# 释放邻居的潜力：用于基于会话推荐的扩散式潜在邻居生成

Yuhan Yang 电子科技大学 中国成都 y.yuhan2000@gmail.com

Jie Zou*电子科技大学 中国成都 jie.zou@uestc.edu.cn

Guojia An 电子科技大学 中国成都 an_guojia@163.com

Jiwei Wei 电子科技大学 中国成都 mathematic6@gmail.com

Yang Yang 电子科技大学，中国成都 yang.yang@uestc.edu.cn

Heng Tao Shen 电子科技大学，中国成都 shenhengtao@hotmail.com

arXiv:2601.03903v1 [cs.IR] 7 Jan 2026

## 摘要

基于会话的推荐旨在根据匿名用户在当前会话中的交互，预测他们可能感兴趣的下一个物品。近期研究表明，检索邻居会话以增强当前会话，可以有效缓解数据稀疏性问题并提升推荐性能。然而，现有方法通常依赖显式观测到的会话数据，忽视了潜在邻居——这些邻居无法被直接观测到，但在兴趣空间中可能相关——因此未能充分发挥邻居会话在推荐中的潜力。

为解决上述局限性，我们提出了一种基于扩散的潜在邻居生成新模型，用于基于会话的推荐，称为**DiffSBR**。具体而言，DiffSBR利用两个扩散模块——检索增强扩散和自增强扩散——来生成高质量的潜在邻居。在检索增强扩散模块中，我们利用检索到的邻居作为引导信号，以约束并重建潜在邻居的分布。同时，我们采用一种训练策略，使检索器能够从生成器提供的反馈中学习。在自增强扩散模块中，我们通过对比学习注入当前会话的多模态信号，明确引导潜在邻居的生成。获得生成的潜在邻居后，我们利用它们增强会话表示，从而提升基于会话的推荐效果。在四个公开数据集上的大量实验表明，DiffSBR能够生成有效的潜在邻居，并且相较于最先进的基线方法提升了推荐性能。

## CCS 概念

• **信息系统** →**推荐系统**。

*通讯作者。

本作品采用知识共享署名-非商业性禁止演绎 4.0 国际许可协议。*KDD '26*，韩国济州岛© 2026 版权所有归所有者/作者所有。

ACM ISBN 978-x-xxxx-xxxx-x/YYYY/MM
https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 关键词

基于会话的推荐、扩散模型、多模态

**ACM 参考格式：**

Yuhan Yang、Jie Zou、Guojia An、Jiwei Wei、Yang Yang 和 Heng Tao Shen。2026。《释放邻居的潜力：用于基于会话推荐的扩散式潜在邻居生成》。载于第 *32* 届 *ACM SIGKDD* 知识发现与数据挖掘会议论文集 *V.1*（*KDD '26*），*2026*年*8*月*9*日至*13*日，大韩民国济州岛。ACM，纽约，NY，美国，10 页。https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 引言

随着多媒体内容呈指数级增长，用户在寻找符合自身偏好的信息时面临越来越多的挑战 [35, 55]。幸运的是，推荐系统 [2, 3, 38, 39, 56] 的出现缓解了这一问题。然而，由于人们日益关注用户隐私，获取用户的个人信息和历史交互往往并不可行。为解决这一问题，基于会话的推荐（SBR） [4, 22, 43] 应运而生，成为一种很有前景的解决方案。SBR 旨在仅根据匿名用户在一个会话中的交互来推荐下一个物品。

在基于会话的推荐中，缺少用户画像，再加上会话长度较短，会加剧数据稀疏性，从而降低整体推荐性能。为缓解这一问题，近期研究主要采用基于检索的邻居方法 [29, 51] 挖掘邻居信息，并增强目标会话表示。这些方法大致可分为基于相似度的方法和基于共现的方法。前者（例如 ICM-SR[25], DIDN [49], TASI-GNN [27],ECCL [1]）通过静态相似度学习筛选语义相似的会话；后者（例如 FGNN [30], MSGAT [29], DGNN [19], DIMO [51]）则利用物品共现模式捕捉结构关系。

尽管这些现有的邻居检索方法在基于会话的推荐领域取得了显著进展，但仍面临固有的局限性。这些方法受数据集范围的限制，只能检索现有的可观测邻居，即数据集中明确包含的兴趣一致的会话。然而，这一范式忽视了潜在邻居，它们未被记录在数据集中，因此无法通过检索方法发现，但仍与用户的潜在兴趣一致，

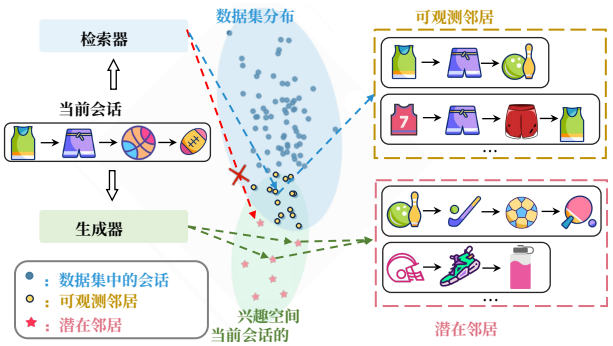



**图1：基于检索的方法与生成式方法的对比。蓝色区域表示数据集的分布，绿色区域表示当前会话的完整兴趣空间。基于检索的方法只能获取这两个区域交集中的兴趣匹配邻居，因此受观测数据限制。相比之下，生成式方法通过探索整个兴趣空间克服了这一局限，从而能够生成可能相关但未被观测到的潜在邻居。**

从而限制了邻居的表达能力，并最终制约推荐结果。具体而言，如图1所示，我们首先将用户当前会话表示投影到统一的兴趣空间中。在检索范式下，可观测邻居取自数据集中存在的样本（图1中的蓝色区域）。尽管检索到的邻居可能只能部分覆盖用户的兴趣空间，但它们从根本上受数据集边界的限制，因而无法触及用户兴趣空间的完整范围。相比之下，生成式范式对用户兴趣的潜在分布进行建模，并扩展邻居的范围（图1中绿色区域内但位于绿蓝重叠区域之外的部分）。这种方法能够主动探索超出观测数据的兴趣区域，从而生成传统检索方法无法获取的潜在邻居。

为克服上述检索范式的固有局限，我们迈出了从基于检索的框架向生成范式转变的第一步，旨在生成超出可观测邻居范围的潜在邻居。受扩散模型在建模复杂数据分布方面取得的显著进展启发 [6, 45], , 我们采用扩散作为框架中的主要生成机制。尽管扩散模型具有颇具前景的生成能力，但要确保生成的潜在邻居的质量和有效性，仍面临两项关键挑战：(1) 如何有效设计引导机制，以提升扩散模型生成的潜在邻居的质量。(2) 如何在扩散生成过程中有效整合多模态信息，以增强潜在邻居的生成。

为应对上述两个挑战，我们提出了一种新颖的**Diffusion-based** 潜在邻居生成模型，用于**Session-based Recommendation**，称为**DiffSBR**，它引导扩散模型生成高质量的潜在邻居。具体而言，为应对第一个挑战，我们提出了一个名为**检索增强扩散模块**的新模块，通过在扩散生成过程中融入检索到的先验邻居信息来控制生成，从而有针对性地生成潜在邻居。此外，该模块还引入了一种训练策略，利用邻居生成器的损失信号反向优化邻居检索器，在检索与生成之间建立闭环协作。为应对上述第二个挑战，我们设计了**自增强扩散模块**，将多模态信息融入潜在邻居生成过程。具体而言，我们利用多模态信息引导生成，随后通过与检索增强扩散输出进行对比学习，更好地融合多模态语义。

在本文中，我们的主要贡献总结如下：

- 我们首次尝试生成潜在邻居，并强调其重要性，旨在缓解传统基于检索的方法在基于会话的推荐中固有的局限性。
- 我们提出了一种新颖的 DiffSBR 模型，该模型通过检索增强扩散模块和自增强扩散模块生成有效的潜在邻居，从而提升推荐性能。
- 我们在四个公开数据集上进行了大量实验，结果表明，所提出的 DiffSBR 模型不仅显著优于现有的基于会话的推荐方法，还证明了生成潜在邻居的有效性和必要性。

## 2 相关工作

### 2.1 基于会话的推荐

基于会话的推荐模型从短期匿名交互中建模用户偏好。FPMC [46] 等早期方法将马尔可夫链与矩阵分解相结合，但无法捕捉高阶依赖。GRU4Rec [11] 和 NARM [17] 通过循环神经网络和注意力机制改进序列建模。近期基于图神经网络的模型（如 SR-GNN [40]、TAGNN [47]）将会话视为图，以更好地捕捉复杂的物品关系。然而，这些模型仅依赖会话内信息，忽略了邻近会话中的宝贵信号。为缓解这一问题，一些方法会检索邻居会话，以丰富当前会话表示。

**基于检索的邻居方法在基于会话的推荐中的应用**。这些方法旨在整合跨会话的邻居信息，以增强稀疏交互下的会话表示。现有研究大致可分为两类。第一类侧重于基于相似度的检索方法 [1, 29], , 通过计算相似度来识别相关邻居会话。例如，ICM- SR [25] 采用意图引导的邻居检测器来定位相关会话，而 DIDN [49] 则利用动态意图感知模块检索语义相似的会话。第二类工作利用共现关系 [37, 51], , 基于物品共现构建全局图，以捕捉会话间的成对转移。例如，CGL [54] 构建全局图来建模会话间相关性，从而增强物品表示。MSGAT [29] 进一步构建会话级关系图，并引入意图感知协作模块，以优化会话表示。

尽管上述基于检索的方法能够取得良好效果，但它们只能从数据集中检索可观测邻居，因此限制了邻居的质量。相比之下，我们利用扩散模型的生成能力来生成潜在邻居，从而发现数据中未明确呈现的语义相关邻居。

**基于多模态的方法在基于会话的推荐中的应用**。由于用户兴趣往往受多模态内容驱动，仅依赖ID特征不足以反映真实偏好。近期研究利用丰富的物品特征来改进用户意图建模。例如，CoHHN[52] 将价格作为一种关键模态；MMSBR [50] 结合物品文本和图像；LLM4SBR [28] 利用文本描述并提示大语言模型进行意图推断；DIMO [51] 则将ID特征与多模态特征解耦并融合。据我们所知，与这些方法不同，我们的方法首次在基于会话的推荐中将多模态信号明确整合到邻居生成过程中。

## 2.2 推荐中的扩散模型

近年来，扩散模型已成为一种颇具前景的推荐生成框架，在不确定性建模方面能力强大，并能灵活生成偏好。早期研究如 DiffRec [34] 将扩散应用于用户偏好建模和物品生成。DiffuRec [18] 和 DiQDiff [24] 通过反向扩散将这一思路拓展到序列推荐。除此之外，DCASR[10] 和 DiffuASR [20] 探索了基于扩散的数据增强，以改进序列表示。在序列之外，DiffKG [15]和 DiffMM [14] 将扩散引入结构化数据，包括知识图谱和多模态交互图。近期，DDRM [53] 和 MCDRec [23], 采用条件扩散，将用户偏好融入生成过程，从而增强个性化和语义对齐。尽管以往研究已验证扩散模型的有效性，但大多数方法聚焦于目标物品生成或序列增强。相比之下，本文首次尝试在基于会话的推荐中生成潜在邻居信息，而这一方向此前鲜有研究。

# 3 问题定义

令 $\mathrm{V} = \{v_1、\ \ldots\ldots、\ v_i、\ \ldots\ldots、\ v_n\}$ 表示数据集中所有物品的集合，其中 $n$ 是唯一物品的总数。每个物品 $v_i \in \mathrm{V}$ 由一个标识符 $v_i^{id}$ 和多模态内容特征 $v_i^{mo}$ 构成，即 $v_i = \{v_i^{id}, v_i^{mo}\}$。在本研究中，多模态内容特征$v_i^{mo}$包括文本和视觉模态：$v_i^{mo} = \left\{v_i^{txt}, v_i^{img}\right\}$，其中 $v_i^{txt} = \{w_1, w_2, \ldots, w_q\}$ 表示描述物品标题和品牌名称的一组 $q$ 词语，而 $v_i^{img}$ 表示物品 $v_i$ 对应的图像。令 $\mathrm{S} = \{s_1, \ldots, s_j, \ldots, s_{|\mathcal{S}|}\}$ 表示所有会话的集合，其中 $|\mathrm{S}|$ 是会话总数。每个会话$s_j$是匿名用户在短时间内产生的有序序列，形式化定义为：$S_j = [v_1, v_2, \ldots, v_m]$，其中 $m$ 是当前会话的长度。给定当前会话历史 $[v_1, v_2, \ldots, v_m]$，基于会话的推荐的目标是预测用户最有可能交互的下一个物品 $v_{m+1}$。

# 4 方法

本节将详细介绍我们提出的 DiffSBR 框架，如图 2 所示。每个输入会话首先经过 ID 和多模态会话表示模块编码，从而得到会话 ID 嵌入和会话模态嵌入。随后，这些表示被输入检索增强扩散模块和自增强扩散模块，以生成潜在邻居，从而增强当前会话，完成最终推荐任务。

## 4.1 ID 与多模态会话表示

在这里，我们旨在通过三个主要步骤，分别为每个会话获取会话 ID 嵌入和会话模态嵌入。

*4.1.1* 物品嵌入的初始化。鉴于各种模态的呈现方式不同，我们采用特定的编码方法，将原始模态数据转换为向量表示。对于每个物品 $v_i$，我们构建三种类型的嵌入：结构化 ID 嵌入、从文本描述中提取的语义嵌入，以及从物品图像中提取的视觉嵌入。

对于每个物品 $v_i$ 的结构化 ID 嵌入 $v_i^{id}$，我们遵循以往工作的常见做法（例如，[17, 40]），构建一个 ID 嵌入表 $\mathbf{E}^{id} \in \mathbb{R}^{n \times d}$，其中每一行对应于特定物品的一个随机初始化 ID 嵌入 $\mathbf{e}_i^{id} \in \mathbb{R}^d$。

文本模态与视觉模态之间存在显著的语义差距。与 Zhang 等人 [51], 类似，我们通过将视觉模态转换为文本形式，来解决这一问题，从而将异构模态对齐到共享语义空间中。具体而言，我们使用 GoogLeNet [31] 预测每个物品图像的前 2 个类别标签，并将其串联起来，形成伪文本表示 $v_i^{ctxt} = \{w_1', w_2', \ldots\ldots, w_o'\}$。随后，将其与原始物品描述 $v_i^{txt} = \{w_1, w_2, \ldots\ldots, w_q\}$进一步串联，构建统一多模态输入 $v_i^{mo} = \{w_1, \ldots\ldots, w_q, w_1', \ldots\ldots, w_o'\}$。我们将 $v_i^{mo}$输入预训练的 BERT [5] 模型，以获取上下文化词元嵌入 $\{\mathbf{e}_1, \mathbf{e}_2, \ldots\ldots, \mathbf{e}_{l+o}\}$，并应用平均池化来初始化多模态物品表示 $\mathbf{e}_i^{mo} \in \mathbb{R}^d$，其计算方式如下：

$$\mathbf{e}_i^{mo} = \frac{1}{q+o} \sum_{r=1}^{q+o} \mathbf{e}_r. \tag{1}$$

*4.1.2* 图增强物品嵌入。获得初始化的物品嵌入后，我们引入图结构，以进一步建模转移依赖关系和协同关系。遵循先前的研究 [42, 51], ，我们根据共现模式构建有向物品图 $\mathrm{G} = (\mathrm{V}, \mathrm{E})$，其中每个节点代表一个物品，两个物品之间的边权重反映它们的共现频率。基于该图，我们得到相应的邻接矩阵 $\mathbf{A} \in \mathbb{R}^{n \times n}$，其中 $\mathbf{A}(i, j)$ 表示从物品 $v_i$指向物品 $v_j$的边权重。我们还构建度矩阵 $\mathbf{D} \in \mathbb{R}^{n \times n}$，其中 $\mathbf{D}(i, i)$ 表示物品 $v_i$的出度，其定义为该物品所有出边权重之和。

接着，我们使用 $L$层图卷积网络（GCN）进行消息聚合。对于 ID 嵌入和多模态嵌入（分别记作$\mathbf{e}_i^{id}$和$\mathbf{e}_i^{mo}$），第 $l$层的更新定义如下：

$$\mathbf{x}_{v_i^c}^{(l)} = \mathrm{Norm}\left(\mathbf{D}^{-1}\mathbf{A}\mathbf{x}_{v_i^c}^{(l-1)}\mathbf{W}^{(l-1)}\right), \tag{2}$$

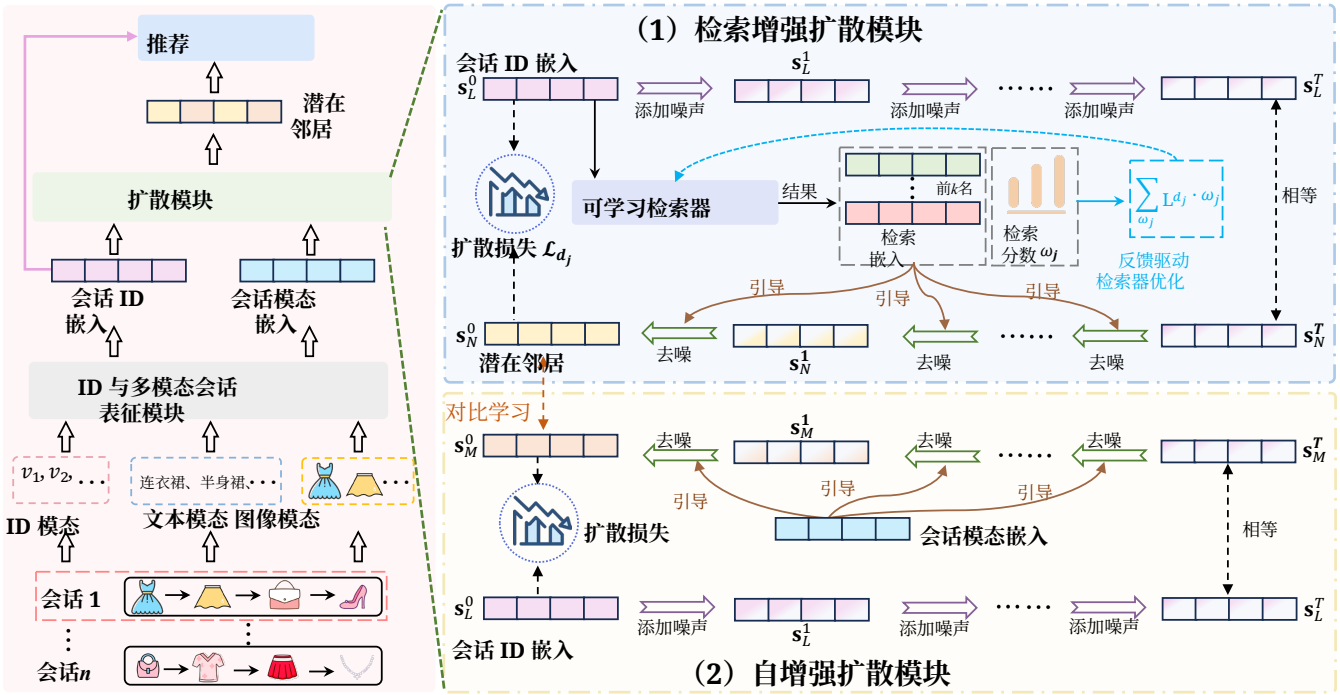



**图 2：DiffSBR 的整体框架。左侧为整体框架，右侧展示了扩散模块的两个主要组成部分：（1）检索增强扩散模块，根据检索先验引导生成潜在邻居；（2）自增强扩散模块，利用会话自身的多模态信息来提升潜在邻居的质量。**

其中，$c \in \{id,\ mo\}$，$\mathbf{x}_{v_i^c}^{(l)}$ 表示第 $l$层图卷积网络层的物品嵌入，$\mathbf{W}^{(l-1)} \in \mathbb{R}^{d\times d}$是一个可学习变换矩阵。Norm(·)表示归一化函数。图卷积网络的输入是初始化后的物品嵌入：$\mathbf{x}_{v_i^{id}}^{0} = \mathbf{e}_i^{id}$、$\mathbf{x}_{v_i^{mo}}^{0} = \mathbf{e}_i^{mo}$。最后，我们聚合所有图卷积网络层的表示，得到最终的图增强物品嵌入$\mathbf{x}^{v_i^c}$：

$$\mathbf{x}_{v_i^c} = \frac{1}{L+1}\sum_{l=0}^{L}\mathbf{x}_{v_i^c}^{(l)}. \tag{3}$$

*4.1.3* 会话表示。为进一步获得会话级表示，我们首先采用对比对齐，促使 ID 嵌入与多模态嵌入保持一致，从而稳定多模态训练；随后，参照 Wu 等人 [40]，计算序列中所有物品的相关性分数。为考虑物品嵌入重要性的差异，我们引入基于注意力的聚合 [7]，对每个物品进行自适应加权。注意力权重 $\alpha_i$ 的计算方式如下：

$$\alpha_i = \sigma\left(\mathbf{W}_1\mathbf{x}_{v_m^c} + \mathbf{W}_2\mathbf{x}_{v_i^c}\right), \tag{4}$$

其中，$\mathbf{W}_1$ 和$\mathbf{W}_2$ 是可学习参数，$\mathbf{x}_{v_m^c}$是序列中的最后点击的物品，$\sigma(\cdot)$表示 sigmoid 激活函数。随后，通过使用学习到的注意力权重聚合物品嵌入，得到最终的会话表示$\mathbf{s}_c$：

$$\mathbf{s}_c = \sum_{i=1}^{m}\alpha_i \cdot \mathbf{x}_{v_i^c}, \tag{5}$$

where $\mathbf{s}_c \in \{\mathbf{s}_{id}, \mathbf{s}_{mo}\}$.

## 4.2 检索增强扩散模块

在本研究中，我们采用扩散模型作为生成骨干网络，因为与生成对抗网络（GANs）[8] 和变分自编码器（VAEs） [26], 相比，扩散模型具有更高的训练稳定性和更高质量的样本生成能力，这一点已在先前研究中得到证实 [18, 24, 45]。此外，扩散模型在推荐任务中已展现出良好的性能，因此非常适合用作我们的生成框架。

为确保扩散模型以目标导向的方式探索兴趣空间，并防止其偏离语义无关或与用户无关的区域，我们设计了一个检索增强扩散模块。该模块利用检索到的先验知识引导扩散过程，从而生成语义连贯且与用户当前偏好高度契合的结果。

*4.2.1* 基于检索的先验构建。为了构建能够为扩散过程提供先验引导的邻居集合，我们首先引入一个可学习检索器，从历史会话数据库中选出与当前会话最相关的 top-$k$ 个邻居。具体而言，给定一个会话 $\mathbf{s}_{id}$，我们使用以下公式计算其相似度分数：

$$\mathrm{sim}^D = \mathcal{F}_{\mathrm{score}}\left(\left[\mathbf{s}_{id} \| \mathbf{s}^D\right]\right), \tag{6}$$

其中 $\mathrm{F}_{\mathrm{score}}$ 表示一个多层感知机，而 $[\cdot\|\cdot]$ 表示向量拼接。然后，我们选出相似度分数最高的前 $k$ 个候选会话

$\mathrm{N}_k$。接下来，我们应用 softmax 函数对相似度分数进行归一化，以获得每个邻居的注意力权重：

$$\omega_j = \frac{\exp\left(\mathrm{sim}_j^D\right)}{\sum_{r=1}^{k}\exp\left(\mathrm{sim}_r^D\right)}, \tag{7}$$

其中 $j \in \{1, \ldots\ldots, k\}$。邻居会话 $\mathrm{N}_k$及其对应的分数权重将作为条件信息，用于引导下游过程中的扩散模型。

*4.2.2* 基于扩散的邻居生成。在获得用于引导扩散过程的先验信息后，我们采用条件去噪扩散概率模型 [44] 训练去噪网络，该网络在推理过程中逐步移除噪声，以生成条件引导的潜在邻居。具体而言，在训练期间，我们构建一个前向马尔可夫链，逐步向原始数据样本（记为干净的会话嵌入 $\mathbf{s}_{id} = \mathbf{s}_L^0$）添加高斯噪声，共经历 $T$ 个时间步 $t \in \{1, \ldots\ldots, T\}$。该过程最终生成一个近似于标准正态分布的带噪向量 $\mathbf{s}_L^T$。每次前向转移定义如下：

$$q\left(\mathbf{s}_L^t \mid \mathbf{s}_L^{t-1}\right) = \mathcal{N}\left(\mathbf{s}_L^t; \sqrt{1-\beta_t}\,\mathbf{s}_L^{t-1}, \beta_t \mathbf{I}\right), \tag{8}$$

其中，$\beta_t \in (0, 1)$ 表示时间步 $t$ 的噪声调度，$\mathbf{I}$ 表示单位矩阵。令 $\alpha_t = 1 - \beta_t$，并定义累积乘积为：$\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_t$。随后，第 $t$ 步的噪声样本可由干净嵌入直接推导得到：

$$\mathbf{s}_L^t = \sqrt{\bar{\alpha}_t}\mathbf{s}_L^0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, \tag{9}$$

其中，$\boldsymbol{\epsilon} \sim \mathrm{N}(0, \mathrm{I})$ [12, 36]。遵循先前研究 [24, 45, 53],，我们不会在训练过程中显式预测所添加的噪声。相反，我们在语义先验 $\mathrm{N}_k$ 的指导下，于每个时间步直接生成潜在邻居表示 $\mathbf{s}_N^0$，从而使模型能够以有针对性的方式生成潜在邻居：

$$\mathbf{s}_N^0 = f_\theta\left(\mathbf{s}_L^t, \mathcal{N}_k, t\right), \tag{10}$$

其中，$f_\theta(\cdot)$ 采用与 Mao 等人 [24] 和 Wu 等人 [41]相同的多层感知机架构。在我们的模型中，逆生成过程在每一步都被概念化为条件高斯分布：

$$p_\theta\left(\mathbf{s}_N^{t-1} \mid \mathbf{s}_N^t, \mathcal{N}_k\right) = \mathcal{N}\left(\mathbf{s}_N^{t-1}; \mu_\theta\left(\mathbf{s}_N^t, \mathcal{N}_k, t\right), \Sigma_\theta\left(\mathbf{s}_N^t, \mathcal{N}_k, t\right)\right), \tag{11}$$

其中，$\mu_\theta(\cdot)$ 和 $\Sigma_\theta(\cdot)$ 是网络预测的可学习均值和协方差。因此，扩散损失 $\mathrm{L}_d$ 定义如下：

$$\mathcal{L}_d = \mathbb{E}_{t, \mathbf{s}_L^0, \boldsymbol{\epsilon}}\left[\left\|\mathbf{s}_{id} - f_\theta\left(\sqrt{\bar{\alpha}_t}\mathbf{s}_L^0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, \mathcal{N}_k, t\right)\right\|^2\right] \tag{12}$$

在推理过程中，我们采用确定性策略生成潜在邻居$\mathbf{s}_N^0$。具体而言，我们首先对输入会话嵌入$\mathbf{s}_{id}$执行 $T'$ 步前向扰动，以获得带噪初始化。然后，从这一状态出发，在检索到的邻居先验 $\mathrm{N}_k$的条件下执行 $T'$ 步反向去噪。在确定性反向过程中，省略方差，并直接使用预测的均值：

$$\mathbf{s}_N^{t-1} = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t} f_\theta\left(\mathbf{s}_N^t, \mathcal{N}_k, t\right) + \frac{\sqrt{\alpha_t}\left(1-\bar{\alpha}_{t-1}\right)}{1-\bar{\alpha}_t}\mathbf{s}_N^t, \tag{13}$$

其中，模型 $f_\theta(\cdot)$需要学习针对每个输入的会话特定去噪策略。最终输出$\mathbf{s}_N^0$ 随后用作生成的潜在邻居表示。

*4.2.3* 反馈驱动的检索器优化。我们的方法旨在使检索器能够持续从生成器的反馈中学习，从而检索出对邻居生成任务更有益的邻居会话。为此，我们评估每个检索到的邻居在引导扩散过程方面的有效性。具体而言，给定目标会话的基于 ID 的表示 $\mathbf{s}_{id}$ 和一组检索到的邻居 $\mathrm{N}_k$，我们通过第 4.2.2 节定义的扩散过程计算生成器的损失 $\mathrm{L}_d$。$\mathrm{L}_d$ 的值越小，表明相应邻居对潜在邻居生成的贡献越大。受 Lu 和 Liu [21], 的启发，我们引入一种基于相对排序的监督机制：对于两个检索到的邻居$\mathbf{s}_i^D$、$\mathbf{s}_j^D \in \mathrm{N}_k$，如果以$\mathbf{s}_i^D$为条件时的扩散损失低于以$\mathbf{s}_j^D$为条件时的扩散损失（即 $\mathrm{L}_{d_i} < \mathrm{L}^{d_j}$），则认为$\mathbf{s}_i^D$对生成更有用，应被赋予更高的检索分数。为根据生成器的偏好优化检索器，我们提出一种纳入该监督信号的训练策略。考虑到实际应用中候选池规模较大，我们仅对排名前 $k$ 的检索到的邻居进行监督，以提高效率。相应的损失函数 $\mathrm{L}_r$ 定义如下：

$$\mathcal{L}_r = \sum_{\omega_j \in \text{ top-}k} \mathcal{L}_{d_j} \cdot \omega_j, \tag{14}$$

其中 $\omega_j$ 表示检索器计算得到的邻居$\mathbf{s}_j^D$的 Softmax 归一化相关性分数，而 $\mathrm{L}^{d_j}$ 表示将$\mathbf{s}_j^D$用作生成条件时产生的扩散损失。

## 4.3 自增强扩散模块

为在不干扰检索增强扩散的情况下有效地将模态特定信息融入扩散过程，我们设计了一个自增强扩散模块。该模块在检索增强扩散路径与自增强扩散路径之间进行对比学习，从而增强跨模态的语义对齐。

具体而言，我们通过添加高斯噪声，对 ID 表示$\mathbf{s}_{id}$应用前向扩散过程，从而得到时间步 $t$处的含噪表示$\mathbf{s}_L^t$。在反向扩散过程中，与在检索到的邻居引导下执行去噪的检索增强扩散模块不同，自增强扩散模块利用当前输入本身的多模态表示来引导去噪。去噪由网络 $f_\psi(\cdot)$执行，得到去噪后的嵌入$\mathbf{s}_M^t = f_\psi\left(\mathbf{s}_L^t, \mathbf{s}_{mo}, t\right)$。与检索增强扩散模块类似，我们计算自增强模块的扩散损失 $\mathrm{L}_s$，如下所示：

$$\mathcal{L}_s = \mathbb{E}_{t, \mathbf{s}_L^0, \boldsymbol{\epsilon}}\left[\left\|\mathbf{s}_{id} - f_\psi\left(\sqrt{\bar{\alpha}_t}\mathbf{s}_L^0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, \mathbf{s}_{mo}, t\right)\right\|^2\right]. \tag{15}$$

随后，为了将模态引导信号间接注入扩散过程，我们采用对比学习策略来增强模态间的语义对齐。给定时间步 $t$处经去噪的嵌入$\mathbf{s}_N^t$和$\mathbf{s}_M^t$，对比损失定义为

$L_m$：

$$\mathcal{L}_m = -\frac{1}{B}\sum_{i=1}^{B}\log\frac{\exp\left(\mathrm{sim}\left(\mathbf{s}_{N,i}^{t},\mathbf{s}_{M,i}^{t}\right)/\tau\right)}{\sum_{j=1}^{B}\exp\left(\mathrm{sim}\left(\mathbf{s}_{N,i}^{t},\mathbf{s}_{M,j}^{t}\right)\tau\right)}, \tag{16}$$

其中，$B$ 为批次大小，$\tau$ 为温度系数，sim(·，·)用于计算余弦相似度。

## 4.4 预测与模型优化

最终的会话表示定义为原始基于 ID 的表示与生成的邻居表示的加权组合：

$$\mathbf{s}_f = \rho\mathbf{s}_{id} + (1-\rho)\mathbf{s}_N^0, \tag{17}$$

其中，$\rho \in [0,1]$ 是一个可学习参数，用于平衡两个组成部分的贡献。对于每个候选物品 $\mathbf{x}_i^{v_i^{id}}$，预测的点击概率 $\hat{}y_i$ 计算如下：

$$\hat{y}_i = \mathbf{s}_f^{\top}\cdot\mathbf{x}_{v_i^{id}}. \tag{18}$$

推荐的训练目标定义为标准的交叉熵损失 $L_e$，其形式如下：

$$\mathcal{L}_e(y,\hat{y}) = -\sum_{i=1}^{m}\left[y_i\log\left(\hat{y}_i\right) + (1-y_i)\log\left(1-\hat{y}_i\right)\right], \tag{19}$$

其中，$y$ 表示真实物品的独热编码向量，$\hat{}y_i$ 表示第 $i$ 个物品的预测概率。

总体训练目标 L 定义为 $L_e$、$L_r$、$L_s$ 和 $L_m$ 的加权组合：

$$\mathcal{L} = \mathcal{L}_e + \gamma(\mathcal{L}_r + \mathcal{L}_s) + \delta\mathcal{L}_m, \tag{20}$$

其中，$\gamma$ 和 $\delta$ 是分别控制扩散损失和对比损失权重的超参数。

# 5 实验

为展示 DiffSBR 的有效性，我们围绕以下研究问题开展了广泛的实验：

- **RQ1**：与现有的基于会话的推荐方法相比，DiffSBR 的表现如何?
- **RQ2**：所提出的各个组件是否都能提升 DiffSBR 的性能?
- **RQ3**：与从已知数据中检索到的可观测邻居相比，生成的潜在邻居是否更有效?
- **RQ4**：不同的超参数设置会如何影响 DiffSBR 的性能?

## 5.1 实验设置

*5.1.1* 数据集与评估指标。 我们在四个广泛使用的公开数据集上评估模型：**手机**、**体育用品**、**杂货**和 **Instacart**，遵循 [51]。前三个数据集来自亚马逊[1] 的不同类别，并已广泛应用于基于会话的推荐。与 Zhang 等人 [51, 52], 的预处理流程类似，我们将同一天内发生的所有用户交互归为一个会话。Instacart 数据集是 Kaggle[2]发布的一项竞赛数据集。为模拟基于会话的推荐场景，按照 Zhang

[1]http://jmcauley.ucsd.edu/data/amazon/ [2] https://www.kaggle.com/c/instacart-market-basket-analysis

**表 1：数据集统计信息。**

| 数据集 | 手机 | 体育用品 | 杂货 | Instacart |
|---|---|---|---|---|
| #item | 9,091 | 14,650 | 7,286 | 10,009 |
| #interaction | 123,186 | 282,102 | 151,251 | 380,230 |
| #session | 40,344 | 90,492 | 43,648 | 88,022 |
| 平均长度 | 3.05 | 3.12 | 3.47 | 4.32 |

等人 [51], 20，选取长度最短的 % 笔交易。就模态而言，Amazon 数据集为每个物品提供文本和视觉信息，而 Instacart 仅使用文本数据。具体来说，文本模态包括物品标题和品牌名称。遵循 Li 等人 [17], 、Wu 等人 [40], ，我们移除了长度为 1 的会话以及出现次数少于五次的物品。所有数据集的统计详情见表 1。对于评估指标，我们遵循先前研究 [16, 17, 51] ，采用两种常用的评估指标：P@K（K 位精确率）和 MRR@K（K 位平均倒数排名），其中 K $\in$ {10,20}，以评估模型性能。

*5.1.2* 基线。 为展示 DiffSBR 的有效性，我们将其与一系列具有代表性的基线方法进行比较，包括传统的和最先进的 SBR 模型，以及基于扩散的序列推荐方法：(1)**SKNN** [13] 通过从历史会话中检索相似度较高的会话邻居来预测下一个物品。(2) **NARM** [17] 使用带有注意力机制的 GRU 对用户意图进行建模。(3) **SR-GNN** [40] 利用图神经网络捕获复杂的物品转移关系。(4) **MSGIFSR** [9] 通过共现模式，利用图神经网络从连续片段中捕获用户偏好。(5) **Atten-Mixer** [48] 利用多层次用户意图，对物品共现转移执行多阶段推理。(6)**MSGAT** [29]通过基于余弦相似度和共现关系检索邻居来增强当前会话。(7) **MGS** [16] 利用物品属性检索相似邻居，并进一步估计用户偏好。(8) **MMSBR** [50] 是 SBR 中首个结合文本和图像来建模用户意图的方法。(9) **DIMO** [51] 揭示共现物品与模态之间的关系，从而解耦 ID 和模态的影响。(10) **DiffuRec** [18] 采用扩散模型处理序列推荐，以概率表示替代传统的静态物品嵌入。(11) **DiQDiff** [24] 使用量化用户序列作为条件，引导序列推荐中的扩散生成。

*5.1.3* 实现细节。 遵循先前研究 [50, 51], ，我们采用 Adam 优化器，初始学习率设为 0.001，小批量大小设为 50。为确保公平比较，所有方法的嵌入维度均设为 100。遵循 Zhang 等人 [51], ，我们使用主成分分析将两种模态的维度均降至 100。我们通过网格搜索选取模型的最佳超参数。图卷积网络层数设为 3，温度系数设为 0.3。对于去噪扩散概率模型，我们遵循 Mao 等人 [24] ，使用 32 个扩散时间步，并采用截断线性噪声调度。

## 5.2 总体性能（RQ1）

表2报告了性能比较的评估结果。结果揭示了以下几点重要发现：

**表2：不同模型在各数据集和指标上的比较。最佳基线结果以下划线标出。* 表示相较于所有基线均有统计显著提升（$p$-值 $< 0.05$）。**

| 模型 | 手机 | | | | 体育用品 | | | | 杂货 | | | | Instacart | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | P@10 | MRR@10 | P@20 | MRR@20 | P@10 | MRR@10 | P@20 | MRR@20 | P@10 | MRR@10 | P@20 | MRR@20 | P@10 | MRR@10 | P@20 | MRR@20 |
| SKNN | 14.31 | 8.84 | 16.48 | 9.06 | 31.79 | 24.23 | 33.98 | 24.39 | 40.40 | 27.64 | 42.40 | 27.78 | 6.78 | 2.06 | 11.79 | 2.41 |
| NARM | 15.42 | 12.43 | 16.80 | 12.53 | 35.55 | 33.40 | 36.67 | 33.57 | 45.67 | 40.39 | 47.14 | 40.59 | 8.27 | 3.02 | 12.19 | 3.25 |
| SR-GNN | 16.36 | 12.96 | 18.11 | 13.09 | 36.31 | 33.36 | 37.69 | 33.66 | 44.33 | 39.44 | 46.24 | 39.64 | 8.96 | 3.27 | 13.00 | 3.64 |
| MSGIFSR | 17.80 | 12.40 | 21.16 | 12.64 | 36.27 | 30.36 | 39.65 | 30.59 | 45.45 | 38.16 | 48.15 | 38.35 | 11.56 | 3.74 | 16.44 | 4.02 |
| Atten-Mixer | 19.51 | 14.54 | 22.28 | 14.71 | 37.30 | 33.63 | 39.19 | 33.86 | 47.65 | 40.71 | 49.56 | 40.84 | 8.11 | 3.12 | 11.53 | 3.36 |
| MSGAT | 17.22 | 13.41 | 20.01 | 13.67 | 37.19 | 33.69 | 38.53 | 33.91 | 45.20 | 39.98 | 47.01 | 40.12 | 9.29 | 3.54 | 13.36 | 3.77 |
| MGS | 21.54 | 14.24 | 25.02 | 14.48 | 36.79 | 32.39 | 38.45 | 32.50 | 46.59 | 38.83 | 48.37 | 38.98 | 8.95 | 2.87 | 13.74 | 3.09 |
| MMSBR | 20.59 | 13.94 | 22.82 | 14.13 | 36.69 | 32.52 | 38.29 | 32.73 | 46.05 | 39.01 | 47.89 | 39.23 | 9.89 | 3.61 | 14.37 | 3.84 |
| DIMO | 31.66 | 16.98 | 38.81 | 17.36 | 45.07 | 34.86 | 49.86 | 35.15 | 53.03 | 41.81 | 57.01 | 41.98 | 12.51 | 4.31 | 18.36 | 4.81 |
| DiffuRec | 25.78 | 15.54 | 30.68 | 15.88 | 41.25 | 33.86 | 47.31 | 34.15 | 50.78 | 38.22 | 54.39 | 38.47 | 9.36 | 3.58 | 13.81 | 3.88 |
| DiQDiff | 28.12 | 16.41 | 33.19 | 16.83 | 43.01 | 34.51 | 49.28 | 34.96 | 52.29 | 39.04 | 56.03 | 39.32 | 10.62 | 3.96 | 15.34 | 4.17 |
| DiffSBR | **37.28*** | **18.03*** | **45.97*** | **18.60*** | **49.94*** | **35.66*** | **55.83*** | **36.07*** | **57.05*** | **42.64*** | **62.33*** | **42.84*** | **13.83*** | **5.12*** | **20.34*** | **5.57*** |
| 提升 ↑ | 17.75% | 6.18% | 18.45% | 7.14% | 10.81% | 2.29% | 11.97% | 2.62% | 7.58% | 1.99% | 9.33% | 2.05% | 10.55% | 18.79% | 10.78% | 15.80% |

（1）SKNN 等早期方法仅依赖会话相似度进行邻居检索，却无法捕捉会话内复杂的物品转移，因此推荐准确率有限。NARM 和 SR-GNN 等后续模型引入注意力机制和图神经网络机制，以增强会话建模能力，从而提升性能。MSGIFSR、Atten-Mixer、MSGAT 和 MGS 等近期方法尝试将邻居检索与基于图的建模相结合。这些模型将历史上相似的会话作为辅助信息，以丰富会话表示，进而提升推荐性能。然而，它们可能仅依赖观测到的邻居，难以捕捉未观测到但在语义上相关的会话。相比之下，DiffSBR 引入了检索增强扩散模块，将检索到的邻居视为语义先验，以引导潜在邻居的生成，从而取得更高的性能。

(2) 我们观察到，MMSBR 和 DIMO 表现良好，这表明引入多模态信息对提升会话建模效果发挥了积极作用。DiffuRec 和 DiQDiff 等基于扩散的模型也展现出良好的效果，证明了扩散模型在重建复杂分布和实现条件控制生成方面的优势。相比之下，所提出的 DiffSBR 进一步提升了推荐性能。DiffSBR 通过在基于扩散的生成过程中持续注入先验知识和多模态信号来生成潜在邻居，从而探索观测数据之外的完整兴趣空间。

(3) DiffSBR 在多个数据集上始终优于所有基线，证明了其有效性。这一性能提升归因于潜在邻居的重要作用，以及检索增强和自增强扩散模块的有效性。通过利用检索到的真实邻居作为先验知识来引导生成过程，并注入多模态信号，DiffSBR 能够生成超出观测数据范围的高质量潜在邻居，从而增强会话表示，进而提高推荐准确率。

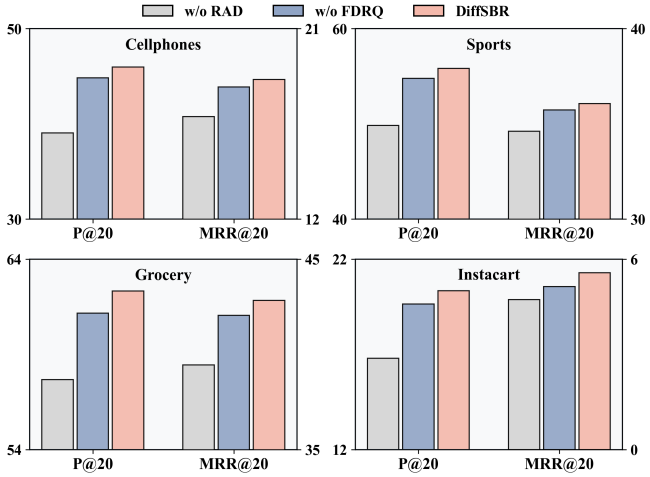



**图3：检索增强扩散模块的效果。**

## 5.3 消融实验（RQ2）

为评估我们方法中各个主要组件的影响，我们通过有选择地移除或调整核心模块进行消融实验。

*5.3.1* 检索增强扩散模块的效果。为验证检索增强扩散模块的效果，我们比较了以下变体：**w/o RAD**：此变体移除了检索增强扩散模块，这意味着不会使用检索到的先验信息来引导扩散过程，也不会通过反馈优化检索器，且不会融入多模态信息。**w/o FDRQ**：在此变体中，检索增强扩散模块仅移除了反馈驱动的检索器优化。四个基准数据集上的性能变化如图3所示。

w/o RAD 变体导致性能显著下降，这表明先验知识对于引导扩散过程生成语义对齐的邻居表示至关重要。缺少有效引导时，扩散过程更容易在潜在空间中偏离，导致推荐质量下降。此外，w/o FDRQ 变体削弱了协作

**表3：自增强扩散模块的效果。**

| 方法 | 手机 | | 体育用品 | | 杂货 | | Instacart | |
|---|---|---|---|---|---|---|---|---|
| | P@20 | MRR@20 | P@20 | MRR@20 | P@20 | MRR@20 | P@20 | MRR@20 |
| 无 SAD | 43.19 | 17.76 | 53.90 | 35.43 | 60.70 | 41.73 | 18.55 | 4.97 |
| -EF | 40.20 | 17.15 | 50.41 | 34.58 | 58.83 | 39.18 | 15.85 | 4.54 |
| -LF | 43.56 | 17.95 | 52.69 | 34.77 | 61.07 | 41.51 | 19.15 | 4.77 |
| -CF | 39.98 | 17.32 | 51.09 | 35.08 | 59.08 | 40.53 | 17.56 | 4.79 |
| -AF | 44.19 | 17.68 | 53.82 | 34.83 | 61.40 | 40.84 | 19.89 | 5.30 |
| **DiffSBR** | **45.97*** | **18.60*** | **55.83*** | **36.07*** | **62.33*** | **42.84*** | **20.34*** | **5.57*** |

检索器与生成器之间的协作。缺少来自生成器的动态监督，检索器无法自适应地优化其评分函数，以更好地支持生成过程，从而限制了检索到的引导信号的质量并降低了性能。

*5.3.2* 自增强扩散模块的作用。我们设计了以下变体：**w/o SAD**：该变体移除了自增强扩散模块，这意味着整个框架不再利用多模态信息，仅使用 ID 模态。**-EF**：该变体移除了自增强扩散模块，并采用早期融合策略，将模态嵌入直接整合到检索增强扩散模块的输入中。**-LF**：该变体移除了自增强扩散模块，并在预测阶段融入模态嵌入，以实现后期融合。**-CF**：该变体移除了自增强扩散模块，并采用条件融合，将模态嵌入直接合并到检索增强扩散模块的条件中，再用该条件引导扩散去噪过程。**-AF**：该变体保留自增强扩散模块，但将对比学习替换为跨模态注意力机制，以整合模态信息。

如表3所示，对于 w/o SAD 变体，移除该变体后性能显著下降，这表明将多模态信息加入扩散过程中有助于提升潜在邻居的质量。我们进一步评估了四种简化的替代方案（EF、LF、CF 和 AF），它们分别通过不同的融合策略整合多模态信息。尽管这些变体在一定程度上纳入了多模态信息，但与

完整模型相比，它们的性能均明显下降，这主要是因为其融合机制过于粗略，或与扩散去噪过程的对齐不足。因此，它们无法提供有效的多模态信号，导致潜在邻居质量较低，模型性能也随之下降。这一结果表明，简单的融合策略不足以有效利用多模态信号。

## 5.4 生成的潜在邻居的有效性（RQ3）

为了全面评估潜在邻居生成的必要性和有效性，我们进行了定量和定性分析。具体而言，我们比较了两个变体：可观测邻居，它根据相似度直接从观测数据中检索邻居（体现了传统的基于检索的方法）；以及潜在邻居，它通过我们提出的 DiffSBR 方法生成潜在邻居。

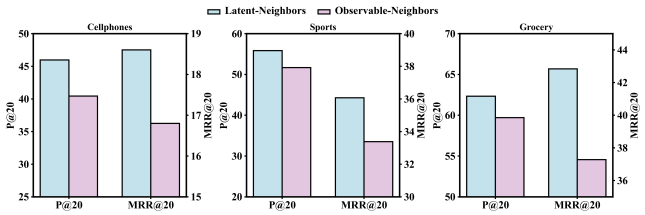
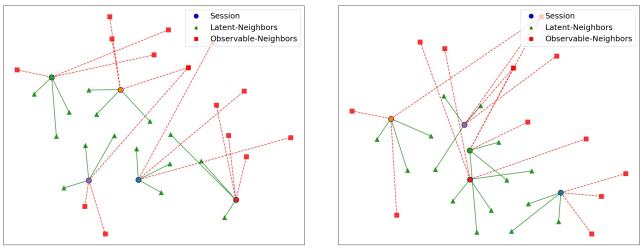



**图4：潜在邻居与可观测邻居的比较结果。**




（a）手机 （b）杂货

**图5：潜在邻居和可观测邻居的定性可视化。**

**定量评估。** 为了评估潜在邻居生成的有效性，我们在三个基准数据集上系统地比较了潜在邻居和可观测邻居这两个变体。如图 4[3], 所示，潜在邻居在三个数据集上的 P@20 和 MRR@20 指标均始终优于可观测邻居。这些提升凸显了我们提出的基于扩散的生成方法的优势，该方法能够合成与语义相关的邻居，突破观测交互数据的局限。潜在邻居不仅弥补了静态检索方法留下的空缺，提升了整体推荐性能，还为生成式邻居建模的可行性提供了数据驱动的证据。

**定性可视化。** 为了直观地理解潜在邻居为何优于静态检索得到的邻居，我们通过可视化分析比较它们的空间分布。具体而言，我们使用 t-SNE[32, 33] 将目标会话及其对应邻居的高维表示投影到二维空间中。我们用蓝色圆点表示不同的会话。图中的红色方块表示可观测邻居，即从已知数据集中检索到的相似度最高的三个邻居。图中的绿色三角形表示使用所提出的方法生成的潜在邻居。

如图 5[3], 所示，潜在邻居（以绿色三角形表示）在嵌入空间中的分布更加集中，紧邻目标会话节点。相比之下，可观测邻居（以红色方块表示）的分布相对分散，其中一些距离目标会话较远。这一现象表明，尽管传统检索方法能够从观测数据中识别出相对相似的邻居，但检索到的邻居通常局限于记录行为的范围内

[3] 值得注意的是，我们在四个数据集上进行了实验。观察到的模式与图中所示一致。为了使可视化更清晰、呈现效果更美观，我们仅展示了部分具有代表性的结果，而非偶然结果。

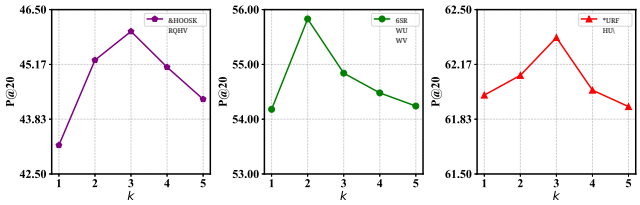


**图 6：检索到的邻居数量的影响 $k$。**

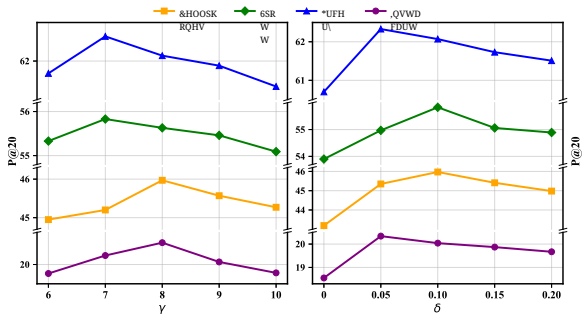


**图 7：损失权重 $\gamma$ 和 $\delta$的影响。**

。相反，潜在邻居能够有效填补嵌入空间中的稀疏区域，从而发现与目标会话在语义上更加接近的邻居。这些定性结果证实了我们提出的潜在邻居生成方法的必要性和有效性。

## 5.5 敏感性分析（RQ4）

*5.5.1* 检索到的邻居数量的影响 $k$。检索到的邻居数量 $k$作为引导潜在邻居生成的先验知识，是一个关键超参数。为探究其影响，我们将 $k$ 设为{1、2、3、4、5}，并在各数据集上评估模型性能。如图 6 所示，随着 $k$ 增加，模型性能起初有所提升，因为更多富含信息的邻居为生成器提供了更丰富的语义先验信号。然而，体育用品数据集上的性能在 $k = 2$ 时达到峰值，手机和杂货数据集上的性能则在 $k = 3$ 时达到峰值；超过这一数量后，进一步增加 $k$ 会导致性能略微下降。这一趋势表明，引入过多邻居可能会带来噪声或冗余信息，最终削弱引导效果。

*5.5.2* 损失权重的影响 $\gamma$ 和 $\delta$。我们进一步研究了两个损失权重的影响：扩散模型损失权重 $\gamma$ 和对比损失权重 $\delta$。如图 7 所示，对于扩散损失权重 $\gamma$，我们分析了 $\gamma = 6$ 的结果，因为初步实验表明，较小的值（例如，$\gamma < 6$）会导致性能欠佳，这是由于扩散损失对生成过程的影响不足。具体而言，体育用品和杂货数据集在 $\gamma = 7$时性能达到峰值，而手机和 Instacart 数据集在 $\gamma = 8$时表现最佳。这表明，恰当设置扩散模型损失的权重，能够使模型有效生成潜在邻居。然而，将 $\gamma$ 设得过高（例如，$\gamma = 9$ 或 10）会导致性能略有下降，这可能是因为扩散损失项主导了总体目标，削弱了其他重要组成部分的作用。

对于对比损失权重 $\delta$，适度的对比信号能够提升性能。具体而言，杂货和 Instacart 数据集在 $\delta = 0.05$时性能最佳，而手机和体育用品数据集则在 $\delta = 0.1$时达到峰值。这证实了在扩散过程中注入多模态信息的有效性。然而，过大的 $\delta$值可能会干扰其他损失组成部分的优化，导致性能略有下降。

## 6 结论

本研究提出了一种新颖的基于扩散的潜在邻居生成模型 DiffSBR，以改进基于会话的推荐。具体而言，我们设计了一个检索增强扩散模块，利用检索到的邻居作为先验知识，引导扩散过程生成潜在邻居。在该模块中，我们还引入了一种新的训练策略，以增强邻居生成过程中检索与扩散之间的协同作用。此外，我们开发了一个自增强扩散模块，以充分利用整个扩散过程中的多模态信息，从而提高生成邻居的质量。在四个基准数据集上的实验结果表明，DiffSBR 的性能始终显著优于最先进方法。

在本研究中，我们强调了潜在邻居生成在基于会话的推荐中的潜力。然而，这仍是对潜在邻居建模的初步研究，仍有改进空间。未来，我们计划研究各种生成模型（例如自回归模型）各自的优势及其互补性，并探索整合多种生成范式的混合框架，以进一步提升生成的潜在邻居的质量。

## 致谢

本研究得到国家自然科学基金（62402093）和四川省科技计划（2025ZNSFSC0479）的资助。

## 参考文献

[1] Guojia An、Jing Sun、Yuhan Yang 和 Fuming Sun。2024年。通过对比学习增强用于基于会话的推荐的协作信息。*Information Processing & Management (IPM)* 61（2024年），103738。[2] Guojia An、Jie Zou、Jiwei Wei、Chaoning Zhang、Fuming Sun 和 Yang Yang。2025年。超越完整对话建模：面向对话式推荐的上下文解耦。载于*SIGIR*。31–41。[3] Xiao Ao、Shuxi Han、Yeming Li、Heli Ma、Pengfei Zhang 和 Jie Zou。2025年。检索增强的多智能体推荐系统。载于*WWW*。2968–2972。[4] Jinpeng Chen、Jianxiang He、Huan Li、Senzhang Wang、Yuan Cao、Kaimin Wei、Zhenye Yang 和 Ye Ji。2025年。结合可插拔大语言模型驱动语义的层次化意图引导优化：面向基于会话的推荐。载于*SIGIR*。1655–1665。[5] Jacob Devlin、Ming-Wei Chang、Kenton Lee 和 Kristina Toutanova。2019年。BERT：用于语言理解的深度双向变换器预训练。载于*NAACL-HLT*。4171–4186。[6] Prafulla Dhariwal 和 Alexander Nichol。2021年。扩散模型在图像合成方面胜过生成对抗网络。载于*NeurIPS*。8780–8794。[7] Lianli Gao、Zhao Guo、Hanwang Zhang、Xing Xu 和 Heng Tao Shen。2017年。利用基于注意力的长短期记忆网络和语义一致性生成视频描述。*Transactions on Multimedia (TMM)* 19卷9期（2017年），2045–2055。[8] Ian J Goodfellow、Jean Pouget-Abadie、Mehdi Mirza、Bing Xu、David Warde-Farley、Sherjil Ozair、Aaron Courville 和 Yoshua Bengio。2014年。生成对抗网络。载于*NeurIPS*。1–9。[9] Jiayan Guo、Yaming Yang、Xiangchen Song、Yuan Zhang、Yujing Wang、Jing Bai 和 Yan Zhang。2022年。学习用于基于会话的推荐的多粒度连续用户意图单元。载于*WSDM*。343–352。

[10] Muskan Gupta、Priyanka Gupta 和 Lovekesh Vig。2024年。用于鲁棒会话式推荐的引导式扩散反事实增强。arXiv预印本 arXiv:2410.21892（2024年）。[11] Balázs Hidasi、Alexandros Karatzoglou、Linas Baltrunas 和 Domonkos Tikk。2015年。基于循环神经网络的会话式推荐。arXiv预印本 arXiv:1511.06939（2015年）。[12] Jonathan Ho、Ajay Jain 和 Pieter Abbeel。2020年。去噪扩散概率模型。载于 NeurIPS。6840–6851。[13] Dietmar Jannach 和 Malte Ludewig。2017年。当循环神经网络遇上用于基于会话的推荐的邻域。载于 RecSys。306–310。[14] Yangqin Jiang、Lianghao Xia、Wei Wei、Da Luo、Kangyi Lin 和 Chao Huang。2024年。DiffMM：用于推荐的多模态扩散模型。载于 MM。7591–7599。[15] Yangqin Jiang、Yuhao Yang、Lianghao Xia 和 Chao Huang。2024年。DiffKG：用于推荐的知识图谱扩散模型。载于 WSDM。313–321。[16] Siqi Lai、Erli Meng、Fan Zhang、Chenliang Li、Bin Wang 和 Aixin Sun。2022年。用于基于会话的推荐的属性驱动镜像图网络。载于 SIGIR。1674–1683。[17] Jing Li、Pengjie Ren、Zhumin Chen、Zhaochun Ren、Tao Lian 和 Jun Ma。2017年。神经注意力会话式推荐。载于 CIKM。1419–1428。[18] Zihao Li、Aixin Sun 和 Chenliang Li。2023年。DiffuRec：用于序列推荐的扩散模型。Transactions on Information Systems (TOIS) 42（2023年），1–28。[19] Zihao Li、Xianzhi Wang、Chao Yang、Lina Yao、Julian McAuley 和 Guandong Xu。2023年。利用显式和隐式物品关系进行基于会话的推荐。载于 WSDM。553–561。[20] Qidong Liu、Fan Yan、Xiangyu Zhao、Zhaocheng Du、Huifeng Guo、Ruiming Tang 和 Feng Tian。2023年。用于序列推荐的扩散增强。载于 CIKM。1576–1586。[21] Hanzhen Lu 和 Zhongxin Liu。2024年。通过检索辅助生成改进检索增强代码注释生成。载于 ICSME。350–362。[22] Mingyang Lv、Xiangfeng Liu 和 Yuanbo Xu。2025年。用于基于会话的推荐的动态多兴趣图神经网络。载于 AAAI。12328–12336。[23] Haokai Ma、Yimeng Yang、Lei Meng、Ruobing Xie 和 Xiangxu Meng。2024年。用于推荐的多模态条件扩散模型。载于 WWW。1733–1740。[24] Wenyu Mao、Shuchang Liu、Haoyang Liu、Haozhe Liu、Xiang Li 和 Lantao Hu。2025年。用于基于扩散的序列推荐的区分量化引导。载于 WWW。425–435。[25] Zhiqiang Pan、Fei Cai、Yanxiang Ling 和 Maarten De Rijke。2020年。用于基于会话的推荐的意图引导协作机器。载于 SIGIR。1833–1836。[26] Yunchen Pu、Zhe Gan、Ricardo Henao、Xin Yuan、Chunyuan Li、Andrew Stevens 和 Lawrence Carin。2016年。用于图像、标签和描述深度学习的变分自编码器。载于 NeurIPS。1–9。[27] Shutong Qiao、Wei Zhou、Fengji Luo 和 Junhao Wen。2023年。结合意图目标协同作用的降噪图神经网络，用于基于会话的推荐。Information Processing & Management (IPM) 60（2023年），103517。[28] Shutong Qiao、Wei Zhou、Junhao Wen、Chen Gao、Qun Luo、Peixuan Chen 和 Yong Li。2025年。利用大语言模型进行多视图意图学习与对齐：面向基于会话的推荐。Transactions on Information Systems (TOIS) 43（2025年），1–25。[29] Shutong Qiao、Wei Zhou、Junhao Wen、Hongyu Zhang 和 Min Gao。2023年。用于基于会话的推荐的双通道多重稀疏图注意力网络。载于 CIKM。2075–2084。[30] Ruihong Qiu、Zi Huang、Jingjing Li 和 Hongzhi Yin。2020年。利用图神经网络中的跨会话信息进行基于会话的推荐。Transactions on Information Systems (TOIS) 38（2020年），1–23。[31] Christian Szegedy、Wei Liu、Yangqing Jia、Pierre Sermanet、Scott Reed、Dragomir Anguelov、Dumitru Erhan、Vincent Vanhoucke 和 Andrew Rabinovich。2015年。借助卷积网络更进一步。载于 CVPR。1–9。[32] Laurens Van der Maaten 和 Geoffrey Hinton。2008年。使用 t-SNE 可视化数据。Journal of Machine Learning Research (JMLR) 9卷11期（2008年）。[33] Bokun Wang、Yang Yang、Xing Xu、Alan Hanjalic 和 Heng Tao Shen。2017年。对抗性跨模态检索。载于 MM。154–162。[34] Wenjie Wang、Yiyan Xu、Fuli Feng、Xinyu Lin、Xiangnan He 和 Tat-Seng Chua。2023年。扩散推荐模型。载于 SIGIR。832–841。[35] Zheng Wang、Zhenwei Gao、Yang Yang、Guoqing Wang、Chengbo Jiao 和 Heng Tao Shen。2024年。跨模态检索的几何匹配。《神经网络与学习系统汇刊》（TNNLS）36卷3期（2024年），5509–5521。[36] Zhendong Wang、Yifan Jiang、Huangjie Zheng、Peihao Wang、Pengcheng He、Zhangyang Wang、Weizhu Chen、Mingyuan Zhou 等。2023年。分块扩散：更快且更具数据效率的扩散模型训练。载于 NeurIPS。72137–72154。[37] Ziyang Wang、Wei Wei、Gao Cong、Xiao-Li Li、Xian-Ling Mao 和 Minghui Qiu。2020年。用于基于会话的推荐的全局上下文增强图神经网络。载于 SIGIR。169–178。[38] Yibiao Wei、Jie Zou、Weikang Guo、Guoqing Wang、Xing Xu 和 Yang Yang。2025年。MSCRS：面向对话式推荐系统的多模态语义图提示学习框架。载于 SIGIR。42–52。[39] Junda Wu、Cheng-Chun Chang、Tong Yu、Zhankui He、Jianing Wang、Yupeng Hou 和 Julian McAuley。2024年。Coral：协同检索增强大语言模型提升长尾推荐效果。载于 SIGKDD。3391–3401。[40] Shu Wu、Yuyuan Tang、Yanqiao Zhu、Liang Wang、Xing Xie 和 Tieniu Tan。2019年。基于图神经网络的会话推荐。载于 AAAI。346–353。[41] Zihao Wu、Xin Wang、Hong Chen、Kaidong Li、Yi Han、Lifeng Sun 和 Wenwu Zhu。2023年。Diff4rec：通过课程调度扩散增强实现序列推荐。载于 MM。9329–9335。[42] Xin Xia、Hongzhi Yin、Junliang Yu、Yingxia Shao 和 Lizhen Cui。2021年。面向基于会话的推荐的自监督图协同训练。载于 CIKM。2180–2190。[43] Xin Xia、Hongzhi Yin、Junliang Yu、Qinyong Wang、Lizhen Cui 和 Xiangliang Zhang。2021年。面向基于会话的推荐的自监督超图卷积网络。载于 AAAI。4503–4511。[44] Ling Yang、Zhilong Zhang、Yang Song、Shenda Hong、Runsheng Xu、Yue Zhao、Wentao Zhang、Bin Cui 和 Ming-Hsuan Yang。2023年。扩散模型：方法与应用综述。《计算调查》（CSUR）56卷（2023年），1–39。[45] Zhengyi Yang、Jiancan Wu、Zhicai Wang、Xiang Wang、Yancheng Yuan 和 Xiang-nan He。2023年。生成你偏好的内容：通过引导式扩散重塑序列推荐。载于 NeurIPS。24247–24261。[46] Feng Yu、Qiang Liu、Shu Wu、Liang Wang 和 Tieniu Tan。2016年。用于下一购物篮推荐的动态循环模型。载于 SIGIR。729–732。[47] Feng Yu、Yanqiao Zhu、Qiang Liu、Shu Wu、Liang Wang 和 Tieniu Tan。2020年。TAGNN：面向基于会话的推荐的目标注意力图神经网络。载于 SIGIR。1921–1924。[48] Peiyan Zhang、Jiayan Guo、Chaozhuo Li、Yueqi Xie、Jae Boum Kim、Yan Zhang、Xing Xie、Haohan Wang 和 Sunghun Kim。2023年。通过注意力混合网络高效利用多层次用户意图进行基于会话的推荐。载于 WSDM。168–176。[49] Xiaokun Zhang、Hongfei Lin、Bo Xu、Chenliang Li、Yuan Lin、Haifeng Liu 和 Fenglong Ma。2022年。面向基于会话的推荐的动态意图感知迭代去噪网络。《Information Processing & Management》（IPM）59卷（2022年），102936。[50] Xiaokun Zhang、Bo Xu、Fenglong Ma、Chenliang Li、Liang Yang 和 Hongfei Lin。2023年。超越共现：多模态基于会话的推荐。《知识与数据工程汇刊》（TKDE）36卷（2023年），1450–1462。[51] Xiaokun Zhang、Bo Xu、Zhaochun Ren、Xiaochen Wang、Hongfei Lin 和 Fenglong Ma。2024年。解耦ID与模态效应，实现基于会话的推荐。载于 SIGIR。1883–1892。[52] Xiaokun Zhang、Bo Xu、Liang Yang、Chenliang Li、Fenglong Ma、Haifeng Liu 和 Hongfei Lin。2022年。价格确实重要！在基于会话的推荐中建模价格与兴趣偏好。载于 SIGIR。1684–1693。[53] Jujia Zhao、Wang Wenjie、Yiyan Xu、Teng Sun、Fuli Feng 和 Tat-Seng Chua。2024年。去噪扩散推荐模型。载于 SIGIR。1370–1379。[54] Pan Zhiqiang、Cai Fei、Chen Wanyu、Chen Chonghao 和 Chen Honghui。2022年。用于基于会话的推荐的协同图学习。《信息系统汇刊》（TOIS）40卷（2022年），72。[55] Jie Zou、Mohammad Aliannejadi、Evangelos Kanoulas、Shuxi Han、Heli Ma、Zheng Wang、Yang Yang 和 Heng Tao Shen。2025年。PSCon：通过对话进行产品搜索。载于 SIGIR。3659–3669。[56] Jie Zou、Yifan Chen 和 Evangelos Kanoulas。2020年。迈向基于问题的推荐系统。载于 SIGIR。881–890。