tasks. Finally, the ID-based DM proves challenging in fully harnessing the associative information within the item semantic hierarchy, potentially resulting in sub-optimal performance.

In contrast, SeeDRec proposes *S2IDM* on the discrete sememes to mine each user's generalized interest distribution through minimal interest units modeling. Hence, the time complexity on the DM side can be simplified from $|\mathcal{I}| \cdot \mathcal{O}(\mathrm{D})$ to $|\mathcal{S}| \cdot \mathcal{O}(\mathrm{D})$, where $\mathcal{O}(\mathrm{D})$ denotes the time complexity of all other facets within DM. According to the consensus of $|\mathcal{S}| \ll |\mathcal{I}|$, *S2IDM* successfully addresses the long-standing scalability challenge that has persistently plagued PDRec. Moreover, SeeDRec exhibits the capability to comprehensively leverage the semantic associative relationships and cross-domain transferability under the sememe hierarchy (see the complexity comparison in Sec. 4.3). This effectively addresses the cross-domain generalization and semantic correlation challenges inherent in PDRec. The experimental results in Sec. 4.2 concurrently affirm that SeeDRec can yield significant and consistent performance improvements upon PDRec. Besides, we also employ the masked discrete DM [Austin *et al.*, 2021] in S2IDM, which yields comparable SR performance. So we just implement S2IDM in the regular mannner to simplify the computational complexity.

4 Experiments

We conduct extensive experiments on nine SR datasets and four CDSR datasets to answer the following questions:

- **RQ1:** Does the proposed SeeDRec outperform the base SR models and the SOTA DM-based SR methods?
- **RQ2:** How does SeeDRec perform in datasets where explicit item taxonomies (e.g., categories) are absent?
- **RQ3:** How does each component proposed in SeeDRec impact the recommendation performance?
- **RQ4:** Is our SeeDRec effective and universal enough with different base SR models and cross-domain SR tasks?
- **RQ5:** How does SeeDRec function in the interest distribution transfer and the few-shot scenarios?

Dataset	PixelRec	Home	Electronic	CD	Toy
Users	24,972	22,788	56,727	110,805	116,677
Items	44,643	23,603	45,279	105,841	77,760
Sememe	108	1,265	781	476	525
#Inter.	45,6813	187,778	507,373	1,342,060	1,018,540

Table 1: Statistics of five real-world SR datasets.

4.1 Experimental Settings

Datasets As shown in Table. 1, we construct five SR datasets from two platforms (i.e., Amazon [Lin *et al.*, 2022] and PixelRec [Cheng *et al.*, 2023]) with *existing categories* viewed as sememes. To verify that SeeDRec could work well without existing taxonomies, we also build four SR datasets and four CDSR datasets (with similar sizes) respectively, intentionally overlooking the inherent categories and using *words* as natural sememes via certain tokenization, lemmatization, and filtering on item titles with the NLTK library.

Baselines We implement SeeDRec on two SR models (SASRec [Kang and McAuley, 2018] and CL4SRec [Xie *et al.*, 2022]) and a SOTA DM-based model (PDRec [Ma *et al.*, 2024a]) to verify its effectiveness and universality. Besides, we also compare with two DM-based CF models (T-DiffRec [Wang *et al.*, 2023e] and TI-DiffRec [Ma *et al.*, 2024a]).

Parameter settings We conduct a comprehensive grid search to select the optimal hyper-parameters. That is, the learning rate is tuned from 0.001 to 0.05. The batch size and the maximum sequence length are defined as 512 and 200 for fair comparisons. It is imperative to underscore that SeeDRec essentially possesses very few parameters (e.g., k in *IPE*). We just assign $k = 3$ via our empirical knowledge. For *S2IDM*, we define $\omega_{min} = 0.1$, $\omega_{max} = 0.5$ and the step T as 10 for all datasets. We use the early-stop strategy to avoid overfitting. Following [Wang *et al.*, 2023c], we randomly sample 999 negative items for each positive instance to speed up the evaluation. All reported results are the average values of five runs with different seeds on the same NVIDIA Tesla V100.

4.2 Performance on SR (RQ1 & RQ2)

In this section, we leverage NDCG@10 (N@10), Hit rate@10 (H@10) and AUC as the evaluation metrics. From Table 2, we can observe that: (1) SeeDRec achieves significant improvements on all metrics and datasets compared to its base sequential models, with the significance level $p < 0.05$. It indicates that SeeDRec can precisely mine the user's interest distribution and successfully incorporate it into SR via the tailored multi-intention prompts on different types of base SR models. (2) Upon a lateral examination across various datasets, SeeDRec is more beneficial on relatively denser datasets. This aligns with its foundational assumption, that is, *S2IDM* can intricately better capture users' generalized interest distributions from the diffusion on minimal interest units on denser datasets where user behaviors are abundant. (3) SeeDRec attains the peak results across most of the datasets on the basic of PDRec. It confirms that SeeDRec exhibits the different underlying mechanisms of DM utilization with PDRec, positioning it as a stalwart support for the evolution of more advanced DM-based recommenders in the future.

Algorithm	Pixel			Home			Electronic			CD			Toy		
Metric	N@10	H@10	AUC	N@10	H@10	AUC	N@10	H@10	AUC	N@10	H@10	AUC	N@10	H@10	AUC
T-DiffRec	0.0436	0.0744	0.5659	0.0889	0.1127	0.5621	0.0953	0.1334	0.5820	0.2362	0.3321	0.7555	0.1279	0.1861	0.6526
TI-DiffRec	0.0420	0.0702	0.5621	0.0900	0.1136	0.5615	0.0990	0.1442	0.6223	0.2455	0.3432	0.7622	0.1316	0.1907	0.6580
SASRec	0.0857	0.1583	0.7769	0.0962	0.1267	0.6286	0.1172	0.1865	0.7483	0.3000	0.4450	0.8847	0.1691	0.2684	0.8025
+SeeDRec	**0.1037**	**0.1883**	**0.7950**	**0.1036**	**0.1387**	**0.6450**	**0.1342**	**0.2130**	**0.7672**	**0.3096**	**0.4574**	**0.8857**	**0.1814**	**0.2869**	**0.8096**
#Improv.	**21.00%**	**18.95%**	**2.33%**	**7.69%**	**9.47%**	**2.61%**	**14.51%**	**14.21%**	**2.53%**	**3.20%**	**2.79%**	**0.11%**	**7.27%**	**6.89%**	**0.88%**
CL4SRec	0.0833	0.1550	0.7739	0.0969	0.1282	0.6274	0.1172	0.1867	0.7481	0.3022	0.4482	0.8875	0.1697	0.2690	0.8026
+SeeDRec	**0.1039**	**0.1885**	**0.7937**	**0.1063**	**0.1427**	**0.6456**	**0.1347**	**0.2120**	**0.7681**	**0.3103**	**0.4590**	**0.8889**	**0.1818**	**0.2869**	**0.8115**
#Improv.	24.73%	**21.61%**	**2.56%**	**9.70%**	**11.31%**	**2.90%**	**14.93%**	**13.55%**	**2.67%**	**2.68%**	**2.41%**	**0.16%**	**7.13%**	**6.65%**	**1.11%**
PDRec	0.0887	0.1643	0.7886	0.0984	0.1310	0.6319	0.1218	0.1929	0.7551	0.3040	0.4556	0.8949	0.1752	0.2750	0.8021
+SeeDRec	**0.1067**	**0.1940**	**0.8038**	**0.1046**	**0.1401**	**0.6431**	**0.1364**	**0.2139**	**0.7744**	**0.3147**	**0.4686**	**0.8969**	**0.1867**	**0.2920**	**0.8116**
#Improv.	**20.29%**	**18.08%**	**1.93%**	**6.30%**	**6.95%**	**1.77%**	**11.99%**	**10.89%**	**2.56%**	**3.82%**	**2.99%**	**0.22%**	**6.56%**	**6.18%**	**1.18%**

Table 2: Results on sequential recommendation (using category as sememe). All improvements are significant ($p<0.05$ with paired t-tests).

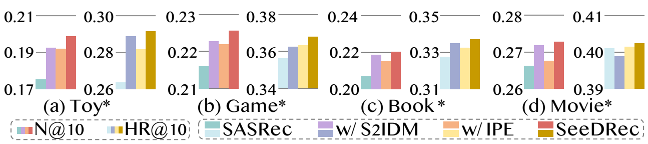

Figure 3: Results of SeeDRec (based on SASRec) and its ablation versions on four SR datasets (using word as sememe).

Dataset	Algorithm	#Par. (M)	GPU (MB)	#Tra. (s)	#Eval. (s)
CD	TI-DiffRec	201.99	6052	409.11	2193.92
	S2IDM	0.92	1408	364.9	827.78
Pixel	TI-DiffRec	85.2	3252	38.79	291.79
	S2IDM	0.21	1270	35.94	178.25

Table 3: Computational complexity comparison between TI-DiffRec and S2IDM on CD and Pixel. "Par.", "Tra." and "Eval." denote parameters, training time and evaluation time for one epoch.

To verify the practical usage of SeeDRec when datasets do not have appropriate categories, we build a more challenging setting that directly uses words of item titles (broadly existed in authentic datasets) as sememes. From Fig. 3, we observe that SeeDRec functions well without existing taxonomies (even comparable with using categories as sememes). It verifies the effectiveness and robustness of SeeDRec.

4.3 Ablation Study (RQ3)

To verify the effectiveness of each component in SeeDRec, we implement two versions, SASRec w/ S2IDM (leveraging a MLP-based feature fusion to replace our prompt-based fusion with the generalized user interest distribution) and SASRec w/ IPE (replacing the generalized user interest distribution given by S2IDM with the original interest distribution) for comparisons. We conduct ablation studies on four SR datasets (word as sememe) in Fig. 3, five SR datasets (category as sememe) in Fig. 4, and four CDSR datasets in Table 4. Here, SeeDRec equals SASRec+IPE+S2IDM.

We notice that: (1) Each component in SeeDRec can bring incremental improvement over its backbone on all datasets, where SeeDRec outperforms all ablation versions on different tasks, error range < 0.003. It validates the effectiveness of both *IPE* and *S2IDM* in SeeDRec. (2) The comparison

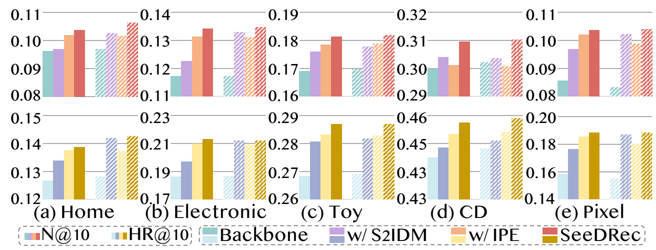

Figure 4: Ablation results of SeeDRec on SASRec (solid) and CL4SRec (slash) on five SR datasets (using category as sememe).

between SeeDRec and SeeDRec w/ IPE demonstrates the indispensability of *S2IDM* in SeeDRec, which simultaneously model the local personalized interests and the global sememe correlations via the tailored DM. Table 3 also verifies that the modification of diffusion objective enables the training and evaluation time of *S2IDM* in a low magnitude, making it hopeful in large-scale recommendation applications.

4.4 Performance on Cross-domain SR (RQ4)

Based on the transferability of SeeDRec when using words as sememes, SeeDRec could also handle CDSR tasks even without overlapped users (which is indispensable for most CDR models [Ma *et al.*, 2024b]). Specifically, we directly adopt the pre-trained and fixed *S2IDM* learned from the source domain (word as sememe) to infer the interest distribution (merely from the original sememe distribution in the target domain), and adopt it via *IPE* to enhance the target-domain SR.

The results in Table 4 demonstrate that: (1) SeeDRec outperforms all its ablation versions in CDSR, indicating the effectiveness and transferability of SeeDRec across domains. Both *IPE* and *S2IDM* are essential in CDSR. (2) It is crucial to emphasize that we directly use the source-domain *S2IDM* for the target domain SR, which is more challenging. Moreover, **SeeDRec does not impose a strict mandate for user overlap**, and thus is more practical and easy-to-deploy. These implicate the potential application of SeeDRec as a universal diffusion-based model for CDSR: training a general multi-domain *S2IDM* based on words, and transferring it to various target domains through domain-specific *IPE*.

Version	Game*→Toy*				Toy*→Game*				Movie*→Book*				Book*→Movie*			
Metric	N@5	N@10	H@5	H@10	N@5	N@10	H@5	H@10	N@5	N@10	H@5	H@10	N@5	N@10	H@5	H@10
SASRec	0.1140	0.1277	0.1470	0.1894	0.1941	0.2236	0.2711	0.3621	0.1903	0.2168	0.2605	0.3427	0.2033	0.2264	0.2645	0.3360
w/ *S2IDM*	0.1207	0.1369	0.1606	0.2109	0.2005	0.2297	0.2777	0.3682	0.2056	0.2304	0.2749	0.3516	0.2093	0.2320	0.2728	0.3431
w/ *IPE*	0.1265	0.1394	0.1600	0.2003	0.1996	0.2292	0.2772	0.3689	0.2036	0.2289	0.2727	0.3509	0.2114	0.2327	0.2694	0.3354
SeeDRec	**0.1274**	**0.1438**	**0.1674**	**0.2183**	**0.2008**	**0.2304**	**0.2783**	**0.3702**	**0.2067**	**0.2321**	**0.2758**	**0.3545**	**0.2130**	**0.2355**	**0.2759**	**0.3457**

Table 4: Results of SeeDRec and its ablation versions on four CDSR datasets. All improvements over the backbone are significant.

Dataset		Pixel		Home		Electronic	
Metric		N@10	H@10	N@10	H@10	N@10	H@10
Full	SASRec	0.0857	0.1583	0.0962	0.1267	0.1172	0.1865
	SeeDRec	0.1037	0.1883	0.1036	0.1387	0.1342	0.2130
	#Improv.	**21.0%**	**19.0%**	**7.7%**	**9.5%**	**14.5%**	**14.2%**
Few Shot	SASRec	0.0894	0.1618	0.1714	0.2072	0.1579	0.2269
	SeeDRec	0.1261	0.2115	0.1900	0.2301	0.1835	0.2624
	#Improv.	**41.0%**	**30.8%**	**10.9%**	**11.1%**	**16.2%**	**15.7%**

Table 5: Results of SeeDRec on full and few-shot settings.

4.5 Universality Analyses (RQ4)

We conduct SeeDRec on different classical SR models in five SR datasets to verify its universality. From Table 2 and Fig. 4, we have: (1) SeeDRec is a model-agnostic method that simultaneously exhibits its superiority on SASRec, CL4SRec, and even diffusion-based PDRec. Meanwhile, each component in SeeDRec achieves its incremental improvement on most datasets with different base models. (2) CL4SRec and PDRec outperform SASRec by advanced techniques such as contrastive learning and diffusion model, which is also reflected when armed with SeeDRec. It implies the compatibility between the sememe-based diffusion of SeeDRec and advanced sequential modeling techniques. Consequently, SeeDRec is likely to retain its universality and effectiveness in cooperating with future sophisticated sequential models.

4.6 In-depth Model Analyses (RQ5)

SeeDRec in Few-shot Setting The diffusion-based generalization exhibited by *S2IDM* serves to facilitate the exploration of associated interests for users with fewer behaviors or discovered interests. It motivates us to investigate SeeDRec's potential in few-shot scenarios. Hence, we set a sememe threshold for each dataset to simulate users that only have very few interests discovered by the real RS application and report the few-shot performance in Table 5.

We observe that: (1) Comparing with the full evaluation, SeeDRec achieves more significant improvements over SASRec on the few-shot setting. It verifies the feasibility of SeeDRec as a few-shot recommender thanks to the diffused interests. (2) Comparing the relative improvements across diverse datasets, it becomes evident that SeeDRec exhibits superior few-shot performance on Pixel, where the average number of sememes per item is comparatively lower. This may stem from the fact that SeeDRec elegantly addresses the twofold challenges of behavioral sparsity and preference unitary in this particular setting, thereby implying some ingenious usages of SeeDRec for new users (e.g., asking new users to pick several "seed" interests and generalize them via SeeDRec).

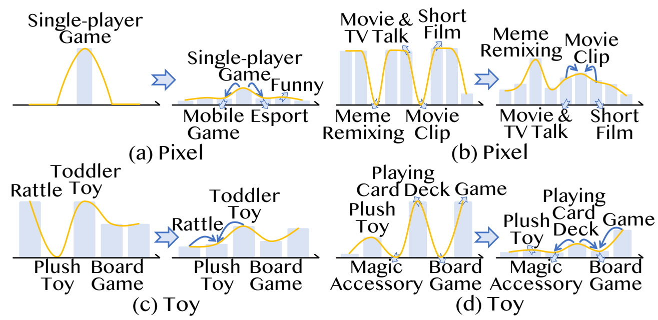

Figure 5: Visualization of *S2IDM* from original to generalized distributions, with rectangles denoting diffused sememe probabilities.

Case Study of User Interest Diffusion We visualize the original sememe distribution and the generalized interest distribution of *S2IDM* to illustrate how *S2IDM* affects the sememe-based interest diffusion. Fig. 5 shows several representative cases on Pixel and Toy. Take Fig. 5 (a) as an example, the user's historical preference singularly comprises one sememe as "Single-player Game". Following the successive inference process of *S2IDM*, it not only generalizes "Esport" and "Mobile Game", which are strongly linked to the core notion of "Single-player Game", but also captures some weakly-associated yet widely-popular sememes like "Funny". This aligns with the conceptual rationale of *S2IDM*, that is, it can accurately mine each user's diffused interests by simultaneously considering the intricate interplay between the local interest evolution and the global interest generalization.

5 Conclusion and Future Work

This paper proposes a Sememe-based Diffusion framework (SeeDRec), which captures each user's sememe-based diffusion-generalized interest distribution to enhance SR. With the proposed recommendation sememe powered by *S2IDM* and *IPE*, SeeDRec is verified to be effective, transferable, and scalable on thirteen SR and CDSR datasets. The proposed SeeDRec could also be easily adopted with different types of sequential models without much additional inference computation costs, which will be welcomed by the industry.

Future investigations should encompass the design of more cogent minimal interest units and the exploration of the semantic correlations among sememes during the diffusion process. Furthermore, we believe that SeeDRec indicates the future direction for the subsequent DM-based recommenders beyond the diffusion on indices. The proposed *recommendation sememe* and *S2IDM* can facilitate seamless integration into diverse scenarios to bring consistent improvements.

Acknowledgments

This work is supported in part by the National Key R&D Program of China (Grant no. 2021YFC3300203), the TaiShan Scholars Program (Grant no. tsqn202211289), the Shandong Province Excellent Young Scientists Fund Program (Overseas) (Grant no. 2022HWYQ-048), the Oversea Innovation Team Project of the "20 Regulations for New Universities" funding program of Jinan (Grant no. 2021GXRC073) and the Young Elite Scientists Sponsorship Program by CAST (2023QNRC001). ChatGPT and Grammarly are utilized to improve grammar and correct spelling. Corresponding Authors: Lei Meng and Ruobing Xie.

References

[Austin *et al.*, 2021] Jacob Austin, Daniel D Johnson, Jonathan Ho, Daniel Tarlow, and Rianne Van Den Berg. Structured denoising diffusion models in discrete state-spaces. *Proceedings of Advances in neural information processing systems (NeurIPS)*, 34:17981–17993, 2021.

[Brempong *et al.*, 2022] Emmanuel Asiedu Brempong, Simon Kornblith, Ting Chen, Niki Parmar, Matthias Minderer, and Mohammad Norouzi. Denoising pretraining for semantic segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, 2022.

[Chen *et al.*, 2023] Linyao Chen, Aosong Feng, Boming Yang, and Zihui Li. Xdlm: Cross-lingual diffusion language model for machine translation. *arXiv preprint arXiv:2307.13560*, 2023.

[Cheng *et al.*, 2023] Yu Cheng, Yunzhu Pan, Jiaqi Zhang, Yongxin Ni, Aixin Sun, and Fajie Yuan. An image dataset for benchmarking recommender systems with raw pixels. *arXiv preprint arXiv:2309.06789*, 2023.

[de Souza Pereira Moreira *et al.*, 2021] Gabriel de Souza Pereira Moreira, Sara Rabhi, Jeong Min Lee, Ronay Ak, and Even Oldridge. Transformers4rec: Bridging the gap between nlp and sequential/session-based recommendation. In *Proceedings of the 15th ACM Conference on Recommender Systems*, pages 143–153, 2021.

[Du *et al.*, 2023] Hanwen Du, Huanhuan Yuan, Zhen Huang, Pengpeng Zhao, and Xiaofang Zhou. Sequential recommendation with diffusion models. *arXiv preprint arXiv:2304.04541*, 2023.

[Fan *et al.*, 2023] Ziwei Fan, Zhiwei Liu, Hao Peng, and Philip S Yu. Mutual wasserstein discrepancy minimization for sequential recommendation. In *Proceedings of International World Wide Web Conferences (WWW)*, pages 1375–1385, 2023.

[Hidasi *et al.*, 2016] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. Session-based recommendations with recurrent neural networks. In *Proceedings of International Conference on Learning Representations (ICLR)*, pages 1–10, 2016.

[Ho *et al.*, 2020] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Proceedings of Advances in neural information processing systems (NeurIPS)*, 33:6840–6851, 2020.

[Kang and McAuley, 2018] Wang-Cheng Kang and Julian McAuley. Self-attentive sequential recommendation. In *Proceedings of International Conference on Data Mining (ICDM)*, pages 197–206, 2018.

[Kingma and Welling, 2013] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[Kong *et al.*, 2021] Zhifeng Kong, Wei Ping, Jiaji Huang, Kexin Zhao, and Bryan Catanzaro. Diffwave: A versatile diffusion model for audio synthesis. In *Proceedings of International Conference on Learning Representations (ICLR)*, pages 1–17, 2021.

[Lester *et al.*, 2021] Brian Lester, Rami Al-Rfou, and Noah Constant. The power of scale for parameter-efficient prompt tuning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3045–3059, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.

[Li *et al.*, 2023] Zihao Li, Aixin Sun, and Chenliang Li. Diffurec: A diffusion model for sequential recommendation. *ACM Transactions on Information Systems (TOIS)*, 42(3):1–28, 2023.

[Lin *et al.*, 2022] Guanyu Lin, Chen Gao, Yinfeng Li, Yu Zheng, Zhiheng Li, Depeng Jin, and Yong Li. Dual contrastive network for sequential recommendation with user and item-centric perspectives. In *Proceedings of International Conference on Research on Development in Information Retrieval (SIGIR)*, 2022.

[Ma *et al.*, 2021] Haokai Ma, Xiangxian Li, Lei Meng, and Xiangxu Meng. Comparative study of adversarial training methods for cold-start recommendation. In *Proceedings of the 1st International Workshop on Adversarial Learning for Multimedia*, pages 28–34, 2021.

[Ma *et al.*, 2023a] Haokai Ma, Zhuang Qi, Xinxin Dong, Xiangxian Li, Yuze Zheng, Xiangxu Meng, and Lei Meng. Cross-modal content inference and feature enrichment for cold-start recommendation. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2023.

[Ma *et al.*, 2023b] Haokai Ma, Ruobing Xie, Lei Meng, Xin Chen, Xu Zhang, Leyu Lin, and Jie Zhou. Exploring false hard negative sample in cross-domain recommendation. In *Proceedings of the 17th ACM Conference on Recommender Systems (Recsys)*, pages 502–514, 2023.

[Ma *et al.*, 2024a] Haokai Ma, Ruobing Xie, Lei Meng, Xin Chen, Xu Zhang, Leyu Lin, and Zhanhui Kang. Plug-in diffusion model for sequential recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 38, pages 8886–8894, 2024.

[Ma *et al.*, 2024b] Haokai Ma, Ruobing Xie, Lei Meng, Xin Chen, Xu Zhang, Leyu Lin, and Jie Zhou. Triple sequence learning for cross-domain recommendation. *ACM Transactions on Information Systems (TOIS)*, 42(4):1–29, 2024.

[Meng *et al.*, 2020] Lei Meng, Fuli Feng, Xiangnan He, Xiaoyan Gao, and Tat-Seng Chua. Heterogeneous fusion of semantic and collaborative information for visually-aware food recommendation. In *Proceedings of the 28th ACM international conference on multimedia (MM)*, pages 3460–3468, 2020.

[Nichol and Dhariwal, 2021] Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *Proceedings of International Conference on Machine Learning (ICML)*, volume 139, pages 8162–8171, 2021.

[Niu *et al.*, 2017] Yilin Niu, Ruobing Xie, Zhiyuan Liu, and Maosong Sun. Improved word representation learning with sememes. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2049–2058, 2017.

[Qi *et al.*, 2022] Fanchao Qi, Chuancheng Lv, Zhiyuan Liu, Xiaojun Meng, Maosong Sun, and Hai-Tao Zheng. Sememe prediction for BabelNet synsets using multilingual and multimodal information. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 158–168, 2022.

[Rendle *et al.*, 2010] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of International World Wide Web Conferences (WWW)*, pages 811–820, 2010.

[Sun *et al.*, 2019] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*, 2019.

[Sun *et al.*, 2022] Weilin Sun, Xiangxian Li, Manyi Li, Yuqing Wang, Yuze Zheng, Xiangxu Meng, and Lei Meng. Sequential fusion of multi-view video frames for 3d scene generation. In *CAAI International Conference on Artificial Intelligence*, pages 597–608. Springer, 2022.

[Sun *et al.*, 2024] Weilin Sun, Manyi Li, Peng Li, Xiao Cao, Xiangxu Meng, and Lei Meng. Sequential selection and calibration of video frames for 3d outdoor scene reconstruction. *CAAI Transactions on Intelligence Technology*, pages 1–15, 2024.

[Tang and Wang, 2018] Jiaxi Tang and Ke Wang. Personalized top-n sequential recommendation via convolutional sequence embedding. In *Proceedings of ACM International Conference on Web Search and Data Mining (WSDM)*, pages 565–573, 2018.

[Walker *et al.*, 2022] Joojo Walker, Ting Zhong, Fengli Zhang, Qiang Gao, and Fan Zhou. Recommendation via collaborative diffusion generative model. In *Proceedings of International Conference on Knowledge Science, Engineering and Management*, pages 593–605. Springer, 2022.

[Wang *et al.*, 2015] Hao Wang, Naiyan Wang, and Dit-Yan Yeung. Collaborative deep learning for recommender systems. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1235–1244, 2015.

[Wang *et al.*, 2023a] Changshuo Wang, Lei Wu, Xu Chen, Xiang Li, Lei Meng, and Xiangxu Meng. Letter embedding guidance diffusion model for scene text editing. In *2023 IEEE International Conference on Multimedia and Expo (ICME)*, pages 588–593. IEEE, 2023.

[Wang *et al.*, 2023b] Changshuo Wang, Lei Wu, Xiaole Liu, Xiang Li, Lei Meng, and Xiangxu Meng. Anything to glyph: Artistic font synthesis via text-to-image diffusion model. In *SIGGRAPH Asia 2023 Conference Papers*, pages 1–11, 2023.

[Wang *et al.*, 2023c] Chenyang Wang, Weizhi Ma, Chong Chen, Min Zhang, Yiqun Liu, and Shaoping Ma. Sequential recommendation with multiple contrast signals. *ACM Transactions on Information Systems*, 41(1):1–27, 2023.

[Wang *et al.*, 2023d] Ran Wang, Zhuang Qi, Xiangxu Meng, and Lei Meng. Learning to fuse residual and conditional information for video compression and reconstruction. In *International Conference on Image and Graphics*, pages 360–372. Springer, 2023.

[Wang *et al.*, 2023e] Wenjie Wang, Yiyan Xu, Fuli Feng, Xinyu Lin, Xiangnan He, and Tat-Seng Chua. Diffusion recommender model. *Proceedings of International Conference on Research on Development in Information Retrieval (SIGIR)*, 2023.

[Wang *et al.*, 2023f] Yu Wang, Zhiwei Liu, Liangwei Yang, and Philip S Yu. Conditional denoising diffusion for sequential recommendation. *arXiv preprint arXiv:2304.11433*, 2023.

[Wang *et al.*, 2023g] Yuqing Wang, Zhuang Qi, Xiangxian Li, Jinxing Liu, Xiangxu Meng, and Lei Meng. Multi-channel attentive weighting of visual frames for multimodal video classification. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2023.

[Wu *et al.*, 2024] Yiqing Wu, Ruobing Xie, Yongchun Zhu, Fuzhen Zhuang, Xu Zhang, Leyu Lin, and Qing He. Personalized prompt for sequential recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 2024.

[Xie *et al.*, 2022] Xu Xie, Fei Sun, Zhaoyang Liu, Shiwen Wu, Jinyang Gao, Jiandong Zhang, Bolin Ding, and Bin Cui. Contrastive learning for sequential recommendation. In *Proceedings of IEEE International Conference on Data Engineering (ICDE)*, 2022.

[Yang *et al.*, 2023] Zhengyi Yang, Jiancan Wu, Zhicai Wang, Xiang Wang, Yancheng Yuan, and Xiangnan He. Generate what you prefer: Reshaping sequential recommendation via guided diffusion. *Proceedings of Advances in neural information processing systems (NeurIPS)*, 2023.

[Zhao *et al.*, 2024] Jujia Zhao, Wenjie Wang, Yiyan Xu, Teng Sun, and Fuli Feng. Denoising diffusion recommender model. *arXiv preprint arXiv:2401.06982*, 2024.