

基于邻域增强的图自监督推荐

文茜琳 杨旭华 马钢峰

(浙江工业大学计算机科学与技术学院 杭州 310023)

摘要 基于图自监督对比学习的推荐算法,通过构造辅助任务对用户项目交互图(U-I图)进行扩充,在不同增强视图中提取监督信号来帮助模型学习节点表示,从而提高模型的鲁棒性和泛化能力,是当前的研究热点。然而,此类方法存在一定的局限性:(1)性能依赖于手工设计的增强策略,这些策略会对图的底层结构发生不同程度的改变,从而导致模型在捕捉网络真实特征方面受到限制;(2)未充分挖掘图结构中的信息,现有方法通常仅从增强视图中构造正样本对,忽视了U-I图中的局部结构和语义关系,这可能限制了模型对节点表征的学习能力。为了解决上述问题,我们提出了基于邻域增强的图自监督推荐(NeSR),分别从子图、结构和语义三个方面提取自监督信号,并设计对比学习策略,提升推荐性能。首先,我们引入扩散模型作为一种数据增强操作,通过前向扩散和反向去噪过程构造高质量的无噪点图,并将子图作为对比学习的增强视图,以提供更稳定的全局结构信息,增强模型对节点表示的鲁棒性。其次,在原始U-I图的基础上,我们通过随机游走提取节点的局部结构信息,将游走序列视为用户(或项目)的正样本对,从而生成基于图结构本身的自监督信号。进一步,为挖掘U-I图中用户(或项目)与其邻域内节点的语义关系,我们采用聚类算法为节点分配聚类伪标签,提出一种基于伪标签提示的目标函数,鼓励用户(或项目)与其语义最相似的邻域节点在表征空间中具有高一致性。我们在五个真实数据集上进行了大量实验,结果表明所提方法在性能上具有竞争性的结果。特别地,在Amazon-Kindle数据集上,NeSR较次优模型分别在Recall@20和NDCG@20上提升了11.79%和12.53%。

关键词 推荐;自监督学习;对比学习;扩散模型;协同过滤

中图法分类号 TP18 DOI号 10.11897/SP.J.1016.2025.02278

Neighborhood-Enhanced Graph Self-Supervised Learning for Recommendation

WEN Xi-Lin YANG Xu-Hua MA Gang-Feng

(College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023)

Abstract Graph-based self-supervised contrastive learning has emerged as a prominent research direction in recommendation systems due to its ability to enhance model robustness and generalization, particularly in sparse and noisy user-item (U-I) interaction scenarios. These methods augment the U-I interaction graph by constructing auxiliary self-supervised tasks and extracting contrastive signals from multiple enhanced graphs. The core idea is to enforce consistency between different enhanced graphs, thereby enabling the model to learn high-quality node representations. Nevertheless, despite their effectiveness, current graph-based self-supervised recommendation methods face several limitations. First, their performance heavily depends on manually designed augmentation strategies—such as node dropping, edge removal, or subgraph sampling—which may distort the intrinsic structure of the original U-I graph to varying degrees. These distortions can obscure crucial topological patterns, thereby limiting the model's

收稿日期:2025-01-02;在线发布日期:2025-07-11。本课题得到国家自然科学基金(62176236)、浙江省自然科学基金重点项目(LZ24F030011)资助。文茜琳,博士研究生,中国计算机学会(CCF)学生会员,主要研究领域为推荐算法、扩散模型、对比学习。E-mail:wenxilin@zjut.edu.cn。杨旭华(通信作者),博士,教授,中国计算机学会(CCF)杰出会员,主要研究领域为图机器学习、推荐算法。E-mail:xhyang@zjut.edu.cn。马钢峰,博士,博士后,主要研究领域为图神经网络、推荐系统。

ability to capture authentic graph characteristics. Second, many existing approaches fail to fully exploit the rich structural and semantic information inherently embedded within the graph. In particular, they often construct positive sample pairs solely from enhanced graphs, neglecting the local structure and neighborhood semantics of the original U-I graph. This limitation may hinder the expressiveness and discriminative capacity of the learned representations. To address these challenges, we propose Neighborhood-enhanced graph Self-supervised learning for Recommendation (NeSR)—a novel framework that extracts self-supervised signals from three complementary perspectives: subgraph, structural, and semantic information. These signals are integrated into the corresponding contrastive learning objective to improve recommendation performance in a principled and robust manner. First, we employ a diffusion model as a generative augmentation mechanism. By simulating the forward diffusion and reverse denoising processes, we construct high-quality, noise-free subgraphs. These subgraphs serve as reliable enhanced graphs for contrastive learning, offering stable global structural cues that improve the robustness of node representation learning. Second, we leverage random walks on the original U-I graph to extract local structural information. The resulting walk sequences are treated as contextually relevant positive samples for users or items. This approach enables the model to generate self-supervised signals that are grounded in the inherent topological patterns of the interaction graph, thus yielding more accurate and informative node embeddings. Furthermore, to capture the semantic relationships between nodes and their neighbors, we apply a clustering algorithm to assign pseudo-labels to all nodes. Based on these labels, we design a pseudo-label-guided contrastive objective that encourages higher embedding consistency between a node and its semantically similar neighbors. This semantic augmentation enables the model to better capture latent user intents and item characteristics that may not be explicitly observable from graph topology alone. We conduct extensive experiments on five real-world benchmark datasets to evaluate the effectiveness of NeSR. Experimental results demonstrate that our method achieves competitive performance across a wide range of metrics. Notably, on the Amazon-Kindle dataset, NeSR outperforms the second-best model by 11.79% in Recall@20 and 12.53% in NDCG@20, indicating substantial improvements in top-N recommendation accuracy. In summary, NeSR presents a comprehensive self-supervised learning framework that integrates subgraph generation via diffusion model, local structure modeling via random walks, and semantic enhancement via clustering. These components jointly improve the quality and robustness of node representations, ultimately leading to significantly enhanced recommendation performance.

Keywords recommendation; self-supervised learning; contrastive learning; diffusion model; collaborative filtering

1 引言

目前,基于深度学习的推荐模型百花齐放,广泛应用于电子商务^[1-2]、新闻媒体^[3-4]以及社交网络^[5-6]等场景,使推荐系统成为网络应用中不可缺少的一部分。与传统的协同过滤方法相比,基于深度学习的推荐算法能够更好处理大规模、高维

度的数据,可以有效地捕捉用户和物品之间复杂的关系,例如AutoRec^[7]、ACF^[8]、NCF^[9]以及NGCF^[10]等。虽然这些算法在一些特定场景下取得很好的效果,但深层架构的设计需要大量数据进行训练。由于推荐系统的数据依赖于用户的交互记录,不可避免地面临数据收集成本高和记录稀疏的问题,例如大部分的用户通常只能消费或点击众多项目中的一小部分^[11],因此导致深度推

荐模型的性能受到一定的限制。

自监督学习是一种特殊的无监督学习,通过利用自身数据特性来进行监督信号的生成和学习,在计算机视觉^[12-13]和自然语言处理^[14-15]等领域取得了巨大成功。由于其具有能够从无标注数据中学习节点表征的特质,可以有效缓解模型标注成本高和数据稀疏的问题,吸引了众多学者对自监督推荐算法的关注,特别是对基于图对比学习的自监督推荐范式^[16-18]研究的热情高涨。基于这些工作我们可以清晰了解到对比推荐的基本流程,先采用数据增强操作扩充额外的用户项目交互二部图(U-I图),再通过联合学习框架最大限度地提取不同视图中同一个节点的一致性。然而我们发现基于对比推荐模型仍存在一些局限:(1)模型性能依赖于手工设计的数据增强策略,例如节点/边丢弃^[16]或边修改^[19]等操作。这些操作可能会对图的底层结构造成不同程度扰动(如图1(a)所示),从而限制模型捕捉网络真实特征的能力。(2)未充分挖掘图结构信息。现有方法通常仅从增强视图中构造正样本对,忽视了U-I图中存在的局部结构和语义关系,从而限制了模型对节点表征的学习能力。因此,设计一种无需依赖手工增强策略且有效挖掘图结构局部信息和语义信息的对比学习方法成为当前研究的重点。

针对上述问题,研究者们提出了多种解决方案,其中扩散模型和无图增强方法备受关注。首先,作为一类强大的生成模型,扩散模型能够通过前向扩散和反向去噪过程捕捉给定数据的潜在分布,在生成任务中展现了强大的能力^[20-22]。近年来,一些研究也探索了扩散模型在推荐系统中的应用潜

力^[23-24],例如利用扩散模型进行数据增强,通过其前向扩散和反向去噪过程生成高质量的无噪声图,为对比学习提供了更鲁棒的增强视图,同时避免了手工设计增强策略带来的结构扰动问题。其次,无图增强方法通过直接利用图结构信息来生成自监督信号,从而减少对额外增强视图的依赖。例如基于扰动的对比^[25-26]和基于聚类的对比^[27-29]方法,分别在数据的表示级别(在嵌入空间加入微小的均匀噪声来生成新的表示)或语义级别(利用聚类算法来寻找锚节点的正样本)进行增强,尽管这些方法取得了一定的性能提升,但仍存在以下问题:(1)基于扰动的方法需要对噪声大小进行精细调优,且未充分考虑图结构的局部信息;(2)基于聚类的方法主要关注原型节点的挖掘,而忽略了图结构中最相似节点对模型学习的贡献。图1展示了一些方法的原理及局限性,尽管它们存在不足,但也为无图增强对比学习的研究提供了新的视角。

基于上述分析,本文提出了一种基于邻域增强的图自监督推荐(NeSR),它是一个层次化的自监督学习框架。首先,我们引入扩散模型作为数据增强操作,通过前向扩散和反向去噪过程生成高质量的无噪声图,并将子图作为对比学习的增强视图,以提供更稳定的全局结构信息。其次,我们设计了一个局部结构增强模块,通过随机游走构造节点局部结构,探索节点与其结构邻域之间更可靠的关联关系。同时为了帮助模型发现数据中隐藏的模式和结构,我们设计了一个邻域语义增强模块,采用基于聚类的方法对U-I图的底层语义信息进行建模,并提出基于伪标签提示的对比目标函数,鼓励锚节点与

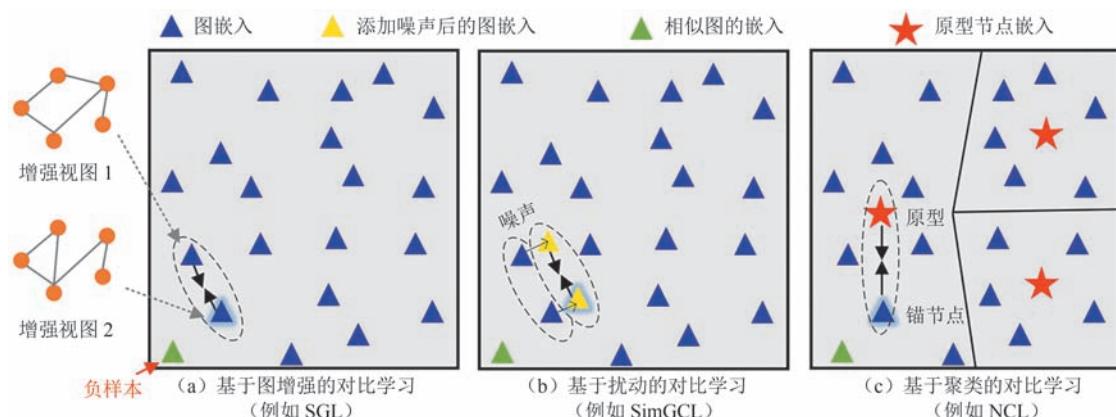


图1 采用不同对比学习方法的推荐算法((子图a)表示采用图增强的对比方法,将同一个图的嵌入视为正对,将其他节点视为负对(包括最相似的节点)。(子图b)和(子图c)表示采用无图增强的对比方法,其中(子图b)表示采用扰动的对比方法,将经过扰动后的节点视为正样本,将其他节点视为负样本(包括最相似的节点)。(子图c)表示基于聚类的对比方法,将锚节点和原型节点视为正对,将其他原型节点视为负对(未考虑最相似的节点))。

其邻域内语义最相似节点表征的一致性。通过在五个真实数据集上的实验验证,NeSR在推荐性能上展现了显著的竞争优势。

本文的主要贡献如下:

(1)提出了一个新颖的基于邻域增强的自监督推荐算法NeSR,分别在子图、结构和语义三个视角下提取用户(或项目)的自监督信号,设计相应的对比学习方案,以改善模型的推荐性能。

(2)引入扩散模型作为一种数据增强操作,通过前向扩散和反向去噪过程生成高质量的无噪声图,并将子图作为对比学习的增强视图,以提供更稳定的全局结构信息。

(3)提出了基于随机游走的局部结构增强方法,通过将用户(或项目)和相应的随机游走产生的游走序列作为锚节点的正对,帮助模型在面对数据噪声和稀疏性时具有更强的鲁棒性。

(4)提出了基于伪标签提示的邻域语义增强方法,先通过聚类算法赋予U-I图中的节点聚类伪标签,然后设计基于伪标签提示的对比目标函数,来鼓励用户(或项目)和其邻域内语义最相似节点的表征具有高一致性。

(5)在五个基准数据集上进行了大量实验,验证了所提算法相较于基准方法的优越性,并通过消融实验验证了不同组件的有效性。

本文第2节从对比学习的图增强和无图增强两个方面的相关研究进行概述。第3节介绍了本文所提出的基于邻域增强的图自监督推荐NeSR算法。第4节对NeSR进行实验,比较其与现有模型之间的优越性。最后在第5节对本文做总结。

2 相关工作

2.1 基于图增强的对比学习

自监督学习在视觉和自然语言的表征学习中很流行^[12-15],其中对比学习为其一大分支^[30],它通过在同一个嵌入空间中对比正样本和负样本来获得判别表征。受局部全局互信息最大化^[31]的启发,DGI^[32]和InfoGraph^[33]首先将图神经网络和对比学习的优势结合起来,在多个节点分类任务中取得具有竞争力的性能,这也为自监督推荐算法提供了新的研究方向。类似基于U-I图的对比推荐,通过设计定制化的图增强技术来生成对比视图,构建代理任务辅助监督任务的学习。此类方法的核心是设计细致的图增强技术,例如SGL^[16]设计了三个算子来生成增

强视图。SSL-DNN^[34]提出了相关特征掩码的模式,将特征集拆分为两个互斥的集合来构建扩展变体。UIR^[35]利用每个用户和项目的邻域信息,对其应用随机扩充来生成不同的视图。GRADE^[36]针对(度分布中)头部和尾部节点应用不同的策略(删除邻居和节点插值)来生成对比视图。AdaMCL^[37]通过Jaccard相似系数来构建额外用户-用户和项目-项目的协同图,设计了自适应融合策略来融合多个图中的嵌入进行对比。LightGCL^[18]利用奇异值分解来指导图的增强,将全局协同关系注入到对比学习的表示一致性中。此外,还有一些算法在U-I图的基础上联合知识图谱来建模用户项目协同知识图,旨在多视角地构建对比学习机制。例如MCCLK^[38]采用KNN来构建局部级项目-项目协同模块,并在多个视图中进行跨视图对比。KACL^[17]提出自适应地去除与任务无关的边的数据扩充方法。考虑到知识图谱中存在大量与特定任务无关的信息,DiffKG^[23]采用扩散模型对知识图谱进行破坏与重构来得到新的视图,并将新视图和U-I图进行整合与原始的协同知识图进行对比。DiffMM^[24]将模态感知图扩散模型与跨模态对比学习范式相结合,以改善模态感知用户表征学习。由于我们所提的算法仅使用了用户项目交互的数据,所以在进行总体性能分析时未涉及此类算法。

2.2 基于无图增强的对比学习

基于无图增强的对比推荐的核心思想是对原始U-I图不依赖增强机制生成额外的图,而是通过原始图结构中已有的信息来获得自监督信号。我们将其分为基于扰动的对比方法和基于聚类的对比方法。基于扰动的对比方法主要是对经过图编码器学习到的表示加入微小噪声,本质上是在数据的嵌入空间中通过扰动操作来获得增强后的数据表示,其中以SimGCL^[25]和XSimGCL^[26]为代表。例如CLDS^[39]受其启发,提出针对社交网络的自对比学习方法,通过加入高斯分布来生成正样本表征以增强原始网络。基于聚类的对比方法通过自监督机制显式建模数据语义关系,例如在GraphLog^[40]中引入了层次原型来捕获全局语义聚类,AFGRL^[41]通过约束来自锚节点的相邻节点或同一簇内的正样本来缩小原始图的语义差距。PGCL^[28]基于正负样本原型间的距离信息,重新对负样本赋予权重来保证正负样本之间的语义差距。上述算法通过设计聚类模块捕获了潜在的语义信息,同时也在分类任务中取得不错的成绩。基于聚类的对比算法NCL^[29],通

过从图结构和语义空间引入用户(或项目)的邻居,并将潜在邻居显式地结合到对比对中。我们的方法和NCL有相通之处,但有以下不同:(1)我们引入扩散模型生成的无噪子图作为可靠的图级对比视图。通过扩散模型的前向扩散和反向去噪过程来生成高质量的子图,不仅保留了原始图结构的关键信息,还减少了噪声对图结构的干扰。(2)在结构增强模块,并未简单使用偶数层的嵌入信息。相反,我们通过随机游走来构造更加可靠的子图结构信息。(3)在语义增强模块中NCL使用节点的原型作为正样本。而我们设计了一个基于伪标签提示的对比目标函数,旨在挖掘与锚节点最为相似的节点作为正样本,利用伪标签提示将负样例限制在与锚节点同一个簇内,通过对比有价值的硬负样例为模型优化提供指导性的梯度信息。

3 方法

在本文中,我们提出基于邻域增强的图自监督推荐算法,该算法通过对原始图的子图、结构和语义三个方面来挖掘自监督信号,能够有效增强用户和项目的表示,从而提升模型的推荐性能。

3.1 问题定义

本研究的目标是利用原始U-I图信息构建对比学习的辅助任务,来增强用户和项目的表征能力,从而提升推荐性能。然而,实现这一目标仍面临以下挑战:(1)如何生成高质量的对比视图,传统的对比方法依赖于手工设计的增强策略来生成对比视图,而这些策略可能会破坏图结构的完整性,导致模型学习到不准确的表征。(2)如何有效识别结构邻域信息,由于U-I图是一个大规模稀疏图,如何精确地定义并获取每个节点的局部结构邻域信息,是一个关键问题。(3)如何挖掘非连接节点之间的邻域语义信息,虽然许多节点之间没有直接连接,但它们可能存在潜在的语义关联。如何在没有显式结构连接的情况下有效捕捉这些语义关联,是提升推荐系统性能的重要挑战。

为了解决上述挑战,我们设计了基于邻域增强的图自监督推荐算法NeSR。该算法利用扩散模型生成的高质量无噪子图作为对比学习的增强视图,子图不仅保留了原始图的关键拓扑特征,还通过去噪过程有效过滤了图中的噪声干扰。其次,通过随机游走的方式,选择锚节点的游走集合,将其作为该节点的局部结构进行对比。同时,我们对经过

warm-up的节点表征执行K-Means聚类,以选择与锚节点最相似的节点作为语义邻居。通过在五个数据集上的系统性实验评估,NeSR在推荐任务中展现了显著的性能提升。这里给出算法的一些符号说明,我们记U-I图为 $\mathcal{G}=(\mathcal{V}, \mathcal{E})$,节点集合 $\mathcal{V}=\mathcal{U} \cup \mathcal{I}$ 由用户集合 $\mathcal{U}=\{u\}$ 和项目集合 $\mathcal{I}=\{i\}$ 组成,边集合为 $\mathcal{E}=\{\mathcal{R}_i^u | u \in \mathcal{U}, i \in \mathcal{I}\}$,若观测到用户 u 和项目 i 之间存在交互行为则 $\mathcal{R}_i^u=1$,否则 $\mathcal{R}_i^u=0$ 。此外我们将每一个用户 $u \in \mathcal{U}$ 的初始状态(step=0)表述为二进制向量 $\mathbf{u}_0^{(u)} = [\mathcal{R}_1^u, \mathcal{R}_2^u, \dots, \mathcal{R}_{|\mathcal{I}|}^u]$ 。特别地,为了方便阅读,当对用户没有区分时,我们将 $\mathbf{u}_0^{(u)}$ 简记为 \mathbf{u}_0 。

3.2 算法框架

NeSR的算法框架如图2所示,它主要由三个模块(三个辅助任务)构成:

(1)扩散子图增强模块,主要采用扩散模型作为数据增强的方法,通过前向扩散和反向去噪过程来生成高质量的无噪子图用于对比。

(2)局部结构增强模块,通过构造节点的局部结构进行对比学习。

(3)邻域语义增强模块,通过聚类算法得到节点的聚类伪标签,在模型对比学习时给予提示信息。

此外,我们采用图编码来学习节点的低维嵌入表示,执行推荐的主任务和辅助任务。

3.3 图编码

近年来,图神经网络^[42-45]旨在将稀疏高维的图结构数据转化为低维稠密的向量表示,已成为图表示学习一种很有前景的方法。为了在U-I图上探索节点的高阶连通性,我们在 \mathcal{G} 上应用了递归消息传播CONST(\cdot)和邻居聚合机制AGGRE(\cdot),即在迭代 l 层消息传播时,用户 $h_u^{(l)}$ 和项目 $h_i^{(l)}$ 的表示定义为

$$\begin{aligned} h_u^{(l)} &= \text{AGGRE}\left(h_u^{(l-1)}, \right. \\ &\quad \left. \text{CONST}\left(\{h_i^{(l-1)} | \forall i \in \mathcal{N}(u)\}\right)\right) \end{aligned} \quad (1)$$

$$\begin{aligned} h_i^{(l)} &= \text{AGGRE}\left(h_i^{(l-1)}, \right. \\ &\quad \left. \text{CONST}\left(\{h_u^{(l-1)} | \forall u \in \mathcal{N}(i)\}\right)\right) \end{aligned} \quad (2)$$

其中,CONST(\cdot)主要是用来提取目标 $u(i)$ 第 $l-1$ 层邻居的嵌入表示,AGGRE(\cdot)是图卷积的核心,它用于更新节点的表示,在很多图编码工作中^[10, 45, 46]都指定了相应的AGGRE(\cdot)。例如,工作[10]采用非线性激活函数聚合节点及邻居信息,工作[45]采

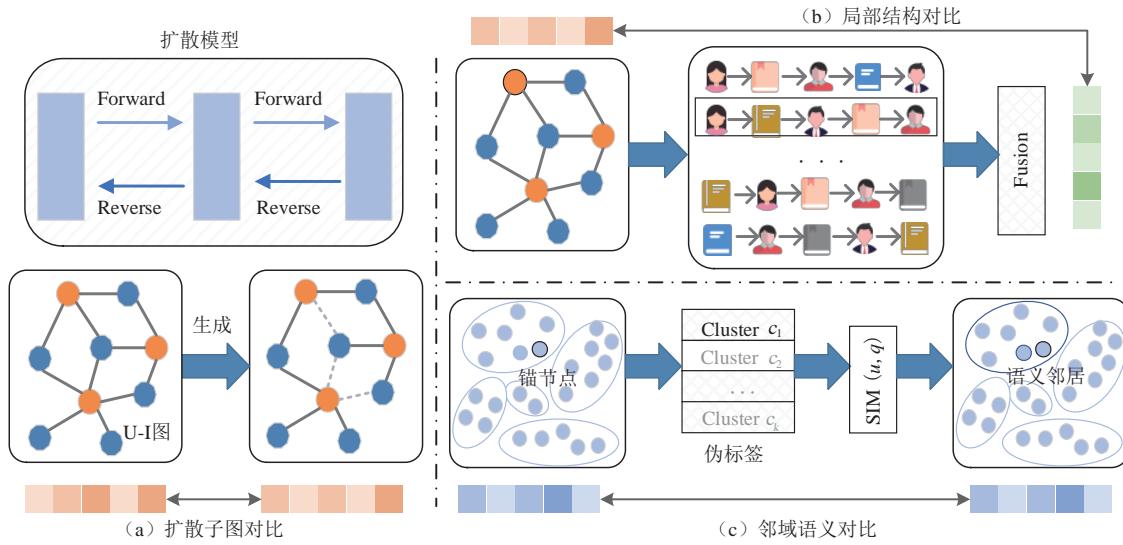


图2 NeSR框架图(图由三部分构成:(a)扩散子图增强模块,以U-I图中的用户初始状态为输入,通过扩散模型的前向扩散和反向去噪过程,构造高质量的无噪声子图。(b)局部结构增强模块,在原始U-I图上进行随机游走,提取节点的局部结构信息。(c)邻域语义增强模块,分别对用户和项目节点进行聚类,获取与锚节点语义最相似的邻域节点。在模型初始化阶段,我们通过扩散子图增强模块和局部结构增强模块进行warm-up训练,以优化节点表示。随后,再训练邻域语义增强模块来提升语义建模能力。)

用轻量级卷积聚合器对邻居进行聚合,工作[46]则采用LSTM聚合器。由于轻量级的卷积操作简化了传统聚合操作繁琐的非线性激活和特征映射操作,所以在这项工作中,我们采用类似的操作进行图编码来获得节点的第 l 层表示,即

$$h_u^{(l)} = \sum_{i \in \mathcal{N}(u)} \frac{1}{\sqrt{|\mathcal{N}(u)|} \sqrt{|\mathcal{N}(i)|}} h_i^{(l-1)} \quad (3)$$

$$h_i^{(l)} = \sum_{u \in \mathcal{N}(i)} \frac{1}{\sqrt{|\mathcal{N}(u)|} \sqrt{|\mathcal{N}(i)|}} h_u^{(l-1)} \quad (4)$$

其中, $|\mathcal{N}(u)|$ 和 $|\mathcal{N}(i)|$ 分别表示 u 和 i 的 1 跳邻居的个数。根据图卷积网络^[42], $1/\sqrt{|\mathcal{N}(u)|} \sqrt{|\mathcal{N}(i)|}$ 为拉普拉斯范数。特别地,当 $l=0$ 时, $h_u^{(0)}$ 和 $h_i^{(0)}$ 初始化为可训练的用户和项目嵌入向量。然后通过读出函数 $\text{READ}(\cdot)$ 来得到用户 u 和项目 i 的最终表示 h_u 和 h_i , 这里使用平均函数来实现 $\text{READ}(\cdot)$:

$$h_u = \text{READ}(h_u^{(0)}, h_u^{(1)}, \dots, h_u^{(L)}) = \sum_{j=0}^L \frac{h_u^{(j)}}{L+1} \quad (5)$$

$$h_i = \text{READ}(h_i^{(0)}, h_i^{(1)}, \dots, h_i^{(L)}) = \sum_{j=0}^L \frac{h_i^{(j)}}{L+1} \quad (6)$$

经过上述卷积操作, 我们不仅用消息传播丰富了初始嵌入, 还通过层数 L 控制了信息传播的范围。并把图编码作为底层模块, 将得到的节点嵌入表示与其他模块共享, 以降低模型复杂度。

3.4 扩散子图增强

为了生成高质量的对比视图并避免传统手工设计增强策略对图结构的扰动, 我们引入了扩散子图增强模块。该模块通过前向扩散和反向去噪过程来捕捉图结构的潜在分布, 并生成保留关键拓扑特征的无噪声图。

3.4.1 扩散模型

在前向扩散过程中, 我们逐步向 U-I 图中用户初始状态 \mathbf{u}_0 引入高斯噪声, 生成一系列噪声逐渐增加的中间状态。这一过程可以表示为

$$q(\mathbf{u}_t | \mathbf{u}_{t-1}) = \mathcal{N}(\mathbf{u}; \sqrt{1 - \beta_t} \mathbf{u}_{t-1}, \beta_t \mathbf{I}) \quad (7)$$

其中, $t \sim U(1, T)$, 在 1 和 扩散步 T 之间均匀采样。 \mathbf{u}_t 表示第 t 步的噪声图状态, \mathbf{I} 是单位矩阵, β_t 是在第 t 扩散步所添加的噪声, 其值由噪声调度器根据预设的噪声规模动态控制。使用再参数化技巧^[47], 可以直接从 \mathbf{u}_0 中采样噪声样本 \mathbf{u}_t :

$$q(\mathbf{u}_t | \mathbf{u}_0) = \mathcal{N}(\mathbf{u}; \sqrt{\bar{a}_t} \mathbf{u}_0, (1 - \bar{a}_t) \mathbf{I}) \quad (8)$$

其中, $\bar{a}_t = \prod_{s=1}^t a_s$, $a_s = 1 - \beta_s$ 。

在反向去噪过程中, 通过对第 T 步的状态分布 $p(\mathbf{u}_T) \sim \mathcal{N}(\mathbf{u}_T; 0, \mathbf{I})$ 迭代去噪来生成目标样本 \mathbf{u}_0 。该过程被定义为优化联合分布 $p_\theta(\mathbf{u}_{0:T})$:

$$p_\theta(\mathbf{u}_{0:T}) = p(\mathbf{u}_T) \prod_{t=1}^T p_\theta(\mathbf{u}_{t-1} | \mathbf{u}_t) \quad (9)$$

$$p_\theta(\mathbf{u}_{t-1}|\mathbf{u}_t) = \mathcal{N}\left(\mathbf{u}_{t-1}; \mu_\theta(\mathbf{u}_t, t), \sum_\theta(\mathbf{u}_t, t)\right) \quad (10)$$

其中, $\mu_\theta(\mathbf{u}_t, t)$ 和 $\sum_\theta(\mathbf{u}_t, t)$ 分别是反向去噪过程中神经网络预测高斯分布的均值和方差。

3.4.2 模型优化

为了训练扩散模型, 我们采用证据下界 (Evidence Lower Bound, ELBO) 作为优化目标。ELBO 的目标是最小化噪声预测误差(即各时间步的 KL 散度), 并保证最终生成数据与真实数据分布的对数似然最大化:

$$\begin{aligned} \mathcal{L}_{ELBO} &= \mathbb{E}_{q(\mathbf{u}_1|\mathbf{u}_0)}[\log p_\theta(\mathbf{u}_0|\mathbf{u}_1)] - \\ &D_{KL}(q(\mathbf{u}_T|\mathbf{u}_0)\|p(\mathbf{u}_T)) - \sum_{t=2}^T \mathbb{E}_{q(\mathbf{u}_t|\mathbf{u}_0)} \\ &\left[D_{KL}(q(\mathbf{u}_{t-1}|\mathbf{u}_t, \mathbf{u}_0)\|p_\theta(\mathbf{u}_{t-1}|\mathbf{u}_t)) \right] \end{aligned} \quad (11)$$

其中, $D_{KL}(\cdot)$ 表示 KL 散度, 用于衡量两个分布之间的差异。通过最大化 ELBO, 能够使生成的无噪子图尽可能接近原始图结构, 同时保留其关键拓扑特征。而工作 [47] 已证明最小化高斯分布 $q(\mathbf{u}_{t-1}|\mathbf{u}_t, \mathbf{u}_0)$ 和 $p_\theta(\mathbf{u}_{t-1}|\mathbf{u}_t)$ 的均值之间的差异可以有效优化模型。因此我们采用以下形式的加权均方误差来训练扩散模型:

$$\mathcal{L}(\theta_{dm}) = \frac{1}{2} \left(\frac{\bar{a}_{t-1}}{1 - \bar{a}_{t-1}} - \frac{\bar{a}_t}{1 - \bar{a}_t} \right) \left[\left\| \hat{\mathbf{u}}_\theta(\mathbf{u}_t, t) - \mathbf{u}_0 \right\|_2^2 \right] \quad (12)$$

其中, 我们将 $\hat{\mathbf{u}}_\theta(\cdot)$ 进行实例化为一个多层次感知机, 该网络包含两个全连接层, 并在每层之间引入 Tanh 激活函数, 其输入为扩散步 t 和 \mathbf{u}_t , 输出为 $\hat{\mathbf{u}}_0$ 。当参数 θ 达到最佳时, $\hat{\mathbf{u}}_\theta(\cdot)$ 可视为 $p_\theta(\cdot)$ 的近似值。

3.4.3 子图生成

与视觉领域随机抽取高斯噪声进行反向去噪的扩散模型相比, 我们采用一种简单的推理策略, 用于 U-I 图中的关系预测, 即在推理阶段, 我们首先逐步破坏原始的图结构数据 \mathbf{u}_0 得到 $\mathbf{u}_{T'}$ 。并设置 $\hat{\mathbf{u}}_T = \mathbf{u}_{T'}$ 为起始步, 使用 $\hat{\mathbf{u}}_{t-1} = \hat{\mathbf{u}}_\theta(\hat{\mathbf{u}}_t, t)$ 进行确定性推理。

随后我们根据 $\hat{\mathbf{u}}_0$ 来重构子图 $\hat{\mathcal{G}}_k$, 即对于每一个用户 u , 从 $\hat{\mathbf{u}}_0$ 中选取与其交互强度最高的前 k 个项目 i 来构建局部子图。这样获得的子图不仅保留了原始图的关键拓扑特征, 还通过扩散模型的去噪过程有效过滤了噪声干扰。因此, 我们将 $\hat{\mathcal{G}}_k$ 作为对比学习的增强视图, 用于提升模型对用户-项目交互关系的表征能力, 以用户侧为例:

$$\mathcal{L}_{dm}^{user} = \sum_{u \in \mathcal{U}} -\log \frac{\exp(h_u^\top \hat{h}_u / \tau)}{\sum_{v \in \mathcal{U}} \exp(h_v^\top \hat{h}_v / \tau)} \quad (13)$$

其中, \hat{h}_u 是在 $\hat{\mathcal{G}}_k$ 上的用户表示(采用第 3.3 节的图编码方法获得), τ 是温度超参数, 用于控制相似性度量的敏感程度。类似地, 可计算得到项目侧的对比损失为 \mathcal{L}_{dm}^{item} , 由此扩散子图增强模块中的总损失为 $\mathcal{L}_{dm} = \mathcal{L}_{dm}^{user} + \mathcal{L}_{dm}^{item}$ 。

3.5 局部结构增强

在复杂网络中, 节点的局部结构信息(如一阶邻域和高阶邻域)对于捕捉用户-项目交互关系至关重要。然而传统的图卷积网络通常依赖于同构层信息^[29] 来生成节点的表示, 这使得学到的节点表示不可避免地会发生平滑现象^[48], 即不同节点的嵌入趋于同质化, 难以保留其独特的局部特性。因此, 我们采用随机游走^[49]的方式遍历图中邻居节点, 从而捕获高阶结构特性。以用户 u 为例, 在用户项目交互二部图 \mathcal{G} 上进行游走, 得到锚节点的游走序列集合 $\text{Walk}(u) = \{w_0, w_1, \dots, w_n | w_i \rightarrow w_{i+1}, w_0 = u\}$, 其中 w_i 表示 u 游走的第 i 个节点, n 为游走的长度, $w_i \rightarrow w_{i+1}$ 表示在 \mathcal{G} 上 i 和 $i+1$ 存在连边, 随后从图编码器中获取游走集合 $\text{Walk}(u)$ 中节点的表示, 并通过融合函数得到节点 u 的局部结构表示:

$$h_u^{sub} = \text{Fusion}(\text{Walk}(u)) \quad (14)$$

其中 $\text{Fusion}(\cdot)$ 可采用均值或线性映射等融合方式, 这里采用均值的方式进行计算。通过多步随机游走, 节点能够间接访问到远邻节点, 不再局限于一阶邻域信息的限制, 同时也为表示学习提供了多样化的上下文信息。我们将这种显式提取节点的结构信息作为自监督信号, 进行用户侧的局部结构的对比学习:

$$\mathcal{L}_{local}^{user} = \sum_{u \in \mathcal{U}} -\log \frac{\exp(h_u^\top h_u^{sub} / \tau)}{\sum_{v \in \mathcal{U}} \exp(h_v^\top h_v^{sub} / \tau)} \quad (15)$$

在局部结构的对比学习中, 我们将节点 u 和其局部结构视为正对, 将 u 和其他局部结构视为负对, 即在同一个嵌入空间中表示为拉近 h_u 和 h_u^{sub} 之间的距离, 推远 h_u 和 h_v^{sub} 之间的距离。这样可以充分利用数据的内部结构信息, 并将结构信息融入对比学习的过程中, 帮助计算节点与其邻域之间更可靠的关系。类似地, 可得到项目侧的对比损失 $\mathcal{L}_{local}^{item}$, 最终得到该模块的总损失为 $\mathcal{L}_{local} = \mathcal{L}_{local}^{user} + \mathcal{L}_{local}^{item}$ 。

3.6 邻域语义增强

传统的对比学习方法大多只考虑了数据之间浅

层的拓扑结构,忽略了图内在语义信息的学习,从而导致无法捕获邻域语义的自监督信息。例如社交网络中,两个没有相互关注的人有非常多的共同好友,这说明他们之间存在着某种潜在的联系或共同的兴趣。由于没有互相关注,在图关系网络中不存在连边,也即传统的对比学习不能通过图结构来获取自监督信号。因此,我们设计了一个邻域语义增强模块,利用伪标签提示进行对比学习。

首先对U-I图的用户(或项目)分别聚类并获取每一个节点的聚类伪标签。在模型初始化阶段,我们通过扩散子图增强模块和局部结构增强模块进行warm-up训练,以优化节点表示。随后,利用K-Means算法对用户(或项目)嵌入进行聚类,将其划分为 K 个簇 $\{c_1, c_2, \dots, c_k\}$,并基于聚类结果为节点分配伪标签。最终,我们依据伪标签提取每个节点的簇标签索引。以用户 u 为例,在其所在的簇中进行查找和 u 最相似的用户 \hat{q} :

$$\text{SIM}(u, q) = \frac{h_u^T h_q}{\|h_u\| \|h_q\|} \quad (16)$$

$$\hat{q} = \arg \max_{q \in c_u} \text{SIM}(u, q) \quad (17)$$

其中, c_u 表示用户 u 所在簇的伪标签,通过对簇中的用户进行遍历,找到得分最高的用户 \hat{q} ,将其视为潜在的语义最相似节点进行对比学习,即构造 (u, \hat{q}) 为正样本对。通常将 u 和批次内的其他用户视为负样本对,但我们认为这种方法过于粗糙:(1)未充分利用伪标签信息。(2)无法区分硬负样例。由此我们设计了基于伪标签提示的对比目标函数:

$$\mathcal{L}_{\text{sema}}^{\text{user}} = \sum_{u \in \mathcal{U}} -\log \frac{\exp(h_u^T h_{\hat{q}} / \tau)}{\sum_{q \in c_u} \exp(h_u^T h_q / \tau)} \quad (18)$$

关于负样本的选择,我们提示模型筛选与用户 u 具有相同簇标签的其他节点作为负样本。这样不仅扩展了使用有限标记数据进行分类的能力(使用伪标签数据),提高节点的自一致性^[50];还能更好地捕获簇内节点的差异,帮助K-Means分类器学习到更准确的边界。类似地,项目侧邻域语义的对比损失为 $\mathcal{L}_{\text{sema}}^{\text{item}}$,最终得到邻域语义增强模块总损失为 $\mathcal{L}_{\text{sema}} = \mathcal{L}_{\text{sema}}^{\text{user}} + \mathcal{L}_{\text{sema}}^{\text{item}}$ 。

3.7 多任务联合优化

我们对图编码阶段学习到的用户项目嵌入(h_u 和 h_i)采用成对学习的BPR损失函数^[51]进行模型优化:

$$\mathcal{L}_{\text{BPR}} = \sum_{(u, i) \in \mathcal{E}^+, (u, j) \in \mathcal{E}^-} -\log [\text{sigmoid}(s(u, i) - s(u, j))] \quad (19)$$

其中 \mathcal{E}^+ 和 \mathcal{E}^- 分别表示边集合条目为 $\mathcal{R}_{(u, i)} = 1$ 和 $\mathcal{R}_{(u, i)} = 0$ 的节点集合, $s(u, i) = h_u^T h_i$ 为用户 u 和项目 i 的预测得分,通过BPR优化模型可以最大化训练集中已观测到用户对项目偏好的概率。

随后,将扩散子图损失 \mathcal{L}_{dm} 、局部结构损失 $\mathcal{L}_{\text{local}}$ 和邻域语义损失 $\mathcal{L}_{\text{sema}}$ 融入推荐主任务 \mathcal{L}_{BPR} 中进行联合优化,得到NeSR的损失目标函数:

$$\mathcal{L} = \mathcal{L}_{\text{BPR}} + \lambda_1 \mathcal{L}_{\text{dm}} + \lambda_2 \mathcal{L}_{\text{local}} + \lambda_3 \mathcal{L}_{\text{sema}} + \lambda_4 \|\theta\|_2^2 \quad (20)$$

其中, λ_1, λ_2 和 λ_3 分别是控制扩散子图、局部结构和邻域语义的自监督信号强度, λ_4 是用来控制参数 θ 的 l_2 正则化强度的系数。通过对 \mathcal{L} 进行优化,使模型学习到的节点嵌入更能体现与其邻域之间可靠的关系。

3.8 复杂度分析

我们记 $|\mathcal{E}|$ 为U-I图中边的数量, $|\tilde{\mathcal{E}}|$ 为最大簇的节点数, $|\mathcal{U}|$ 和 $|\mathcal{I}|$ 分别表示用户和项目的总个数, L 为卷积层数, d 为用户项目的嵌入维度, B 是批次大小, n 为随机游走长度, K 为簇的大小, l 为K-Means迭代数。特别地,在扩散模型中, d_l 为反向去噪过程中多层感知机第 l 层的维度。NeSR的时间复杂度由以下部分组成:(a)图编码,在此阶段执行图卷积和BPR损失的时间复杂度分别为 $\mathcal{O}(|\mathcal{E}|Ld)$ 和 $\mathcal{O}(Bd)$ 。(b)扩散子图增强模块中,训练扩散模型的时间复杂度为 $\mathcal{O}(|\mathcal{U}| \cdot \sum_{l=1}^{L-1} d_l d_{l+1})$,对比损失时间复杂度为 $\mathcal{O}(B^2 d)$ 。(c)局部结构增强模块中,随机游走和局部结构对比的时间复杂度分别为 $\mathcal{O}(Bnd)$ 和 $\mathcal{O}(B^2 d)$ 。(d)邻域语义增强模块中,执行K-Means聚类和基于伪标签提示的对比学习的时间复杂度分别为 $\mathcal{O}(IK|\mathcal{E}|d + B|\mathcal{E}|)$ 和 $\mathcal{O}(B|\tilde{\mathcal{E}}|d)$ 。

4 实验

在本节中我们进行了全面的实验,以回答以下关键问题:

RQ1:与基线相比NeSR的表现如何?

RQ2:不同超参对NeSR的影响大吗?

RQ3:不同增强模块对NeSR性能贡献如何?

RQ4:NeSR集成到不同主流模型中性能如何?

RQ5:NeSR挖掘长尾项目的能力是否有提升?

RQ6:NeSR的鲁棒性如何?

4.1 实验设置

4.1.1 数据集描述

我们在 ML-1M^[52]、Last-FM^[44]、Yelp^[25]、Amazon-Book^[44]和 Amazon-Kindle^[26]五个基准数据集上进行了实验,表1是数据集的统计信息,以7:1:2的比例把数据集划分为训练集、验证集和测试集。根据文献[16,26],我们首先在验证集上搜索最佳超参数,然后将训练集和验证集进行合并用于模型训练,最终在测试集上进行评估。我们在推荐长度为N的列表中使用Recall和NDCG的度量指标,分别从模型预测项目的正确率和项目的相关性排序两个维度进行评估,默认设置为20。

表1 实验数据集的统计

数据集	用户	项目	交互	稀疏度
ML-1M	6,038	3,492	575,281	97.27%
Last-FM	23,566	48,123	1,712,638	99.85%
Yelp	31,668	38,048	1,561,406	99.87%
Amazon-Book	70,679	24,915	846,434	99.95%
Amazon-Kindle	138,333	98,572	1,909,965	99.99%

4.1.2 基准模型

我们采用三组不同类型的模型与NeSR进行对比,分别是基于采样的模型(BPRMF^[51]和DNS^[53])、基于图嵌入的模型(NGCF^[10]和LightGCN^[45])以及基于自监督学习的模型(SSL-DNN^[34]、BUIR^[35]、SGL^[16]、NCL^[29]、XSimGCL^[26]和DiffRec^[54])。

(1) BPRMF^[51]: 使用贝叶斯个性化排序的方式,将推荐任务转化为偏好排序问题。通过优化目标函数,使用户对已交互物品的预测评分高于未交互物品,从而学习个性化的排序关系。

(2) DNS^[53]: 提出了动态负采样策略来优化排名top-N的协同过滤任务,即选择得分较高的样本作为负例进行模型训练。

(3) NGCF^[10]: 基于GCN的推荐算法,旨在将用户项目之间的高阶连通性显式编码到嵌入中,从而更有效地建模协同过滤信号。

(4) LightGCN^[45]: 它是NGCF的简化版本,仅保留邻居聚合操作在U-I图上进行迭代学习节点嵌入。

(5) SSL-DNN^[34]: 基于大规模项目推荐的多任务自监督学习框架,采用对比学习的方式用于推荐中召回模型的训练。

(6) BUIR^[35]: 采用类似学生-老师的网络,通过使用两个不同的编码器网络(在线编码器和目标编码器)引导用户和项目的表示。

(7) SGL^[16]: 基于传统的图扩充方法设计了三种不同的增强策略(节点/边丢弃、随机游走)生成对比视图,在实验中我们选择性能最佳的边丢弃策略。

(8) NCL^[29]: 它是一种基于邻域丰富的对比推荐算法,分别从图结构和语义空间引入节点的邻居来进行对比。

(9) XSimGCL^[26]: 这个算法是在SimGCL^[25]的改进版本,提出一种简单而高效的图对比学习方法,对节点表示进行跨层对比。

(10) DiffRec^[54]: 它利用扩散模型强大的生成能力,将其引入推荐领域,通过迭代去噪过程来恢复原始的用户项目交互数据。

4.1.3 参数设置

实验运行的服务器配置为12th Gen Intel(R) Core(TM) i7-12700KF CPU和NVIDIA GeForce RTX 3090 GPU。在不同算法设置中,我们将所有模型的嵌入维度设置为固定值64,采用Adam来优化模型以及使用Xavier初始化可训练参数,对于具有卷积网络的基准模型我们将层数L设置为2,学习率设置为 $1e^{-3}$, l_2 正则化项系数是 $1e^{-4}$,批处理大小为2048,附录3给出了基准模型的超参数设置。特别地,对于扩散模型的参数设置为:噪声规模为 $1e^{-3}$,扩散步T初始设置为5,推理步T'在 $[0, T/4, T/2, T]$ 中选取,重构交互条目数k设置为10。NeSR在初始化学习嵌入时参数设置为:卷积层设置为2,并在{1,2,3,4}中调整,在 $\{0, 1e^{-3}, 5e^{-3}, \dots, 1e^{-1}, 5e^{-1}, 1\}$ 中选择自监督信号 λ_1 、 λ_2 和 λ_3 ,温度超参设置为0.1,局部结构增强模块中随机游走长度设置为5,邻域语义增强模块中簇划分设置为500。

4.2 对比实验结果及分析(RQ1)

为了全面评估所提算法NeSR的性能优势,我们将其与三组不同类型的推荐模型进行了对比实验。如表2所示,我们在五个基准数据集上报告了当推荐列表长度N=5和N=20时的Recall和NDCG指标,我们有以下发现:

(1) NeSR的性能远超过经典的采样算法BPRMF和DNS。例如NeSR较BPRMF在五个数据集上都有17%以上的提升,特别在Yelp上Recall@5和NDCG@5分别提升了57.32%和59.18%。作为经典的采样算法,BPRMF通过对用户未交互的项目随机采样负样本,优化用户对正负样本的偏好得分差异来训练模型。这一方法因其简单有效,已成为推荐系统中广泛使用的目标函数。DNS在BPR基础上进行改进,动态筛选得分较高的

表2 N=5和N=20时不同模型Top-N的推荐性能

数据集	评估指标	BPRMF	DNS	NGCF	LightGCN	SSL-DNN	UIR	SGL	NCL	XSimGCL	DiffRec	NeSR
ML-1M	Recall@5	0.0970	0.0918	0.0954	0.1099	0.0536	0.0853	0.1107	0.1155	0.1154	0.1137	0.1196
	NDCG@5	0.2787	0.2655	0.2864	0.3191	0.1230	0.2645	0.3275	0.3350	0.3387	0.3320	0.3395
	Recall@20	0.2455	0.2267	0.2413	0.2688	0.1515	0.2062	0.2637	0.2779	0.2713	0.2696	0.2873
	NDCG@20	0.2696	0.2522	0.2713	0.3022	0.1398	0.2409	0.3032	0.3142	0.3128	0.3065	0.3211
Last-FM	Recall@5	0.0293	0.0360	0.0236	0.0375	0.0422	0.0110	0.0354	0.0377	0.0380	0.0431	0.0457
	NDCG@5	0.0494	0.0570	0.0400	0.0605	0.0709	0.0231	0.0610	0.0619	0.0658	0.0727	0.0740
	Recall@20	0.0669	0.0796	0.0529	0.0782	0.0898	0.0309	0.0793	0.0791	0.0817	0.0883	0.0915
	NDCG@20	0.0577	0.0676	0.0464	0.0692	0.0784	0.0273	0.0699	0.0702	0.0733	0.0815	0.0820
Yelp	Recall@5	0.0157	0.0159	0.0157	0.0197	0.0109	0.0150	0.0224	0.0218	0.0244	0.0231	0.0247
	NDCG@5	0.0294	0.0294	0.0297	0.0368	0.0191	0.0294	0.0422	0.0410	0.0457	0.0445	0.0468
	Recall@20	0.0486	0.0497	0.0498	0.0588	0.0364	0.0467	0.0666	0.0661	0.0725	0.0668	0.0737
	NDCG@20	0.0393	0.0400	0.0402	0.0481	0.0281	0.0384	0.0548	0.0539	0.0596	0.0559	0.0599
Amazon-Book	Recall@5	0.0496	0.0589	0.0382	0.0592	0.0488	0.0376	0.0682	0.0627	0.0674	0.0736	0.0756
	NDCG@5	0.0393	0.0460	0.0301	0.0462	0.0392	0.0309	0.0539	0.0495	0.0537	0.0598	0.0607
	Recall@20	0.1181	0.1315	0.0981	0.1376	0.1012	0.1001	0.1556	0.1452	0.1577	0.1511	0.1680
	NDCG@20	0.0623	0.0703	0.0500	0.0726	0.0569	0.0515	0.0832	0.0770	0.0840	0.0870	0.0916
Amazon-Kindle	Recall@5	0.0752	0.0973	0.0458	0.1005	0.1105	0.0367	0.0903	0.0932	0.1012	0.0810	0.1165
	NDCG@5	0.0681	0.0880	0.0413	0.0916	0.1017	0.0314	0.0815	0.0824	0.0935	0.0886	0.1061
	Recall@20	0.1475	0.1817	0.1030	0.1877	0.1911	0.0847	0.1855	0.1831	0.1985	0.0816	0.2219
	NDCG@20	0.0912	0.1150	0.0596	0.1193	0.1273	0.0471	0.1115	0.1114	0.1237	0.1003	0.1392

注:加粗表示最优结果,下同。

负样本,所以性能优于BPRMF。

(2)NeSR的自监督任务提供了明显的性能增益。NGCF和LightGCN属于图嵌入算法,通过将用户项目的高阶关系嵌入在节点表示中,而NeSR在轻量级的卷积操作的基础上,设计了额外的扩散子图、局部结构和邻域语义的自监督任务,使推荐性能有大幅提升,例如NeSR较LightGCN在Yelp数据集上有24%以上的涨幅。

(3)NeSR在自监督推荐模型中有明显的竞争优势。首先NeSR和UIR性能对比最为显著,我们分析认为,UIR性能较差的原因在于随机生成正对增强视图时,未能有效利用用户和项目的邻域信息。其次,即使与同时建模结构和语义的NCL相比,NeSR在Amazon-Book和Amazon-Kindle数据集上仍能实现15%以上的提升,验证了NeSR中局部结构增强和邻域语义增强模块的有效性。最后,NeSR在SSL-DNN、SGL、XSimGCL和DiffRec中展现了竞争优势,说明其既能适应稀疏大规模数据场景,又能兼容不同范式(图增强/无图增强对学习、生成式方法),具有广泛的适用性。

4.3 参数敏感性分析(RQ2)

我们在Last-FM和Amazon-Book两个数据集上选取Recall@20和NDCG@20作为评估指标,对

NeSR中的不同超参的性能进行分析与验证,包括子图、结构和语义的自监督强度 λ_1 、 λ_2 和 λ_3 ,对比学习中的温度 τ ,局部随机游走长度 n ,语义簇数量 K 以及扩散子图增强模块中所涉及到的噪声规模、扩散步 T 和推理步 T' 。

4.3.1 自监督强度 λ_1 、 λ_2 和 λ_3

自监督学习通过构造代理任务辅助模型学习,因此在总损失函数中自监督强度的占比对算法性能具有重要影响。我们先固定其他超参数,将其设置为初始值,来验证扩散子图增强系数 λ_1 、局部结构增强系数 λ_2 和邻域语义增强系数 λ_3 对模型性能的影响。实验结果如图3所示,从中可以观察到,随着自监督强度的增加, λ_1 和 λ_3 对NeSR性能的影响呈现从平缓上升到逐渐下降的趋势。例如在Amazon-Book上,当 λ_1 取0.005, λ_3 取0.05时,Recall和NDCG达到峰值;但当自监督强度进一步增加时,模型性能开始下降。这一现象表明,过强的扩散子图增强和邻域语义增强可能导致模型过拟合。对于局部结构增强系数 λ_2 ,我们发现在Last-FM和Amazon-Book上当 $\lambda_2=1$ 时,模型的性能达到峰值。这一结果表明,局部结构增强在NeSR中具有重要作用,适度地增强强度能够显著提升模型性能。

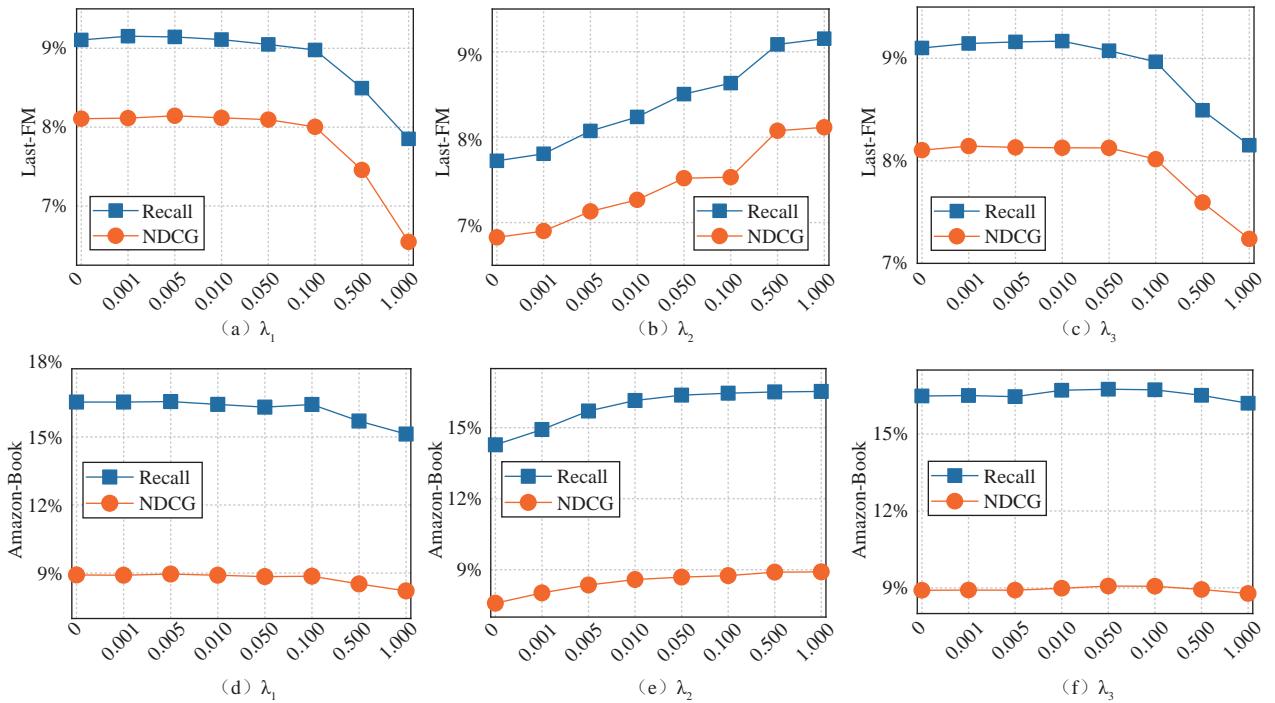


图3 不同自监督强度对NeSR的影响

4.3.2 对比学习温度 τ

在自监督推荐中,通常使用InfoNCE作为对比损失函数^[16,18,25-26]进行模型优化。而温度 τ 是一个超参数,用于调节样本之间的相似性度量,其中较高的 τ 会使损失函数输出更平滑,从而使得正负样本之间的相似度更加接近,较低的 τ 会给模型带来更多的对比损失的梯度信息,使模型更加关注正负样本之间的差异。因此,我们将自监督强度设为对应数据集的峰值,其他超参仍保持预先设置,尝试不同温度参数对NeSR性能的影响。通过表3可知,当温度 τ 取 0.1, Recall 和 NDCG 性能取得最佳。

表3 不同温度 τ 对NeSR的影响

τ	Last-FM		Amazon-Book	
	Recall@20	NDCG@20	Recall@20	NDCG@20
0.1	0.0914	0.0814	0.1675	0.0907
0.2	0.0822	0.0729	0.1675	0.0887
0.3	0.0737	0.0654	0.1527	0.0790
0.4	0.0684	0.0608	0.1386	0.0710
0.5	0.0658	0.0583	0.1296	0.0657

4.3.3 局部随机游走长度 n

在NeSR的局部结构增强模块中,我们验证了不同游走长度($5, 10, 15, 20$)对模型性能的影响。实验中,已验证的参数被设置为对应数据集的最优值,其余超参数则保持预先设置。从表4的实验结果中可

以观察到,在Last-FM和Amazon-Book上,当随机游走长度分别取10和5时,模型性能达到峰值。然而,随着游走长度 n 的增加,模型性能并未显著提升,反而呈现下降趋势。我们推测较长的游走序列可能引入了过多的噪声信息,导致模型难以有效捕捉有用的局部结构特征,从而影响了其性能表现。

表4 局部随机游走长度 n 对NeSR的影响

n	Last-FM		Amazon-Book	
	Recall@20	NDCG@20	Recall@20	NDCG@20
5	0.0912	0.0814	0.1675	0.0907
10	0.0915	0.0817	0.1661	0.0900
15	0.0910	0.0815	0.1646	0.0898
20	0.0896	0.0805	0.1626	0.0895

4.3.4 聚类簇个数 K

在邻域语义增强模块中,我们使用K-Means聚类的方法来获取节点伪标签,让模型在目标节点簇内对比硬负样例。随后在数据集上将已验证过的参数设置为最优值,我们尝试不同的 K 值(100, 500, 1000, 1500, 2000)来观察对性能的影响大小。根据表5的实验结果可以观察到较小的聚类簇个数($K=100$)在Last-FM和Amazon-Book上均能带来最佳性能。随着 K 的增加,模型性能并未显著提升,反而呈下降趋势。这表明过多的簇可能导致语义信息分散,从而影响模型对局部结构的有效学习。

表5 不同聚类簇个数K对NeSR的影响

K	Last-FM		Amazon-Book	
	Recall@20	NDCG@20	Recall@20	NDCG@20
100	0.0912	0.0825	0.1678	0.0913
500	0.0909	0.0817	0.1675	0.0907
1000	0.0905	0.0817	0.1668	0.0905
1500	0.0907	0.0818	0.1660	0.0900
2000	0.0901	0.0809	0.1671	0.0908

4.3.5 扩散模型噪声规模、扩散步T和推理步T'

在本节中,我们对扩散子图增强模块中的三个关键参数进行了分析:噪声规模(noise scale)、扩散步

(T)和推理步(T')。噪声规模控制扩散过程中每一步所添加噪声的强度,扩散步是指在前向扩散过程中添加噪声的步数,而推理步是反向去噪过程中从噪声逐步重建数据所需的步数。从图4的实验结果中可以得出以下观察:首先,较小的噪声规模(例如,噪声规模取值为0.001)能够生成更高质量的样本。其次,在低噪声水平下,增加扩散步数对模型性能的影响较小。为了在性能和计算效率之间取得平衡,我们选择扩散步T=5。值得注意的是,当推理步T'=0时,模型达到了最佳性能,这可能是由于避免了原始U-I图的过度破坏,从而保留更多的有用结构信息。

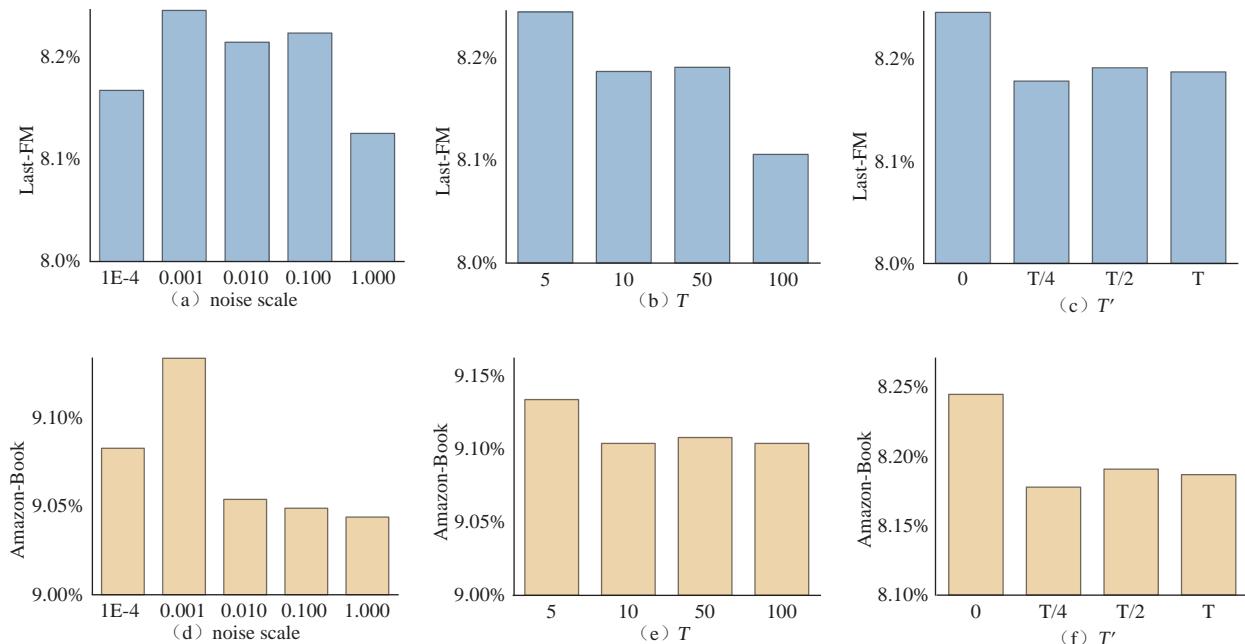


图4 扩散模型的噪声规模、扩散步和推理步的选择

4.4 NeSR模块分析(RQ3)

为了验证本文所提算法不同模块的单独贡献,我们进行了消融实验。接着,针对扩散子图增强、局部结构增强和邻域语义增强模块进行了更细粒度的分析与验证。

4.4.1 消融实验

为了验证本文所提出的算法NeSR不同模块对性能的贡献,我们设计了三个变体模型在Last-FM和Amazon-Book上进行实验,其中“NeSR w/o d”表示NeSR移除扩散子图增强模块,“NeSR w/o l”表示移除局部结构增强模块,“NeSR w/o s”表示NeSR移除邻域语义增强模块。从表6中观察得知,NeSR算法始终优于其变体模型,缺少任何一部分都会导致性能有不同程度的损失:

(1)局部结构增强模块,移除该模块(“NeSR w/o

l”)对性能的影响最大,在Last-FM和Amazon-Book上Recall@20分别下降17.05%和14.58%,NDCG@20分别下降18.41%和17.03%。这表明局部结构信息是模型的核心,不仅直接贡献于节点关系建模,还为邻域语义模块提供了必要的初始化,缺失会导致整体性能显著退化。

(2)对于扩散子图增强和邻域语义增强模块来说,移除后性能降幅较小。例如“NeSR w/o d”在两个数据集上的指标下降幅度在1.09%-5.65%之间,“NeSR w/o s”的下降幅度在0.44%-3.06%之间。虽然这些模块对最终性能的提升幅度较小,但其在特定场景下的补充作用仍不可忽视。例如邻域语义增强模块能够挖掘额外的潜在信息,为模型提供更深层的用于微调推荐性能的自监督信号。

此外,我们针对NeSR不同模块设计的合理性

表6 整体消融实验

模型	Last-FM		Amazon-Book	
	Recall@20	NDCG@20	Recall@20	NDCG@20
NeSR w/o d	0.0905	0.0809	0.1585	0.0870
NeSR w/o l	0.0759	0.0669	0.1435	0.0760
NeSR w/o s	0.0911	0.0812	0.1647	0.0888
NeSR	0.0915	0.0820	0.1680	0.0916

进行验证，并设计了三个变体模型，其中“NeSR-drop”表示在扩散子图增强模块中，采用随机丢弃边策略来生成对比视图。“NeSR-Layer”表示在局部结构增强模块中采用图编码器的偶数层作为锚节点的正样本对，“NeSR-Louvain”在邻域语义增强模块中采用基于图的社区检测方法Louvain作为伪标签的生成方式。从表7中可以观察到：

(1)“NeSR-Layer”变体性能较差，相较于NeSR在Last-FM和Amazon-Book数据集上有14%-18%的下降幅度。对于NeSR单一使用GCN的同构层

表7 局部消融实验

模型	Last-FM		Amazon-Book	
	Recall@20	NDCG@20	Recall@20	NDCG@20
NeSR-drop	0.0898	0.0801	0.1657	0.0895
NeSR-Layer	0.0778	0.0683	0.1427	0.0751
NeSR-Louvain	0.0794	0.0721	0.1597	0.0874
NeSR	0.0915	0.0820	0.1680	0.0916

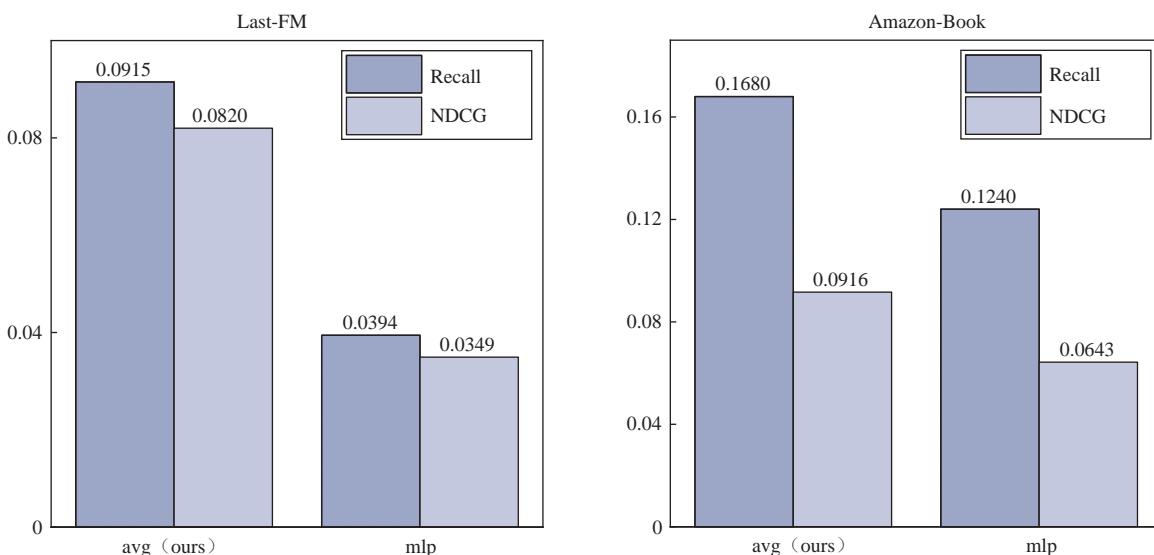


图5 选择不同子图聚合方式

4.4.3 邻域语义增强模块中负样例的选择

NeSR在邻域语义增强中提出了基于伪标签提示的对比目标函数，我们将负样本的选择压缩到和锚节点同一个簇内，即在模型优化时只使用硬负样

信息作为锚节点的正样本对，这使学习到的节点表示不可避免地会发生平滑现象，从而导致性能下降。

(2)作为邻域模块变体“NeSR-Louvain”来说，在两个数据集上都有不同程度的下降。这说明NeSR在邻域语义增强模块中采用简单高效的K-Means聚类具有一定的优势。

(3)当我们选择采用随机边丢弃策略“NeSR-drop”来生成对比视图时，在两个数据集上的性能都有所下降，这说明利用扩散模型能够生成高质量的U-I对比视图。

4.4.2 局部结构增强模块中聚合方式的选择

局部结构增强模块中，我们采用随机游走得到子图集合，并通过融合函数Fusion(•)来得到局部结构的嵌入表示。关于随机游走长度n我们在4.3.3节中已验证，这里对Fusion(•)进行进一步的探索，我们采用两种常用的聚合方式进行验证，即均值(avg)和线性映射(两层的mlp)。图5报告了在Last-FM和Amazon-Book数据集上选择不同聚合方式下NeSR的性能，我们发现采用线性映射的方式并未展现出其灵活性，在两个数据集上的表现欠佳，而采用均值聚合的性能指标较好，它通过对游走集合中节点的特征进行了平均，能够简单快速得到局部结构的嵌入表示，所以我们在局部结构增强模块中选择均值聚合。

本对(“in-cluster”)的梯度信息。为了验证此方法的有效性，我们在训练批次内构建负样本对(“in-batch”)为对照组进行实验，结果如图6所示。在Last-FM和Amazon-Book上，“in-cluster”在Recall和

NDCG 指标上均显著优于“in-batch”变体。这说明了我们提出伪标签提示的对比学习方法,能够通过

使用伪标签数据在簇内选择硬负样例,为模型提供更有价值的梯度信息。

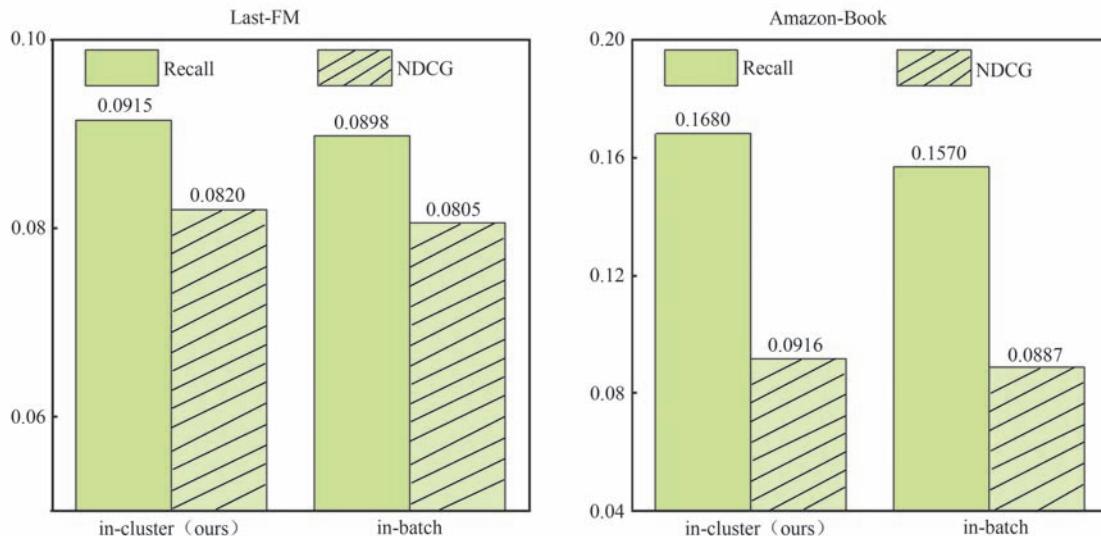


图6 负样例的选择

4.5 Backbone 对比(RQ4)

NeSR作为一种与模型无关的对比学习框架,通过对原始图数据从子图、结构和语义三个方面扩充自监督信号,具有即插即用的通用性。为验证其广泛适应性,我们将 NeSR 集成到不同主流推荐模型(BPRMF、NGCF 和 LightGCN)中进行实验评估。实验结果如表8所示,所提方法整体上能够有效提升基线模型的性能,例如BPRMF在Last-FM 上 Recall@20 和 NDCG@20 分别提升了 9.87% 和 12.13%,NGCF 在两个数据集上有 23%-41% 的涨幅,LightGCN 有 17%-26% 的提升。值得注意的是,由于BPRMF本身不包含图卷积操作,无法有效利用扩散子图增强模块所带来的拓扑感知能力,这一局限性在交互稀疏的 Amazon-Book 中表现尤为明显:在该场景下,以BPRMF为基础的模型性能出现了明显下降。我们推测这主要是由于两个因素:一方面,数据稀疏使得模型难以捕捉有效的结构信息;另一方面,BPR 算法无法建模用户与项目之间的高阶关系,导致表示学习不充分。相比之下,当 NeSR 与包含图卷积机制的 NGCF 或 LightGCN 结合时,模型能够更好地捕捉并利用高阶邻接关系,从而显著缓解上述问题。因此,将 NeSR 集成到主流图神经网络推荐模型中可带来一定的性能增益。

4.6 提升长尾项目的能力(RQ5)

为了验证 NeSR 在具有不同稀疏级别的推荐

表8 关于不同架构上 NeSR 的性能对比

模型	Last-FM		Amazon-Book	
	Recall@20	NDCG@20	Recall@20	NDCG@20
BPRMF	0.0669	0.0577	0.1181	0.0623
+ours	0.0735	0.0647	0.0763	0.0433
NGCF	0.0529	0.0464	0.0981	0.0500
+ours	0.0745	0.0658	0.1209	0.0635
LightGCN	0.0782	0.0692	0.1376	0.0726
+ours(NeSR)	0.0915	0.0820	0.1680	0.0916

性能是否有所提升,我们将数据集中用户的交互条目分为 5 组。例如 Last-FM 中将交互条目小于等于 51、103、183、351 和 2881 的用户分为 5 个组别,尽可能地保证每小组占总交互条目的 20%。越靠后的组别,其包含的热门项目越多。然后我们在 Last-FM 和 Amazon-Book 数据集上进行实验,以检查 Recall@20 在每小组的性能指标,结果如图 7 所示。我们从中可以观察到 XSimGCL 倾向于推荐热门项目,在 Last-FM 最后三个组别上取得了较高的 Recall 值,这一现象表明该模型存在明显的热门项目推荐偏好,其推荐结果更倾向于流行度较高的项目。而 NeSR 对长尾项目的推荐效果尤为突出,例如在 Last-FM 上,NeSR 在长尾组别的 Recall 值较 XSimGCL 提升显著。这表明 NeSR 通过扩散子图增强、局部结构增强和邻域语义增强进行建模,能够更好地捕捉长尾项目的特征表示。

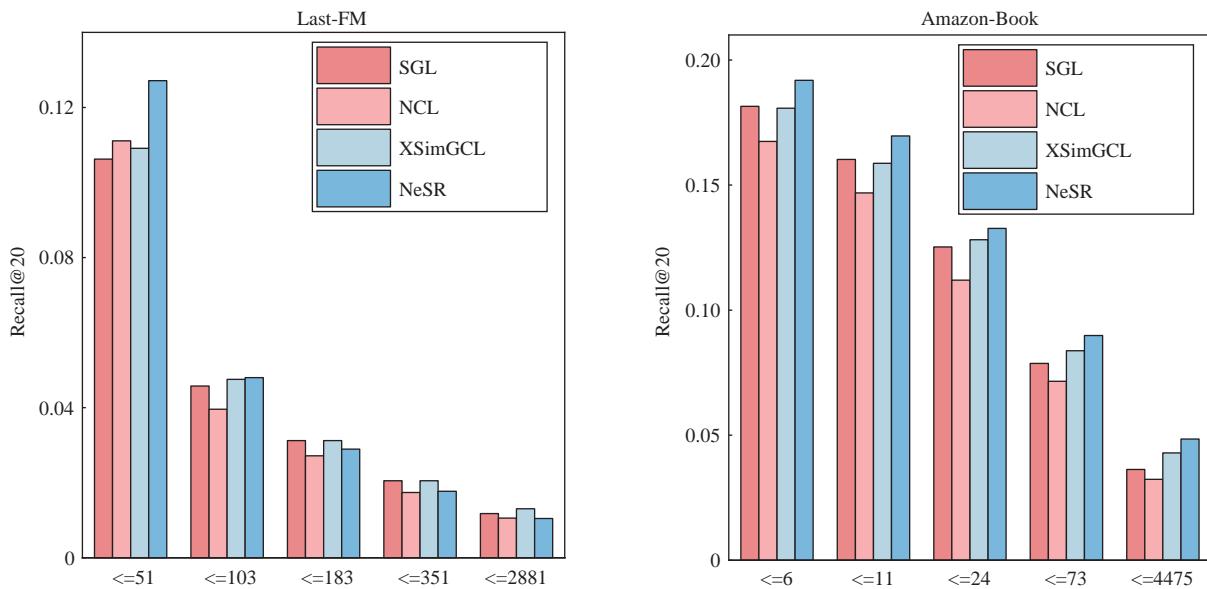


图7 不同稀疏级别的性能比较

4.7 模型的鲁棒性(RQ6)

为了评估所提出模型的鲁棒性,我们通过在训练数据中随机丢弃用户-项目交互数据进行实验。以Amazon-Book数据集为例,我们设置了不同的丢弃比例(10%、20%、30%、40%和50%),并在每种丢弃情况下重新训练模型,实验结果如图8所示。随着丢弃比例的增加,所有模型的性能均出现不同程度的下降。这是由于交互数据的减少降低了用户和项目节点的语义信息完整性,导致模型学习的表示质量下降。相比基准模型(如SGL、XSimGCL和NCL),NeSR在所有丢弃比例下均保持相对较高的指标,特别是在50%丢弃情况下,仍显著优于基准模型。这表明NeSR能够有效捕捉自监督信号,即使在数据稀疏的情况下也能生成高质量的节点表示。

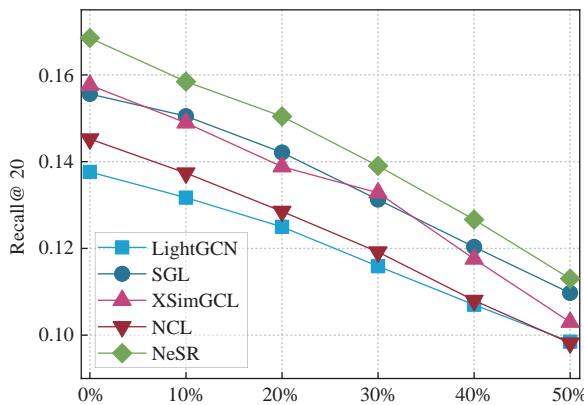


图8 模型的鲁棒性分析

5 总结与展望

在这项工作中,我们提出基于邻域增强的图自监督推荐算法(NeSR),在不破坏原始图结构的情况下,从子图、结构和语义三个方面提取自监督信号,是一种即插即用、与模型无关的对比学习框架。具体来说我们引入扩散模型作为数据增强操作,通过前向扩散和反向去噪过程生成高质量的无噪音图,并将子图作为对比学习的增强视图,以捕捉图结构的多视角信息。其次,为了充分利用数据的结构信息,设计了局部结构增强模块,来加强节点与其邻域之间的可靠的关联关系。随后,我们为了探索语义邻域中的自监督信号,提出了基于伪标签提示的目标函数,通过对嵌入进行聚类来获得节点的伪标签以指导模型优化。我们在五个数据集中进行了大量的实验并证明所提算法NeSR的有效性。在未来的工作中,我们将进一步探索生成式自监督推荐算法,通过引入多模态辅助信息(如文本、图像等),构建多模态扩散模型推荐框架,以更全面地捕捉用户-项目交互的潜在关系,从而进一步提升推荐系统的性能。

参考文献

- [1] Zhao K, Zheng Y, Zhuang T, et al. Joint learning of e-commerce search and recommendation with a unified graph neural network// Proceedings of the Fifteenth ACM International Conference on

- Web Search and Data Mining. Online, 2022: 1461-1469
- [2] Singer U, Roitman H, Eshel Y, et al. Sequential modeling with multiple attributes for watchlist recommendation in e-commerce //Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. Online, 2022: 937-946
- [3] Raza S, Ding C. News recommender system: a review of recent progress, challenges, and opportunities. Artificial Intelligence Review, 2022, 55: 749-800
- [4] Christner C, Urman A, Adam S, et al. Automated tracking approaches for studying online media use: A critical review and recommendations. Communication Methods and Measures, 2022, 16(2): 79-95
- [5] Li Q, Wang X, Wang Z, et al. Be causal: De-biasing social network confounding in recommendation. ACM Transactions on Knowledge Discovery from Data, 2023, 17(1): 1-23
- [6] Liu H, Zhang X, Yang B, et al. An enhanced influence diffusion model for social recommendation. Chinese Journal of Computers, 2023, 46(3): 626-642 (in Chinese)
(刘会, 张璇, 杨兵, 等. 用于社交推荐的增强影响扩散模型. 计算机学报, 2023, 46(3): 626-642)
- [7] Sedhain S, Menon A K, Sanner S, et al. Autorec: Autoencoders meet collaborative filtering//Proceedings of the 24th International Conference on World Wide Web. Florence, Italy, 2015: 111-112
- [8] Chen J, Zhang H, He X, et al. Attentive collaborative filtering: Multimedia recommendation with item-and component-level attention //Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval. Tokyo, Japan, 2017: 335-344
- [9] He X, Liao L, Zhang H, et al. Neural collaborative filtering//Proceedings of the 26th International Conference on World Wide Web. Perth, Australia, 2017: 173-182
- [10] Wang X, He X, Wang M, et al. Neural graph collaborative filtering //Proceedings of the 42nd international ACM SIGIR conference on Research and Development in Information Retrieval. Paris, France, 2019: 165-174
- [11] Sarwar B M. Sparsity, scalability, and distribution in recommender systems. University of Minnesota, USA, 2001
- [12] Zhang B, Li L, Zha Z, et al. Contrastive cross-modal representation learning based active learning for visual question answer. Chinese Journal of Computers, 2022, 45(8): 1730-1745 (in Chinese)
(张北辰, 李亮, 查正军, 等. 基于跨模态对比学习的视觉问答主动学习方法. 计算机学报, 2022, 45(8): 1730-1745)
- [13] Zhao H, Wang W, Li J, et al. Fine-grained multimodal entity linking for videos. Journal of Software, 2022, 34(03): 1140-1153 (in Chinese)
(赵海全, 王续武, 李金亮, 等. 面向视频的细粒度多模态实体链接. 软件学报, 2022, 34(03): 1140-1153)
- [14] Elnaggar A, Heinzinger M, Dallago C, et al. Prottrans: Toward understanding the language of life through self-supervised learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(10): 7112-7127
- [15] Baevski A, Hsu W N, Xu Q, et al. Data2vec: A general framework for self-supervised learning in speech, vision and language//Proceedings of the International Conference on Machine Learning. Baltimore, USA, 2022: 1298-1312
- [16] Wu J, Wang X, Feng F, et al. Self-supervised graph learning for recommendation//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. Online, 2021: 726-735
- [17] Wang H, Xu Y, Yang C, et al. Knowledge-Adaptive Contrastive Learning for Recommendation//Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. Singapore, 2023: 535-543
- [18] Cai X, Huang C, Xia L, et al. LightGCL: Simple Yet Effective Graph Contrastive Learning for Recommendation//Proceedings of the Eleventh International Conference on Learning Representations. Kigali, Rwanda, 2023: 1-16
- [19] Duan J, Wang S, Zhang P, et al. Graph anomaly detection via multi-scale contrastive learning networks with augmented view//Proceedings of the AAAI Conference on Artificial Intelligence. Washington, USA, 2023, 37(6): 7459-7467
- [20] Xing Z, Feng Q, Chen H, et al. A survey on video diffusion models. ACM Computing Surveys, 2024, 57(2): 1-42
- [21] Hatamizadeh A, Song J, Liu G, et al. Diffit: Diffusion vision transformers for image generation//Proceedings of the European Conference on Computer Vision. Milan, Italy, 2024: 37-55
- [22] Xu Z, Zhang J, Liew J H, et al. Magicanimate: Temporally consistent human image animation using diffusion model//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA, 2024: 1481-1490
- [23] Jiang Y, Yang Y, Xia L, et al. Diffkg: Knowledge graph diffusion model for recommendation//Proceedings of the 17th ACM International Conference on Web Search and Data Mining. Mérida, México, 2024: 313-321
- [24] Jiang Y, Xia L, Wei W, et al. Diffmm: Multi-modal diffusion model for recommendation//Proceedings of the 32nd ACM International Conference on Multimedia. Melbourne, Australia, 2024: 7591-7599
- [25] Yu J, Yin H, Xia X, et al. Are graph augmentations necessary? simple graph contrastive learning for recommendation//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022: 1294-1303
- [26] Yu J, Xia X, Chen T, et al. XSimGCL: Towards extremely simple graph contrastive learning for recommendation. IEEE Transactions on Knowledge and Data Engineering, 2023, 36(2): 913-926
- [27] Li J, Zhou P, Xiong C, et al. Prototypical contrastive learning of unsupervised representations//Proceedings of the Ninth International Conference on Learning Representations. Online, 2021: 1-16
- [28] Lin S, Liu C, Zhou P, et al. Prototypical graph contrastive learning. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(2): 2747-2758
- [29] Lin Z, Tian C, Hou Y, et al. Improving graph collaborative

- filtering with neighborhood-enriched contrastive learning// Proceedings of the ACM Web Conference 2022. Lyon, France, 2022: 2320-2329
- [30] Yu J, Yin H, Xia X, et al. Self-supervised learning for recommender systems: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(1): 335-355
- [31] Hjelm R D, Fedorov A, Lavoie-Marchildon S, et al. Learning deep representations by mutual information estimation and maximization//Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA, 2019: 1-24
- [32] Veličković P, Fedus W, Hamilton W L, et al. Deep graph infomax//Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA, 2019: 1-17
- [33] Sun F Y, Hoffmann J, Verma V, et al. Infograph: Unsupervised and semi-supervised graph-level representation learning via mutual information maximization//Proceedings of the 8th International Conference on Learning Representations. Online, 2020: 1-16
- [34] Yao T, Yi X, Cheng D Z, et al. Self-supervised learning for large-scale item recommendations//Proceedings of the 30th ACM International Conference on Information & Knowledge Management. Gold Coast, Australia, 2021: 4321-4330
- [35] Lee D, Kang S K, Ju H, et al. Bootstrapping user and item representations for one-class collaborative filtering// Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. Online, 2021: 317-326
- [36] Wang R, Wang X, Shi C, et al. Uncovering the Structural Fairness in Graph Contrastive Learning. *Advances in Neural Information Processing Systems*, 2022, 35: 32465-32473
- [37] Zhu G, Lu W, Yuan C, et al. AdaMCL: Adaptive Fusion Multi-View Contrastive Learning for Collaborative Filtering// Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. Taipei, China, 2023: 1076-1085
- [38] Zou D, Wei W, Mao X L, et al. Multi-level cross-view contrastive learning for knowledge-aware recommender system//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022: 1358-1368
- [39] Ma G F, Yang X H, Long H, et al. Robust social recommendation based on contrastive learning and dual-stage graph neural network. *Neurocomputing*, 2024, 584: 127597
- [40] Xu M, Wang H, Ni B, et al. Self-supervised graph-level representation learning with local and global structure// Proceedings of the International Conference on Machine Learning. Vienna, Austria, 2021: 11548-11558
- [41] Lee N, Lee J, Park C. Augmentation-free self-supervised learning on graphs//Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence. Online, 2022: 7372-7380
- [42] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks//Proceedings of the 5th International Conference on Learning Representations. Toulon, France, 2017: 1-14
- [43] Wu F, Souza A, Zhang T, et al. Simplifying graph convolutional networks//Proceedings of the 36th International conference on machine learning. Long Beach, USA, 2019: 6861-6871
- [44] Wang X, He X, Cao Y, et al. Kgat: Knowledge graph attention network for recommendation//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Anchorage, USA, 2019: 950-958
- [45] He X, Deng K, Wang X, et al. Lightgen: Simplifying and powering graph convolution network for recommendation// Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. Xi'an, China, 2020: 639-648
- [46] Hamilton W, Ying Z, Leskovec J. Inductive representation learning on large graphs//Proceedings of the Neural Information Processing Systems. Long Beach, USA, 2017: 1025-1035
- [47] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 2020, 33: 6840-6851
- [48] Li Q, Han Z, Wu X M. Deeper insights into graph convolutional networks for semi-supervised learning// Proceedings of the AAAI Conference on Artificial Intelligence. New Orleans, USA, 2018: 3538-3545
- [49] Grover A, Leskovec J. node2vec: Scalable feature learning for networks//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, USA, 2016: 855-864
- [50] Xia W, Wang Q, Gao Q, et al. Self-consistent contrastive attributed graph clustering with pseudo-label prompt. *IEEE Transactions on Multimedia*, 2023, 25: 6665-6677
- [51] Rendle S, Freudenthaler C, Gantner Z, et al. BPR: Bayesian personalized ranking from implicit feedback//Proceedings of the Conference on Uncertainty in Artificial Intelligence. Montreal, Canada, 2009: 452-461
- [52] Harper F M, Konstan J A. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, 2015, 5(4): 1-19
- [53] Zhang W, Chen T, Wang J, et al. Optimizing top-n collaborative filtering via dynamic negative item sampling// Proceedings of the 36th international ACM SIGIR conference on Research and Development in Information Retrieval. Dublin, Ireland, 2013: 785-788
- [54] Wang W, Xu Y, Feng F, et al. Diffusion recommender model// Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. Taipei, China, 2023: 832-841
- [55] Gao T, Yao X, Chen D. Simcse: Simple contrastive learning of sentence embeddings//Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. Punta Cana, Dominican Republic, 2021: 6894-6910
- [56] Wang C, Yu Y, Ma W, et al. Towards representation alignment and uniformity in collaborative filtering//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington, USA, 2022: 1816-

1825

附录1. 模型的局部收敛性分析

在优化问题中, 算法的收敛性依赖于目标函数的梯度信息, 梯度 Lipschitz 平滑性直接与梯度下降算法的行为相关, 能够保证函数下降、收敛性和收敛速率。在我们的工作中, NeSR 的目标函数主要由推荐损失(\mathcal{L}_{BPR})和对比损失(\mathcal{L}_{dm} 、 \mathcal{L}_{local} 和 \mathcal{L}_{sema})多个函数组成, 对其局部收敛性进行分析可以先将 NeSR 的优化目标拆解为推荐损失和对比损失的函数, 分别验证这些损失是否具有梯度 Lipschitz 平滑性。以单个用户 u 为例, 关于推荐损失, 能够找到 BPR 损失的梯度 Lipschitz 常数 $L_1 \leq 0.25$ 是有界的, 使其满足:

$$\|\nabla \mathcal{L}_{BPR}(h_u^{(1)}) - \nabla \mathcal{L}_{BPR}(h_u^{(2)})\| \leq L_1 \|h_u^{(1)} - h_u^{(2)}\| \quad (1)$$

因此 \mathcal{L}_{BPR} 损失的梯度是 Lipschitz 连续的。关于 NeSR 多个对比损失, 以 \mathcal{L}_{dm} 为例, 其梯度为

$$\nabla_{h_u} \mathcal{L}_{dm} = -\frac{1}{\tau} \left(\hat{h}_u - \sum_{v \in \mathcal{U}} \hat{h}_v \frac{\exp(h_u^T \hat{h}_v / \tau)}{\sum_{v' \in \mathcal{U}} \exp(h_u^T \hat{h}_{v'} / \tau)} \right) \quad (2)$$

而其梯度的范数为

$$\|\nabla_{h_u} \mathcal{L}_{dm}\| \leq \frac{1}{\tau} \left(\|\hat{h}_u\| + \sum_{v \in \mathcal{U}} \|\hat{h}_v\| \frac{\exp(h_u^T \hat{h}_v / \tau)}{\sum_{v' \in \mathcal{U}} \exp(h_u^T \hat{h}_{v'} / \tau)} \right) \quad (3)$$

其中,

$$\sum_{v \in \mathcal{U}} \frac{\exp(h_u^T \hat{h}_v / \tau)}{\sum_{v' \in \mathcal{U}} \exp(h_u^T \hat{h}_{v'} / \tau)} = 1 \quad (4)$$

因此公式(3)整理可得

$$\|\nabla_{h_u} \mathcal{L}_{dm}\| \leq \frac{2\|\hat{h}_u\|}{\tau} \quad (5)$$

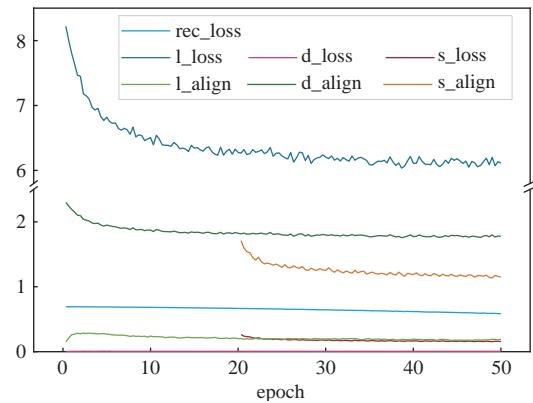
由公式(5)可知该范数是有界的, 因此 \mathcal{L}_{dm} 是 Lipschitz 连续的, 其 Lipschitz 常数为 $L_2 = 2\|\hat{h}_u\|/\tau$, 并满足

$$\|\nabla \mathcal{L}_{dm}(h_u^{(1)}) - \nabla \mathcal{L}_{dm}(h_u^{(2)})\| \leq L_2 \|h_u^{(1)} - h_u^{(2)}\| \quad (6)$$

类似地, \mathcal{L}_{local} 和 \mathcal{L}_{sema} 损失的梯度也是 Lipschitz 连续的。对于多个满足 Lipschitz 梯度连续性函数的相加后的函数也满足 Lipschitz 梯度连续性, 所以 NeSR 满足 Lipschitz 梯度连续性。这确保了 NeSR 的梯度变化速率平稳, 优化过程中不会出现剧烈振荡, 有助于局部收敛的实现。为了验证 NeSR 的优化过程及收敛性能, 我们在 Amazon-Book 数据集上进行了实验, 并绘制了损失函数的变化曲线。我们可以从附图1中观察得到, 局部结构增强损失(l_loss)随着 epoch 的增加而逐渐收敛。由于自监督任务的权重系数 λ_1 和 λ_3 的设置较小, 所以扩散子图增强损失(d_loss)和邻域语义增强损失(s_loss)在数值上收敛接近零。

附录2. 一致性的理论分析

节点一致性是节点表征在训练过程中逐渐趋于一致的



附图1 在 Amazon-Book 数据集上的 NeSR 损失曲线

重要部分。我们给出 NeSR 一致性的公式:

$$l_{align} = \mathbb{E}_{(u, u^+) \sim p_{user}} \|\tilde{h}(u) - \tilde{h}(u^+)\|^2 + \mathbb{E}_{(i, i^+) \sim p_{item}} \|\tilde{h}(i) - \tilde{h}(i^+)\|^2 \quad (7)$$

其中, (u, u^+) 和 (i, i^+) 是从不同模块中收集到的用户和项目的正样本对, $\tilde{h}(\cdot)$ 表示经过 l_2 规范化后的嵌入表示。根据工作[55, 56], 通过优化对比目标函数, 我们能够最大化正样本对的相似度, 同时最小化负样本对的相似度。具体而言, 对比目标函数的梯度更新方向推动正样本对的嵌入逐渐接近, 而负样本对的嵌入逐渐远离。此外, 正如在附录1中所分析的, 对比学习的目标函数是局部收敛的, 即随着训练的进行, 节点表征在嵌入空间中逐渐形成清晰的簇结构, 从而能够实现高一致性。我们进一步在 Amazon-Book 上对一致性损失曲线进行了可视化分析, 如附图1所示。实验结果表明, NeSR 框架通过局部结构增强模块(l_align)实现了强有力的跨视图一致性引导。同时扩散子图增强模块(d_align)与邻域语义增强模块(s_align)的协同优化机制, 有效促进了推荐任务的性能提升。

附录3. 基准模型的参数设定

在这一部分, 我们详细说明了基准模型在五个数据集上的超参数设置及实验环境, 以确保实验的透明性和结果的可复现性。所有实验均在配备 NVIDIA GeForce RTX 3090 GPU 的机器上进行, 使用 PyTorch 1.13.0 版本实现。对于基准模型的超参数设置:(1)BPRMF 没有超参数选择。(2)DNS 的负样例选择数设置为 8。(3)针对基于图卷积神经网络的基准模型, 我们将其卷积层统一设置为 2。(4)其余基准模型的超参数具体设置列于附表1中。其中, 对于某些超参数的选择, 我们主要参考了原论文中的推荐值。例如, 在 SSL-DNN 中 softmax 温度设置为 0.07; 在 BUIR 中 momentum 系数设置为 0.995。对部分模型超参数的调优, 我们采用了单变量验证的方法, 调整关键参数以优化验证集上的性能。

附录4. 选择邻域语义节点时的案例研究

为了进一步分析邻域语义节点, 我们在 Amazon-Book

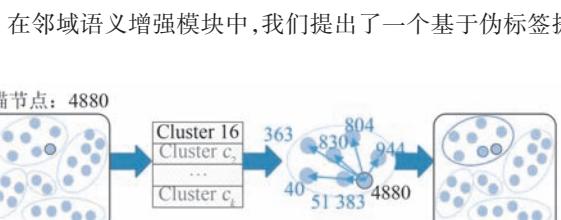
附表1 基准算法超参设置

模型	ML-1M	Last-FM	Yelp
SSL-DNN	$\alpha=0.1, dr=0.1, \tau=0.07$	$\alpha=0.1, dr=0.1, \tau=0.07$	$\alpha=0.1, dr=0.1, \tau=0.07$
BUIR	$\rho=0.2, \tau=0.995$	$\rho=0.2, \tau=0.995$	$\rho=0.2, \tau=0.995$
SGL	$\lambda=0.1, \rho=0.1, \tau=0.2$	$\lambda=0.1, \rho=0.1, \tau=0.2$	$\lambda=0.1, \rho=0.1, \tau=0.2$
NCL	$\lambda_1=1e^{-6}, \lambda_2=1e^{-7}, \tau=0.05, \alpha=1.5, k=2000$	$\lambda_1=1e^{-6}, \lambda_2=1e^{-7}, \tau=0.05, \alpha=1.5, k=2000$	$\lambda_1=1e^{-6}, \lambda_2=1e^{-7}, \tau=0.05, \alpha=1.5, k=2000$
XSimGCL	$\lambda=0.2, \epsilon=0.2, l^*=1, \tau=0.15$	$\lambda=0.2, \epsilon=0.2, l^*=1, \tau=0.15$	$\lambda=0.2, \epsilon=0.2, l^*=1, \tau=0.15$
DiffRec	$lr=1e^{-3}, dims=[200, 600], scale=1e^{-3}, min=5e^{-3}, max=1e^{-2}, T=40, T'=0$	$lr=1e^{-5}, dims=[1000], scale=1e^{-2}, min=1e^{-3}, max=1e^{-2}, T=5, T'=0$	$lr=1e^{-5}, dims=[1000], scale=1e^{-2}, min=1e^{-3}, max=1e^{-2}, T=5, T'=0$
NeSR	$\lambda_1=\lambda_3=1e^{-3}, \lambda_2=5e^{-3}, \tau=0.2, n=5, K=100, scale=1e^{-3}, T=5, T'=0$	$\lambda_1=5e^{-3}, \lambda_2=1, \lambda_3=1e^{-3}, \tau=0.1, n=10, K=100, scale=1e^{-3}, T=5, T'=0$	$\lambda_1=\lambda_3=1e^{-3}, \lambda_2=5e^{-3}, \tau=0.1, n=5, K=500, scale=1e^{-2}, T=10, T'=0$

附表1 基准算法超参设置(续表)

模型	Amazon-Book	Amazon-Kindle
SSL-DNN	$\alpha=0.1, dr=0.1, \tau=0.07$	$\alpha=0.1, dr=0.1, \tau=0.07$
BUIR	$\rho=0.2, \tau=0.995$	$\rho=0.2, \tau=0.995$
SGL	$\lambda=0.1, \rho=0.1, \tau=0.2$	$\lambda=0.05, \rho=0.1, \tau=0.2$
NCL	$\lambda_1=1e^{-6}, \lambda_2=1e^{-7}, \tau=0.05, \alpha=1.5, k=2000$	$\lambda_1=1e^{-6}, \lambda_2=1e^{-7}, \tau=0.05, \alpha=1.5, k=2000$
XSimGCL	$\lambda=0.2, \epsilon=0.1, l^*=1, \tau=0.2$	$\lambda=0.2, \epsilon=0.2, l^*=1, \tau=0.2$
DiffRec	$lr=5e^{-5}, dims=[1000], scale=1e^{-4}, min=5e^{-4}, max=5e^{-3}, T=5, T'=0$	$lr=5e^{-6}, dims=[1000], scale=1e^{-4}, min=5e^{-4}, max=5e^{-3}, T=5, T'=0$
NeSR	$\lambda_1=5e^{-3}, \lambda_2=1, \lambda_3=5e^{-2}, \tau=0.1, n=5, K=100, scale=1e^{-3}, T=5, T'=0$	$\lambda_1=\lambda_3=1e^{-3}, \lambda_2=5e^{-3}, \tau=0.1, n=5, K=100, scale=1e^{-3}, T=5, T'=0$

数据集中随机选择一个用户(id为4880)为锚节点,提取与锚节点同一个簇(伪标签为16)的用户节点。通过计算节点相似性 $SIM(u, q)$, 我们选取与锚节点最相似的用户(id为51383)作为邻域语义节点, 具体获取流程如附图2所示。



附图2 邻域语义节点的获取

示的优化损失函数, 并通过引入邻域语义节点为对比学习提供了两方面的增强:(1)提供高质量的正样本。邻域语义节点选择算法能够根据节点间的语义相似性, 从簇内筛选出语义最相关的节点。这使得对比学习中的正样本对更加精准, 减少了随机选择正样本可能引入的语义偏差, 从而提升了模型的学习效果。(2)加快模型优化。在传统对比学习中, 负样本通常从全图中随机选择, 这可能引入无关或噪声样本, 导致梯度信号较弱。而通过将负样本的选择范围压缩到锚节点所在簇(伪标签16)内的节点, 显著提高了负样本的质量, 减少了不必要的干扰。模型能够获得更加有价值的梯度信息, 从而加快优化过程, 提高训练效率。



WEN Xi-Lin, Ph. D. candidate. Her research interests include recommender systems, diffusion model and contrastive learning.

YANG Xu-Hua, Ph. D., professor. His research interests include graph machine learning, recommender systems.

MA Gang-Feng, Ph. D., post-doctor. His research interests include graph neural networks, recommender systems.

Background

Deep learning based recommendations have a significant impact on people's lives in various scenarios and have become an indispensable part of people's daily lives. In comparison to traditional collaborative filtering, deep learning-based recommendation algorithms excel in handling large-scale, high-dimensional data. Despite these algorithms demonstrating outstanding performance in specific scenarios, the design of deep architectures necessitates a substantial amount of data for training. As recommendation system relies on user-item (U-I) interaction records, they face challenges of data sparsity and high collection costs.

Graph-based self-supervised contrastive learning has emerged as a prominent research direction in recommendation systems due to its ability to enhance model robustness and generalization, particularly in sparse and noisy U-I interaction scenarios. These methods augment the U-I interaction graph by constructing auxiliary self-supervised tasks and extracting contrastive signals from multiple enhanced graphs. The core idea is to enforce consistency between different enhanced graphs, thereby enabling the model to learn high-quality node representations. Nevertheless, despite their effectiveness, current graph-based self-supervised recommendation methods face several limitations. First, their performance heavily depends on manually designed augmentation strategies—such as node dropping, edge removal, or subgraph sampling—which may distort the intrinsic structure of the original U-I graph to varying degrees. These distortions can obscure crucial topological patterns, thereby limiting the model's ability to capture authentic graph characteristics. Second, many existing approaches fail to fully exploit the rich structural and semantic information inherently embedded within the graph. In particular, they often construct positive sample pairs solely from enhanced graphs, neglecting the local structure and neighborhood semantics of the original U-I graph. This limitation may hinder the expressiveness and discriminative capacity of the learned representations. To address these challenges, we propose Neighborhood-enhanced graph Self-supervised learning for Recommendation (NeSR)—a novel framework that extracts self-supervised signals from three complementary perspectives: subgraph, structural, and semantic information. These signals are integrated into the corresponding contrastive learning objective

to improve recommendation performance in a principled and robust manner. First, we employ a diffusion model as a generative augmentation mechanism. By simulating the forward diffusion and reverse denoising processes, we construct high-quality, noise-free subgraphs. These subgraphs serve as reliable enhanced graphs for contrastive learning, offering stable global structural cues that improve the robustness of node representation learning. Second, we leverage random walks on the original U-I graph to extract local structural information. The resulting walk sequences are treated as contextually relevant positive samples for users or items. This approach enables the model to generate self-supervised signals that are grounded in the inherent topological patterns of the interaction graph, thus yielding more accurate and informative node embeddings. Furthermore, to capture the semantic relationships between nodes and their neighbors, we apply a clustering algorithm to assign pseudo-labels to all nodes. Based on these labels, we design a pseudo-label-guided contrastive objective that encourages higher embedding consistency between a node and its semantically similar neighbors. This semantic augmentation enables the model to better capture latent user intents and item characteristics that may not be explicitly observable from graph topology alone. We conduct extensive experiments on five real-world benchmark datasets to evaluate the effectiveness of NeSR. Experimental results demonstrate that our method achieves competitive performance across a wide range of metrics. Notably, on the Amazon-Kindle dataset, NeSR outperforms the second-best model by 11.79% in Recall@20 and 12.53% in NDCG@20, indicating substantial improvements in top-N recommendation accuracy. In summary, NeSR presents a comprehensive self-supervised learning framework that integrates subgraph generation via diffusion model, local structure modeling via random walks, and semantic enhancement via clustering. These components jointly improve the quality and robustness of node representations, ultimately leading to significantly enhanced recommendation performance.

This work was supported by the National Natural Science Foundation of China (Grant No. 62176236) and Key Project of Zhejiang Provincial Natural Science Foundation (Grant No. LZ24F030011).