

基于反事实用户行为生成的会话推荐方法

卢香葵^{1,3}, 邬俊^{1,2,3*}

(1. 北京交通大学计算机科学与技术学院, 北京 100044; 2. 数字化学习技术集成与应用教育部工程研究中心, 北京 100039;
3. 交通大数据与人工智能教育部重点实验室, 北京 100044)

摘要: 为保护用户隐私, 许多平台为用户提供了匿名登录选项, 迫使推荐系统仅能访问当前会话中的有限用户行为记录, 进而催生了会话推荐(Session-Based Recommendation, SBR)系统. 现有SBR技术在很大程度上沿用了传统非匿名用户行为建模思路, 聚焦于序列建模以习得会话表征. 然而, 当会话长度偏短时, 现有SBR技术性能衰减严重, 难以应对以短会话为主的真实会话推荐场景. 有鉴于此, 提出一种通过频繁模式引导长会话生成的反事实推理方法(Counterfactual inference by frequent pattern guided Long Session Generation, CLSG), 试图回答反事实问题: “如果会话内包含更丰富的交互物品, SBR模型预测结果将会如何?” CLSG遵循反事实理论的“归纳-行动-预测”经典三阶段推理流程. “归纳”: 从已观测会话集合中构建频繁模式知识库; “行动”: 基于所构建知识库生成反事实长会话; “预测”: 度量已观测会话和反事实会话预测结果间的差异, 并将其作为正则化项并入目标函数, 以达到表征一致性的目的. 值得注意的是, CLSG具有模型无关的技术特点, 可对现有SBR模型实现普惠式赋能. 三个基准数据集上的实验结果表明, CLSG提升了五款现有SBR模型的预测性能, 在命中率(Hit Rate, HR)和平均倒数排名(Mean Reciprocal Rank, MRR)评价指标上均取得6%左右的平均性能提升.

关键词: 会话推荐; 反事实推理; 频繁模式挖掘; 用户行为建模; 匿名会话

基金项目: 中央高校基本科研业务费重点项目(No.2023JBZY038); 北京市自然科学基金(No.L232033); 数字化学习技术集成与应用教育部工程研究中心创新基金重点项目(No.1321004)

中图分类号: TP181

文献标识码: A

文章编号: 0372-2112(2025)04-1264-15

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20240783

Counterfactual User Behavior Generation for Session-Based Recommendation

LU Xiang-kui^{1,3}, WU Jun^{1,2,3*}

(1. School of Computer Science and Technology, Beijing Jiaotong University, Beijing 100044, China;

2. Engineering Research Center of Integration and Application of Digital Learning Technology, Ministry of Education, Beijing 100039, China;

3. Key Laboratory of Big Data & Artificial Intelligence in Transportation, Ministry of Education, Beijing 100044, China)

Abstract: To protect user privacy, many platforms offer anonymous login options, limiting recommendation systems to accessing only user behavior records within the current session, thereby leading to the development of session-based recommendation (SBR). Existing SBR approaches mainly follow the traditional paradigms of non-anonymous user behavior modeling, focusing on learning session representations through sequential modeling. However, when sessions are short, the performance of these techniques drops significantly, making it challenging to address real-world SBR scenarios dominated by short sessions. To this end, we propose a method called counterfactual inference by frequent pattern guided long sequence generation (CLSG), which aims to answer the counterfactual question: “what would be the model’s prediction if the session contained richer interactions?” CLSG follows the classical three-stage counterfactual inference process of “induction-action-prediction”. The induction stage constructs a frequent pattern knowledge base from the observed session set. The action stage generates counterfactual long sessions with the guide of the knowledge base. The prediction stage measures the discrepancy between the predictions of the observed and counterfactual sessions, and incorporates such discrepancy as a regularization term into the objective function to achieve representation consistency. Notably, CLSG is model-agnos-

tic and can be easily applied to enhancing current SBR models. Experimental results on three benchmark datasets demonstrate that CLSG significantly improves the recommendation performance of five existing SBR models, with an average improvement of 6% in terms of both hit rate (HR) and mean reciprocal rank (MRR) metrics.

Key words: session-based recommendation; counterfactual inference; frequent patterns mining; user behavior modeling; anonymous sessions

Foundation Item(s): Fundamental Research Funds for the Central Universities (No.2023JBZY038); Natural Science Foundation of Beijing (No.L232033); Innovation Fund from the Engineering Research Center of Integration and Application of Digital Learning Technology, Ministry of Education (No.1321004)

1 引言

推荐系统作为对抗信息过载的有效技术手段,在电子商务^[1]、新闻资讯^[2]、影音视听^[3]等各类互联网平台中起到举足轻重的作用. 历经长久运营,平台积累了大量的用户历史行为,进而为用户行为建模(User Behavior Modeling, UBM)^[4,5]奠定了必要的数据库. 平台借助 UBM 能够实现对用户信息需求的细粒度感知,并为其提供个性化的信息服务. 然而,伴随互联网法律法规(如《中华人民共和国数据安全法》《互联网信息服务算法推荐管理规定》等)的日益完善,推荐系统在数据使用方面受到诸多限制. 为了保护用户隐私,许多平台为用户提供匿名登录选项,迫使推荐系统仅能访问当前会话中的有限用户行为记录,进而催生了会话推荐(Session-Based Recommendation, SBR)系统^[6,7]. 该技术通过会话表征学习推知匿名用户的即时信息需求,并为其提供短期个性化信息服务.

从方法论角度来看,现有 SBR 技术在很大程度上沿用了传统 UBM 的设计思路:通过序列建模习得会话表征,并据此预测匿名用户的“下一次点击”. 早期 SBR 方法^[8,9]使用循环神经网络(Recurrent Neural Networks, RNNs)从会话集合中提炼物品之间的时序变化特征;为了进一步捕获多跳依赖关系,一些研究^[10-13]将会话(即物品序列)转化为物品图,并借助图神经网络(Graph Neural Networks, GNNs)建模物品间的高阶协同关系;为了有效应对物品间的长程依赖问题,一些工作^[14-17]引入注意力机制从全局视角建模物品间的上下文依赖关系,达到了目前该领域的先进(State Of The Art, SOTA)性能.

然而,当会话长度偏短时,现有 SBR 技术性能衰减严重. 图 1 展示了三款经典 SBR 模型在 Tmall 数据集上的性能表现,评价指标为 TOP-10 命中率(HR@10). 这三款经典模型包括:基于 RNNs 的 NARM 模型(Neural Attentive Recommendation Machine)^[9]、基于 GNNs 的 SRGNN 模型(Session-based Recommendation with Graph Neural Networks)^[10]以及基于注意力机制的 CORE 模型(encoding session and item embeddings into a COnsistent REpresentation space)^[15]. 实验结果显示,相对于常规会话(交互物品数量>5),三款模型在短会话(交互物品数

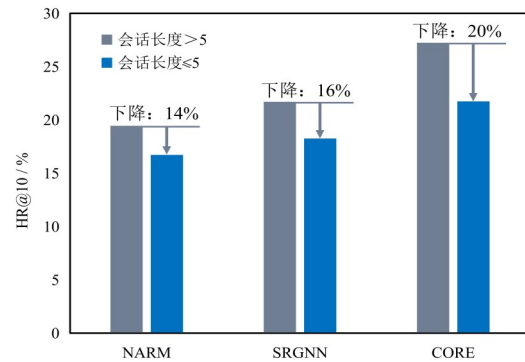


图 1 三款经典 SBR 模型在 Tmall 数据集上的性能表现

量≤5)上的性能下降十分明显,下降幅度分别为 14%、16% 和 20%. 造成该现象的主要原因可归结为,短会话中缺少足够的上下文信息,难以推测用户的真实意图. 例如,在电商场景下,当出现会话“果篮→鲜花”时,系统会大致推断用户意图为“挑选教师节礼物”;如果将该会话扩展为“尿裤→奶粉→果篮→鲜花”,系统便可精准感知用户意图为“探望产妇”. 值得注意的是,在传统 UBM 场景下,平台能够访问用户的历史行为记录,相应的行为序列比较长,如 ML-1M 数据集中序列的平均长度可达 165.5;相比之下,SBR 系统仅能访问当前会话,相应的行为序列(即会话)较短,如 Tmall 数据集中短会话数量过半,平均会话长度仅为 6.8. 因此,如何有效缓解短会话问题,是现有 SBR 技术能否突破瓶颈并走向成熟的关键.

有鉴于此,提出一种通过频繁模式引导长会话生成的反事实推理方法(Counterfactual inference by frequent pattern guided Long Session Generation, CLSG). 主张从反事实推理视角切入,试图回答反事实问题:“如果会话内包含更丰富的交互物品,SBR 模型预测结果将会如何?” CLSG 遵循反事实理论的“归纳-行动-预测”经典三阶段推理流程.(1)归纳:通过分析会话集合中物品间的序列依存规律,构建频繁模式知识库;(2)行动:依据知识库对已观测短会话施加干预,生成相应的反事实长会话;(3)预测:设计干预效应函数,度量已观测会话和反事实会话预测结果间的差异,并将其作为正则项并入目标函数,达到表征一致性的目的. CLSG 具有模型无关的技术特点,可

对现有形态各异的SBR模型实现普惠式赋能,从而促进SBR任务整体推荐性能的提升.三个基准数据集上的实验结果表明,CLSG显著提升五款现有SBR模型的预测性能,在命中率(Hit Rate, HR)和平均倒数排名(Mean Reciprocal Rank, MRR)两项评价指标上均取得了6%左右的平均性能提升.

2 相关工作

2.1 会话推荐

会话推荐(SBR)通过会话表征学习,预测匿名用户的下一次交互行为.早期研究利用循环神经网络(RNNs)捕获物品之间的时序依赖特征.GRU4Rec^[8]作为该类方法的先驱工作,率先将RNNs的变体门控循环单元(Gated Recurrent Unit, GRU)应用于会话表征学习;随后,NARM^[9]通过多层堆叠的GRU网络建模物品时序变化特征,同时借助基础注意力(vanilla attention)模块捕捉用户意图.为了进一步建模物品间的多跳近邻关系,图神经网络(GNNs)被引入SBR领域.SRGN^[10]作为开创性工作,利用门控GNNs学习会话的高阶依赖关系;随后,SGNN-HN(Star Graph Neural Networks with Highway Networks)^[18]利用星形图结构对会话进行建模,并集成高速公路网络以缓解图学习过拟合问题;DHCN(Dual channel Hypergraph Convolutional Networks)^[19]通过双通道超图卷积网络建模多个交互行为之间的高阶协同特征.为有效处理长程依赖问题,注意力机制被引入SBR领域.STAMP(Short-Term Attention/Memory Priority)^[14]提出短期注意力/记忆优先网络,利用基础注意力模块从会话整体捕获用户的一般兴趣,并从最近一次交互行为中捕捉用户的当前兴趣;CORE^[15]借助基于自注意机制的Transformer学习会话内物品的线性组合系数,捕捉交互行为全局依赖性的同时,保证会话表征与物品表征位于相同表示空间.ATTMIX(multi-level ATTention MIXture network)^[16]引入多头多层次注意力网络,实现对物品依赖关系的多层次推理.近年来,大语言模型(Large Language Models, LLMs)^[20,21]的兴起激发学者们将其应用于文本丰富的SBR场景^[22,23],LLMs以其卓越的文本分析和推理能力,为SBR提供更优的可解释性^[24]和迁移能力^[25]等,然而,基于LLMs的方法依赖于能同时提供文本信息和匿名交互行为序列的数据集,不具有普适性.

总体而言,尽管当前SBR方法在会话表征学习方面取得了重要进展,然而它们主要聚焦于序列建模,未从根本上解决SBR场景中的短会话问题.文献[26]针对传统UBM场景下冷启动用户短序列问题,提出一种基于逆向Transformer的序列延长方法.该方法借助一定数量的长会话训练反向Transformer,并据此为短序

列生成伪前置物品,以达到“取长补短”的效果.然而,SBR场景以短序列为主,长序列极度匮乏,导致该方法失效.此外,该方法受限于Transformer架构,难以灵活适配于现有形态各异的SBR模型.与此不同,所提CLSG通过频繁模式引导反事实长会话生成,避免了对大量已观测长会话的依赖,更适用于应对SBR场景中的短会话挑战,同时CLSG具备模型无关属性,可为各平台千差万别的SBR基模型实现普惠式赋能.

2.2 基于反事实推理的推荐

近年来,反事实推理(Counterfactual Inference)^[27]在推荐系统领域展现出了巨大的应用潜力^[28].依据所应对的问题类型,现有方法可大致分为三类:针对数据偏差的方法、针对数据缺失的方法,以及针对准确性以外指标(如可解释性、公平性、多样性等)的方法.针对数据偏差的方法^[29-31]通过消除观测数据中存在的流行度偏差、曝光偏差、位置偏差等混淆变量,实现数据去偏,从而学习到用户的真实偏好.例如,MACR(Model-Agnostic Counterfactual Reasoning)^[29]通过构建因果图,将物品属性与流行度对评分的影响进行解离;在推理阶段,MACR使用消除流行度效应的反事实因果图进行推理,有效削弱流行度的影响.针对数据缺失的反事实推理方法(Counterfactual Inference for Data Missing, CIDM)通过样本增广方式生成观测以外的反事实样本,实现对训练数据集的扩充.例如,CASR(Counterfactual data-Augmentation Sequential Recommendation)^[32]和CauseRec(Counterfactual user sequence synthesis for sequential Recommendation)^[33]分别通过不同的物品选择策略,对已观测历史交互序列中的物品进行替换,从而生成更多训练样本.值得注意的是,此过程中并未改变序列长度.针对准确性以外指标的方法借助因果图的透明性和可控性,提升模型可解释性、促进群体间公平性或增强推荐结果多样性等.例如,MCT(Maximum Causal Tree)^[34]已从观测数据中挖掘因果关系,并将其与现有推荐方法相结合,为推荐结果提供可解释性;UCRS(User Controllable Recommender System)^[35]提出因果增强的用户可控推理框架,允许用户主动调整推荐系统的行为,并利用反事实推理缓解信息茧房问题,从而提升推荐结果多样性;CBDF(Counterfactual Bandit with Delayed Feedback)^[36]利用反事实重要性采样模型重新加权已观测反馈并生成修改后的奖励,从而提升推荐结果的公平性.

所提出CLSG方法可归类为CIDM方法.然而,现有CIDM方法聚焦于训练样本总量不足问题,生成大量反事实样本,但无法解决SBR场景的短会话样本问题.与此不同,CLSG旨在生成包含更多交互行为的反事实长会话,以缓解短会话挑战.

3 方法设计

本文所提出的方法的设计思路如图2所示. 首先, 基于因果推理相关理论重述 SBR 任务; 然后, 引入 CLSG 正则项, 对现有 SBR 模型进行改进, 以应对短会话挑战; 最后, 通过“归纳-行动-预测”三阶段流程实

现 CLSG 方法, 其中“归纳”以已观测会话集合为输入, 通过频繁模式挖掘, 构建频繁模式知识库;“行动”以已观测会话为输入, 通过频繁模式引导的会话干预, 生成反事实长会话;“预测”以已观测会话和反事实会话为输入, 通过 RBO 度量两者的预测排序列表差异, 获得干预效应值. 所涉及的重要符号及其说明如表1所示.

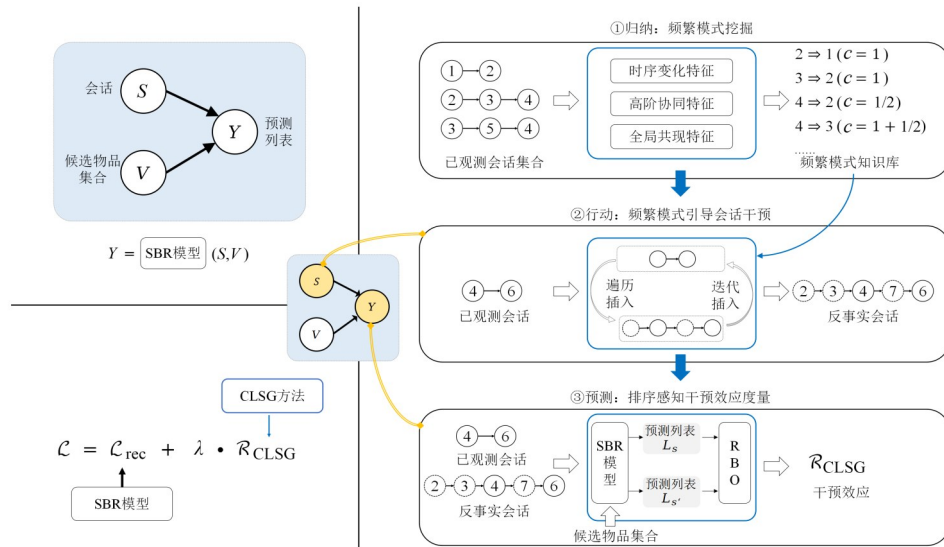


图2 基于反事实用户行为生成的会话推荐方法

表1 符号及其描述

| 符号 | 描述 |
|-----------|------------------------|
| M | 结构因果模型 |
| G, F | 因果图, 结构方程组 |
| S, V, Y | 节点变量: 会话, 候选物品集合, 预测列表 |
| e | 干预效应值 |
| S, V | 会话集合, 物品集合 |
| m, n, t | 会话总数, 物品总数, 会话长度 |
| s, v | 单个会话, 单个物品 |

3.1 预备知识

结构因果模型 (Structure Causal Model) 结构因果模型 M 由因果图 G 和结构方程组 F 构成. G 是一个有向无环图, 图中的节点表示变量, 连接节点的边表示变量间的因果关系. 依据因果图, 可以构造对应结构方程组 F , 其中每个方程表示某一变量由其父节点变量推导而来的具体过程.

反事实推理 (Counterfactual Inference) 假设因果图 G 中存在两个节点 X 和 Y , 且 X 是 Y 的祖先节点, 那么基于 X 的观测值 x 和 Y 的观测值 y , 反事实推理旨在回答以下反事实问题: “如果 X 的值被设定为 x' , 那么 Y 的值将会如何?” 反事实推理包括“归纳-行动-预测”经典三阶段流程.

(1) 归纳 (abduction). 从已观测数据中分析潜在规

律或挖掘潜在模式, 为后续的行动和预测奠定基础.

(2) 行动 (action). 对节点 X 实施干预 (intervention), 获得反事实值 $X=x'$.

(3) 预测 (prediction). 计算 $X=x'$ 的反事实结果 Y , 并度量干预所产生的因果效应 (简称干预效应).

干预效应 (Intervention Effect) 度量一个变量对另一个变量的影响程度, 具体表现为当祖先节点 X 发生改变时, 目标节点 Y 所产生的变化量. 不失一般性, 该变化量可形式化定义为 $e = \varphi(Y_x, Y_{x'})$, 其中 Y_x 和 $Y_{x'}$ 分别表示祖先节点在不同取值下 ($X=x$ 和 $X=x'$), 目标节点 Y 的取值, 而函数 $\varphi(\cdot)$ 用于量化这两个值的差异, 其具体形式可依据实际需求进行设计与选择.

3.2 因果视角下的 SBR 任务

在因果视角下重述 SBR 任务是开展反事实推理的基础. SBR 任务的结构因果模型表示为 $M = \{G, F\}$, 因果图 G 中包含三个节点变量 S, V, Y , 分别表示会话、候选物品集合和预测排序列表, 三者的因果关系为 $Y = F(S, V)$, 其中结构方程组 F 可由任意 SBR 模型实现, 如 NARM^[9]、SRGNN^[10] 等. 变量 S 的所有观测值构成了会话集合 $\mathcal{S} = \{s_1, s_2, \dots, s_m\}$, 其中 m 表示会话总数, 每个会话 $s \in \mathcal{S}$ 是由用户交互的物品按照交互时间顺序排列而成的序列, 表示为 $s = \langle v_{s,1}, v_{s,2}, \dots, v_{s,t} \rangle$, t 为会话长度. 变量 V 通常赋值为全量物品集合 $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$, 其中 n 表示物品总数.

给定 $S=s, V=v$, 通过 F 可获得每个候选物品 v_i 相对于会话 s 的推荐概率 $\hat{y}_{s,i}$, 依据推荐概率对物品进行降序排序, 即可获得预测排序列表 $Y=L_s$. 最终, L_s 中排序最靠前 (TOP-1) 的物品将作为用户可能交互的下一个物品 $v_{s,t+1}$ 进行推荐. SBR 任务的目标函数通常采用二元交叉熵 (Binary Cross Entropy, BCE), 表示如下:

$$\mathcal{L}_{\text{rec}} = - \sum_{i=1}^n (y_{s,i} \log(\hat{y}_{s,i}) + (1 - y_{s,i}) \log(1 - \hat{y}_{s,i})) \quad (1)$$

其中, $\hat{\mathbf{y}}_s = [\hat{y}_{s,1}, \hat{y}_{s,2}, \dots, \hat{y}_{s,n}]$ 表示推荐概率向量, $\mathbf{y}_s = [y_{s,1}, y_{s,2}, \dots, y_{s,n}]$ 表示真实标签的独热编码 (one-hot) 向量, $y_{s,i} = 1$ 表示物品 v_i 是用户真实交互的下一个物品, 否则 $y_{s,i} = 0$.

由于 SBR 系统仅能访问当前会话, 变量 S 的观测值 (即会话) 普遍较短, 对应观测结果 $Y=L_s$ 在 SBR 任务上表现不佳. 由此引出反事实问题: “如果会话内包含更丰富的交互物品, SBR 模型预测结果将会如何?” 为回答该问题, 提出一种通过频繁模式引导长会话生成的反事实推理方法 CLSG. 该方法以正则化项的形式并入现有 SBR 模型的目标函数中, 促进表征学习一致性, 最终的优化目标函数为

$$\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \cdot \mathcal{R}_{\text{CLSG}} \quad (2)$$

其中, λ 为超参数, 用于调整正则项权重.

3.3 CLSG 方法

CLSG 遵循反事实理论的“归纳-行动-预测”经典三阶段推理流程. 接下来, 将对三个阶段展开介绍.

3.3.1 归纳: 频繁模式挖掘

该阶段旨在利用频繁模式挖掘^[37]技术, 从已观测会话集合 \mathcal{S} 中分析物品间的序列依存规律, 从而为反事实长会话生成提供可靠的先验知识. 该过程借鉴经典的 ERMiner (Equivalence class based sequential Rule Miner)^[38]算法, 并针对 SBR 场景下的数据特征进行必要的修改. 具体而言, 考虑到 SBR 场景下的会话普遍较短, 将传统算法中两个物品集合之间的频繁模式简化为两个物品之间的频繁模式 (即二项模式). 此外, 为了更全面地捕获用户复杂多样的行为模式, 在挖掘过程中综合考虑了用户交互行为中的三个关键特征: 时序变化特征、高阶协同特征和全局共现特征. 其中, 时序变化特征聚焦于交互行为间的顺序依赖关系, 体现为单个会话中相邻物品的交互先后顺序; 高阶协同特征则进一步探索交互行为间的高阶依赖关系, 体现为单个会话中交互物品之间的距离跳数; 全局共现特征从会话集合视角出发, 考虑多个交互行为在整个会话集合中的共现频次.

挖掘频繁二项模式的具体过程如下:

(1) 候选模式生成. 从会话集合 \mathcal{S} 中挖掘所有潜在的二项模式, 该过程体现交互行为之间的时序变化特征. 以

会话 $s = \langle v_{s,1}, v_{s,2}, \dots, v_{s,t} \rangle$ 为例, 遍历会话内物品 $v_{s,i}$ ($i \in \{2, t\}$), 可构建 $v_{s,i}$ 与其所有前置物品的频繁二项模式, 即 $\{v_{s,i} \Rightarrow v_{s,1}, \dots, v_{s,i} \Rightarrow v_{s,i-1}\}$, 其中模式 $v_{s,i} \Rightarrow v_{s,i-k}$ 表示右部 $v_{s,i-k}$ 是左部 $v_{s,i}$ 的前置物品.

(2) 局部支持度计算. 通过量化二项模式中两个物品在同一会话中的距离跳数, 衡量物品之间的高阶协同特征强度. 这一量化结果也被称为该频繁模式在当前会话中的局部支持度. 具体而言, 在 s 中, 候选模式 $v_{s,i} \Rightarrow v_{s,i-k}$ 的局部支持度为 $c_s(v_{s,i} \Rightarrow v_{s,i-k}) = 1/k$, 其中 k 为物品 $v_{s,i}$ 与物品 $v_{s,i-k}$ 在会话 s 中的距离跳数.

(3) 全局支持度计算. 通过计算二项模式在整个会话集合 \mathcal{S} 中的共现频次, 量化物品全局共现特征强度, 这一量化结果也被称为该频繁模式在会话集合 \mathcal{S} 中的全局支持度. 具体而言, 候选模式 $v_i \Rightarrow v_j$ 的全局支持度为 $c(v_i \Rightarrow v_j) = \sum_{s \in \mathcal{S}} c_s(v_i \Rightarrow v_j)$, 其中 $c_s(v_i \Rightarrow v_j)$ 为候选模式 $v_i \Rightarrow v_j$ 在会话 s 中的局部支持度, 当 s 中不存在该模式时, $c_s(v_i \Rightarrow v_j) = 0$.

(4) 全局支持度调整. 鉴于热门物品可能导致相应频繁模式的全局支持度偏高, 采用流行度惩罚机制调整全局支持度. 具体而言, 物品 v_j 的流行度惩罚因子定义为 $p_{v_j} = 1/\log(1 + N(v_j))$, 其中 $N(v_j)$ 表示物品 v_j 在数据集 \mathcal{S} 中被交互的次数. 经调整, 模式 $v_i \Rightarrow v_j$ 最终支持度为: $\tilde{c}(v_i \Rightarrow v_j) = p_{v_j} \times c(v_i \Rightarrow v_j)$. 通过上述过程, 可以系统地挖掘会话集合 \mathcal{S} 中的所有频繁模式及其支持度, 作为先验知识库 \mathcal{K} 引导后续反事实“行动”阶段.

3.3.2 行动: 频繁模式引导会话干预

该阶段以先验知识库 \mathcal{K} 为引导, 对已观测会话 s 实施干预, 生成反事实长会话 s' . 该过程伪代码如算法 1 所示, 其中函数 $\text{len}(s)$ 用于计算会话 s 包含的交互数 (即会话长度); $\text{query}(q, \mathcal{K}, \text{des})$ 用于从 \mathcal{K} 中检索所有以物品 q 为左部的频繁模式, 并按照其支持度进行降序排序, 最终输出检索结果中的所有右部元素, 构成物品 q 的候选前置物品集合; $\text{merge}(L_1, L_2)$ 用于将两个列表按顺序合并.

算法 1 对已观测会话集合 \mathcal{S} 中的每一个会话 s 进行多轮迭代式干预, 直到达到预设的最大迭代轮次 M_E 或所生成的反事实会话达到预设的会话长度阈值 M_L (对应算法 1 第 1~5 行). 对于每一个会话 $s = \langle v_{s,1}, v_{s,2}, \dots, v_{s,t} \rangle$, 依次以会话内的物品 $v_{s,i}$ 作为查询, 从 \mathcal{K} 中检索其前置物品集合中的“遍历插入”及算法 1 第 6 行). 若查询物品位于会话首部 (即 $v_{s,1}$), 则直接返回支持度最高的频繁模式 $v_{s,i} \Rightarrow v_c$ 所对应前置物品 v_c , 并将其插入 $v_{s,i}$ 之前, 获得反事实会话 $s' = \langle v_c, v_{s,1}, v_{s,2}, \dots, v_{s,t} \rangle$ (对应算法 1 第 7、第 8 行); 若查询物品 $v_{s,i}$ 位于会话中部, 需进行双重判断, 要求候选物品 v 既是 $v_{s,i}$ 的前置物品 (即存在频繁模式 $v_{s,i} \Rightarrow v$), 也是 $v_{s,i-1}$ 的后置物品 (即存在频繁模式

算法 1 频繁模式引导的会话干预

输入: 已观测会话集合 \mathcal{S} , 频繁模式知识库 \mathcal{K} , 最大迭代轮次 M_E , 最大会话长度 M_L

输出: 反事实会话集合 \mathcal{S}'

```

1. FOR  $s$  IN  $\mathcal{S}$ 
2.    $s' = s$ 
3.   WHILE  $i$  IN RANGE( $M_E$ )//迭代插入
4.      $idx = 0$ 
5.     WHILE  $idx < \text{len}(s')$  AND  $\text{len}(s') < M_L$ 
6.        $D_1 = \text{query}(s'[idx], \mathcal{K}, \text{des})$ 
7.       IF  $idx == 0$  //查询物品位于首部
8.          $s' = \text{merge}(D_1[0], s')$ 
9.       ELSE //查询物品位于中部
10.        FOR  $v$  IN  $D_1$ 
11.           $D_2 = \text{query}(v, \mathcal{K}, \text{des})$ 
12.          IF  $s'[idx - 1]$  IN  $D_2$ 
13.             $\tilde{s} = \text{merge}(s'[idx], v)$ 
14.             $s' = \text{merge}(\tilde{s}, s'[idx:])$ 
15.          BREAK
16.        ELSE
17.          CONTINUE
18.       $idx += 2$ 
19.    $\mathcal{S}'.\text{append}(s')$ 

```

$v \Rightarrow v_{s,i-1}$), 在满足上述条件的 v 中返回支持度最高的频繁模式 $v_{s,i} \Rightarrow v_c$ 对应的前置物品 v_c , 插入会话中物品 $v_{s,i-1}$ 与物品 $v_{s,i}$ 之间, 获得反事实会话 $s' = \langle \dots, v_{s,i-1}, v_c, v_{s,i}, \dots \rangle$ (对应算法 1 第 9~17 行). 对 s 中的所有物品遍历完毕后, 将获得的反事实会话 s' 作为新的待插入会话, 通过上述步骤进行迭代式会话干预 (对应图 2(c) 中的“迭代插入”), 直到达到停止条件, 输出最终的 s' . 对于每一个会话可获得一个对应的反事实会话, 构成反事实会话集合 \mathcal{S}' . 将其与已观测会话集合 \mathcal{S} 一同送入反事实“预测”阶段, 度量干预效应.

3.3.3 预测: 排序感知干预效应度量

该阶段旨在利用排序感知因果效应函数, 量化会话 s 及其反事实会话 s' 的预测列表差异.

如图 2 所示, 给定 s 和 s' , 分别将它们输入 SBR 模型 F 中, 获得对应的预测排序列表: $L_s = F(s, \mathcal{V})$, $L_{s'} = F(s', \mathcal{V})$, 其中 \mathcal{V} 为候选物品集合. 度量预测结果差异的通用公式为 $e = \varphi(Y_x, Y_{x'})$. 在 SBR 排序学习任务中, $Y_x = L_s, Y_{x'} = L_{s'}$. 由于 $L_s, L_{s'}$ 为物品排序列表, 差异性函数 $\varphi(\cdot, \cdot)$ 的设计需具备排序感知特性. 受到文献 [39] 的启发, 引入排序偏置重叠 (Rank-Biased Overlap, RBO) 指标, 度量两个列表在 TOP- K 排序位置上的一致性, 其负值即可表示差异性. RBO 指标定义如下:

$$R(L_s, L_{s'}, K) = \sum_{k=1}^K p^{\frac{k}{q}} \frac{|L_s[1:k] \cap L_{s'}[1:k]|}{k} \quad (3)$$

其中, 缩放因子 q 默认值为 10, 排序折扣因子 p 为超参数, 用于调整不同排序位置的权重. 考虑到推荐系统通常仅关注头部排序结果, 超参数 K 的取值远小于列表总长度 n , 从而将排序一致性问题聚焦于 TOP- K 物品, 既满足实际应用需求, 又可降低计算复杂度. $R(L_s, L_{s'}, K)$ 函数旨在计算排序列表 L_s 和 $L_{s'}$ 在一系列 k 值 (k 依次取 $[1, 2, \dots, K]$) 下的 TOP- k 子列表交集元素数量 (即重叠度) 的加权平均值, 该值越大, 表明两个列表中物品的排序一致性越高. 为便于后续论述, 将式 (3) 中两个排序列表 TOP- k 重叠度表示如下:

$$O(L_s, L_{s'}, k) = |L_s[1:k] \cap L_{s'}[1:k]| \quad (4)$$

尽管 RBO 指标可以有效度量两个排序列表的一致性, 但直接将其作为正则化项并入模型优化过程却面临挑战. 这一挑战主要源自式 (4) 中涉及的集合交集运算, 该运算既不可求导也不可微分, 从而无法适用于依赖梯度信息的优化算法, 如广泛应用于现有 SBR 模型训练的梯度下降法. 换言之, “集合交运算不可梯度优化” 成为了 CLSG 无缝集成至现有 SBR 模型中的主要障碍.

有鉴于此, 提出一种近似算法, 将集合运算转化为向量计算, 使模型可通过梯度下降法实现端到端训练. 具体而言, 首先确定推荐列表的 TOP- k 边界, 即求取推荐概率向量 \hat{y}_s 和 $\hat{y}_{s'}$ 中排在第 k 位的预测分数值 (从大到小排序) \hat{y}_s^k 和 $\hat{y}_{s'}^k$. 随后, 通过比较每个物品的预测分数值与对应阈值 \hat{y}_s^k 或 $\hat{y}_{s'}^k$, 判断该物品是否被包含在 TOP- k 排序列表中:

$$\begin{aligned} \mathbf{d}_s &= [\hat{y}_{s,1}, \hat{y}_{s,2}, \dots, \hat{y}_{s,n}] - [\hat{y}_s^k, \hat{y}_s^k, \dots, \hat{y}_s^k] \\ \mathbf{d}_{s'} &= [\hat{y}_{s',1}, \hat{y}_{s',2}, \dots, \hat{y}_{s',n}] - [\hat{y}_{s'}^k, \hat{y}_{s'}^k, \dots, \hat{y}_{s'}^k] \\ \tilde{\mathbf{d}}_s &= [I(d_{s,1} > 0), I(d_{s,2} > 0), \dots, I(d_{s,n} > 0)] \\ \tilde{\mathbf{d}}_{s'} &= [I(d_{s',1} > 0), I(d_{s',2} > 0), \dots, I(d_{s',n} > 0)] \end{aligned} \quad (5)$$

其中, $I(\cdot)$ 表示指示函数, 当括号内条件成立时取值为 1, 否则为 0. 由于 $I(\cdot)$ 不可导, 采用带温度系数 τ 的 Sigmoid 函数 $\sigma(x; \tau) = 1/(1 + e^{-x\tau})$ 对指示函数 $I(\cdot)$ 进行平滑近似, 可得:

$$\begin{aligned} \tilde{\mathbf{d}}_s &\approx [\sigma(d_{s,1}; \tau), \sigma(d_{s,2}; \tau), \dots, \sigma(d_{s,n}; \tau)] \\ \tilde{\mathbf{d}}_{s'} &\approx [\sigma(d_{s',1}; \tau), \sigma(d_{s',2}; \tau), \dots, \sigma(d_{s',n}; \tau)] \end{aligned} \quad (6)$$

当候选物品 v_i 属于 s 的 TOP- k 排序列表, $\tilde{\mathbf{d}}_s$ 的第 i 维近似为 1, 否则近似为 0. 最终, 式 (4) 中两个列表的交集运算可以近似为对应向量的点积运算:

$$O(L_s, L_{s'}, k) \approx \tilde{\mathbf{d}}_s \cdot \tilde{\mathbf{d}}_{s'} = \sum_{i=1}^n \tilde{d}_{s,i} \cdot \tilde{d}_{s',i} \quad (7)$$

至此, 由 RBO 定义的排序感知干预效应函数如下:

$$\mathcal{R}_{\text{CLSG}} = - \sum_{k=1}^K p^{\frac{k}{q}} \frac{\sigma(\hat{y}_s - \hat{y}_{s'}^k; \tau) \cdot \sigma(\hat{y}_{s'} - \hat{y}_s^k; \tau)}{k} \quad (8)$$

3.3.4 时间复杂度分析

CLSG 总体时间复杂度由三个阶段的复杂度共同构

成,表示为 $O(\text{CLSG})=O_1+O_2+O_3$ 。“归纳”阶段对已观测会话集合进行两次遍历,一次遍历用于获取物品流行度,一次遍历用于获取频繁模式及其支持度.上述操作在训练过程中仅执行一次,因此,该阶段时间复杂度为 $O_1=2\times m\approx O(m)$,其中 m 为集合 S 所包含的会话总数.“行动”阶段对集合进行一次遍历,旨在对集合中每一个会话生成反事实长会话,此过程时间复杂度为 $O_2=m\times M_E\times M_L\times O(1)\approx O(m)$,其中 M_E 和 M_L 分别表示最大迭代轮次和最大会话长度, $O(1)$ 为从 K 中检索前置物品的时间复杂度.“预测”阶段利用现有SBR模型对集合中的样本依次执行排序列表预测与列表差异计算,该阶段的时间复杂度为 $O_3=2\times O_F+K\times m\times O(n)\approx O_F+O(m\times n)$,其中, O_F 为SBR模型的预测时间复杂度, K 为一致性函数所关注的列表顶端位置值, n 为候选物品总数.综合上述三阶段时间复杂度,CLSG总体时间复杂度为 $O(\text{CLSG})=O(m)+O(m)+O_F+O(m\times n)$,化简后可得到 $O(\text{CLSG})\approx O_F+O(m\times n)$.总体而言,CLSG的时间复杂度取决于SBR模型的时间复杂度:当 $O_F<O(m\times n)$ 时,CLSG是会话总数及物品总数的线性时间复杂度;当 $O_F\geq O(m\times n)$ 时,CLSG与SBR模型为等量级时间复杂度.

4 实验

首先介绍数据集、对比方法、评价指标等实验设置,然后通过大量实验尝试回答如下5个研究问题,以全面评估所提出CLSG方法.

研究问题1:CLSG对现有SBR模型提升如何?

研究问题2:CLSG在短会话上表现如何?

研究问题3:CLSG较现有CIDM方法如何?

研究问题4:CLSG如何改进会话推荐性能?

研究问题5:超参数设置对CLSG影响如何?

4.1 实验设置

4.1.1 数据集

在3个SBR基准数据集上评估CLSG方法性能.

(1)Diginetica.该数据集记录了Diginetica平台上用户的交易历史,数据来源于2016年CIKM竞赛.

(2)Nowplaying.该数据集记录了Twitter平台用户的音乐收听行为,数据来源于文献[40].

(3)Tmall.该数据集记录了Tmall平台上用户的在线购物日志,数据来源于2015年IJCAI竞赛.

为确保实验公平性,采用现有工作^[10,15,16]中普遍使用的数据预处理手段.首先,移除数据集中交互次数少于5次的冷物品,并对长度超过20的会话进行截断,仅保留最近交互的20个物品.然后,以交互时间为参考,选取最近一周的交互行为构成测试集,其余交互行为构成训练集.最后,为充分利用序列信息,对训练集和测试集中的每个会话进行序列分割处理,具体

而言,给定会话 $s=\langle v_{s,1},v_{s,2},\dots,v_{s,t}\rangle$,依次按位分割,可生成一系列子会话及其对应标签,即 $\{(\langle v_{s,1}\rangle,v_{s,2}),\dots,(\langle v_{s,1},\dots,v_{s,t-1}\rangle,v_{s,t})\}$,从而丰富训练数据和测试数据,该做法的有效性已在诸多研究^[14,41]中得以验证.经过数据预处理后的数据集统计信息详见表2.

表2 数据集统计信息

| | Diginetica | Nowplaying | Tmall |
|--------|------------|------------|-----------|
| 交互数量 | 981 620 | 5 900 621 | 2 863 467 |
| 会话数量 | 204 591 | 904 439 | 366 556 |
| 物品数量 | 42 596 | 60 416 | 40 719 |
| 会话平均长度 | 3.80 | 5.52 | 6.81 |
| 短会话占比 | 78% | 64% | 55% |

4.1.2 对比方法

为验证CLSG的有效性,选取两组对比方法:SBR基线方法和CIDM基线方法.

SBR基线方法CLSG被设计为即插即用模块,可与现有SBR模型组合使用,以缓解其所面对的短会话挑战.为验证CLSG的通用性和有效性,选取五款基于不同深度神经网络架构的主流SBR模型作为基线,以比较各基线模型与其CLSG升级版本(记为X+CLSG,其中X为基线模型)的性能.5款基线模型如下:

(1)NARM^[9].利用多层堆叠GRU建模物品时序变化特征,并借助基础注意力模块捕捉用户意图.

(2)STAMP^[14].提出短期注意力/记忆优先网络,利用基础注意力模块从完整会话中捕获用户的一般兴趣,并从最近一次交互行为中捕获用户的当前兴趣.两者融合作为最终的会话表征.

(3)SRGNN^[10].引入门控GNNs,建模会话图中物品间的多跳近邻关系,以捕捉物品高阶协同特征.

(4)CORE^[15].通过基于自注意力的Transformer^[42]模块学习物品的线性组合系数,以获得会话表征.

(5)ATTMIX^[16].设计多头多层次注意力混合网络,建模物品间的多层次依赖关系.值得注意的是,CLSG聚焦于会话内交互行为数量稀缺问题,因此选择仅建模交互行为(纯ID)的SBR基线.当数据集存在边信息(如文本、图像)时,可扩展至利用边信息的模型.

CIDM基线方法CLSG可归类为针对数据缺失的反事实推理(CIDM)方法,为验证其先进性,选取2款现有CIDM方法作为对比:

(1)CASR^[32].对已观测历史交互序列中的物品进行替换生成反事实样本以扩充训练集,其中替换物品的选择策略包括启发式方法、模型引导的方法和基于强化学习的方法,原文实验结果表明,模型引导的方法在多数情况下表现最优,因此本实验部分采用该策略实现CASR算法.

(2)CauseRec^[33].从已观测交互序列中识别出不可

或缺和可有可无的概念(即物品或偏好嵌入),以此为依据对序列中的物品进行替换,生成反事实样本,作为已观测交互序列的正样本或负样本参与对比学习.为公平比较,所有方法中的用户编码器均使用上述5款SBR基线模型.

4.1.3 评价指标

与现有工作^[14,16]相同,采用SBR的基准评测指标:命中率 $HR@K$ (Hit Rate)和平均倒数排名 $MRR@K$ (Mean Reciprocal Rank)评价各方法性能.其中, $HR@K$ 计算用户真实交互物品在预测列表TOP-K项中的概率,用于评价模型的推荐准确性. $MRR@K$ 则进一步计算用户真实交互物品在推荐列表中排序位置倒数的平均值,排序越靠前, $MRR@K$ 值越高,当排序位置大于 K 时, $MRR@K$ 值为0.实验中, K 值设置为5和10.

4.1.4 实现细节

所有方法均采用网格搜索寻找最佳超参数,其中SBR基线模型的超参数搜索范围参考原文建议.在CLSG方法中,正则项系数(λ)搜索范围为 $\{0, 0.1, 0.2, 0.3, 0.4\}$,排序折扣因子(ρ)搜索范围为 $\{0.9, 0.8, 0.7, 0.6\}$,温度系数(τ)搜索范围为 $\{100, 1000\}$.列表一致性函数中 K 取值为100,为提高计算效率,从中采样一个子集进行计算,即 $k \in \{20, 30, 40, 50, 60, 70, 80, 90, 100\}$,最大会话长度 M_L 设置为20,最大迭代轮次 M_E 设置为20.获得最佳超参数后,采用不同随机种子进行三次独立实验,并取平均值作为最终结果.

4.2 对现有SBR模型的提升

表3展示了5款SBR基线与其升级版(X+CLSG)在三个数据集上的性能比较,其中“提升比”以粗体表示,该指标反映CLSG升级版相对于原基线模型的性能改进,“*”表示统计结果显著(采用成对样本 t 检验,显著性水平设置为0.05).总体实验结果表明,所有CLSG升级模型在三个数据集和两个评价指标上的性能表现均好于相应基线模型,验证了CLSG方法的有效性和通用性.具体观测结论如下:

(1)CLSG对基于RNNs和基础注意力的基线(NARM、STAMP)性能提升尤为显著,在HR和MRR指标上分别取得了10.78%和11.23%的平均性能提升.NARM和STAMP均采用基础注意力机制,该机制在候选项(会话中交互物品)较少时效果不明显,而CLSG通过引入反事实长会话,为注意力计算提供更多候选项,因此获得大幅性能提升.

(2)CLSG对基于GNNs的基线(SRGNN)提升效果也比较明显,在HR和MRR指标上分别取得了4.47%和2.77%的平均性能提升.CLSG借助反事实长会话进一步丰富了图结构,使得GNNs能够捕获更多物品关联,进而获得较明显的性能改进.

(3)即使对于两款较强的基于复杂注意力机制(CORE、ATTMIX)的基线模型,CLSG仍然可进一步改善其性能,在HR和MRR指标上分别取得了2.20%和2.39%的平均性能提升.CLSG通过生成反事实长会话,为复杂注意力模块提供了更丰富的交互上下文,故而模型的预测性能被进一步优化.

总体而言,无论基线模型采用简单或复杂的神网络结构,CLSG方法均能为其带来进一步的性能改善(在HR和MRR指标上均取得了6%左右的平均性能提升).该观测结果表明,CLSG可有效缓解短会话问题,为各类既有SBR模型提供普惠式赋能.

4.3 在短会话上的表现

为了进一步验证CLSG方法是否有效提升了现有SBR模型在短会话上的性能,从各个数据集的测试集中挑选出长度 ≤ 5 的会话,构建短会话测试集.图3展示了5款SBR基线模型与其升级版(X+CLSG)在短会话测试集上的性能表现.具体而言,在HR和MRR指标上,CLSG分别实现了4.98%和4.05%的平均性能增益.同时,所生成的反事实长会话平均长度达到18,有效缓解了短会话问题.更多实验观察如下:

(1)在部分情况下,本文所提出的CLSG在短会话上的提升比略低于在所有会话(见表3)上的提升比,其主要原因如下:在用户行为建模任务中,序列越长越有利于推荐系统准确地捕捉用户偏好,进而获得更优推荐性能;反之,短序列中的用户行为有限,模型难以充分学习用户表征,导致性能不佳.因此,对于长会话,CLSG可以与其既有长序列优势相结合,从而实现更显著的性能提升;对于短会话,CLSG可以应对现有SBR模型表现不足问题,提升其在短会话上的推荐性能.所有会话集合中同时包含短会话和长会话,因此,在部分情况下,在所有会话上的提升比略优于在短会话上的提升比.

(2)在Nowplaying数据集上,SRGNN+CLSG的性能提升未达到预期.这一现象由数据集特点和模型特点两方面因素共同决定.在数据集方面,Nowplaying作为音乐数据集,其中的近邻交互行为相较于电商数据集(如Diginetica、Tmall等)具有更强的依赖性,主要表现为用户倾向于连续播放相同风格的歌曲.使用CLSG对已观测会话进行干预后,近邻依赖受到一定破坏;在模型方面,相较于使用注意力机制通过权重灵活分配来聚焦用户核心兴趣的其他基线模型,SRGNN依赖于近邻传播机制进行建模,难以从反事实长会话中重建原始的近邻关系.因此,在数据集和模型两方面因素的共同作用下,CLSG出现了部分情况下的失效.如何在反事实会话中更好地保持和重建已观测会话中的近邻关系,以使CLSG更适用于强近邻依赖性的推荐场景及基模型,将在未来工作中探索.

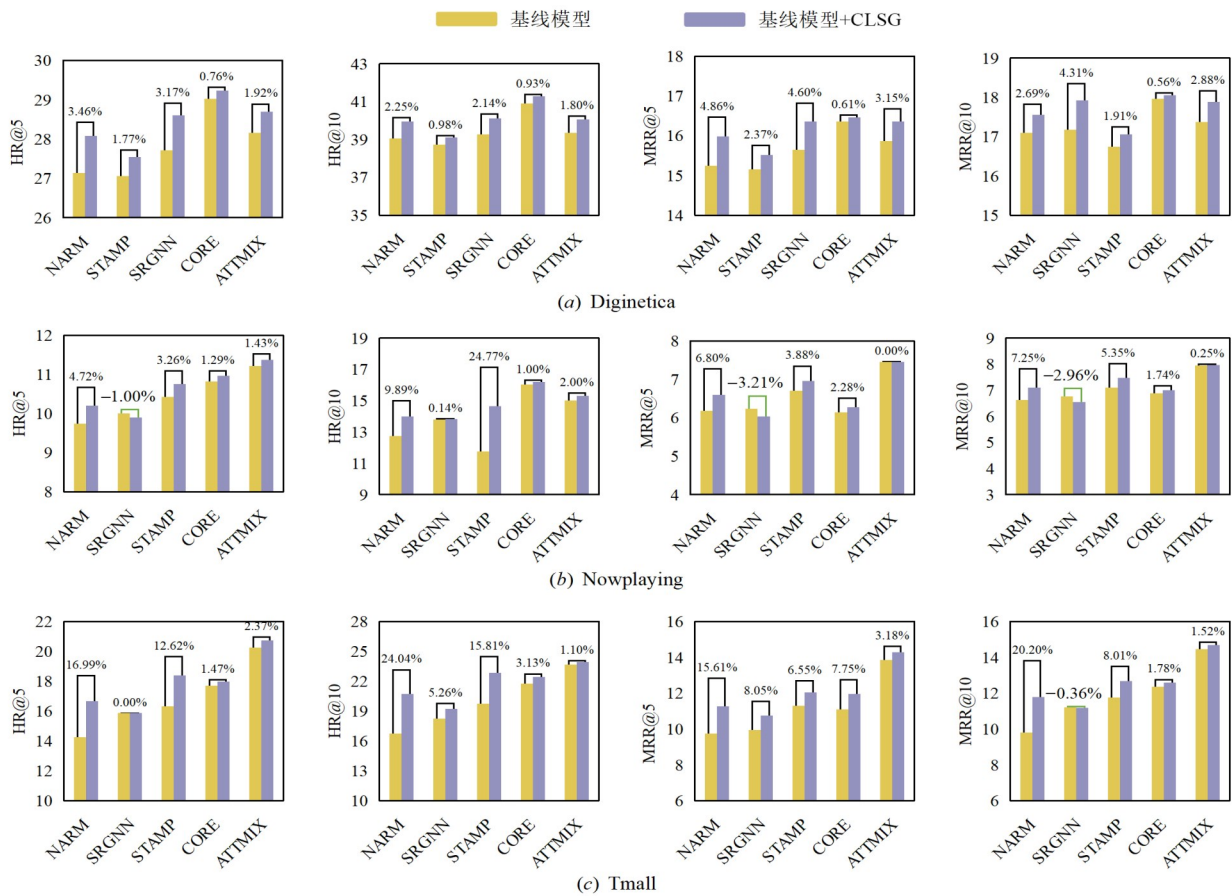


图3 CLSG在短会话上的性能表现

表3 基线模型与其CLSG升级版(即X+CLSG)在基准数据集上的性能比较

| | Diginetica | | | | Nowplaying | | | | Tmall | | | |
|--------|------------|--------|--------|--------|------------|--------|-------|--------|--------|--------|--------|--------|
| | HR@5 | HR@10 | MRR@5 | MRR@10 | HR@5 | HR@10 | MRR@5 | MRR@10 | HR@5 | HR@10 | MRR@5 | MRR@10 |
| NARM | 26.69 | 38.40 | 14.97 | 16.52 | 10.31 | 14.07 | 6.62 | 7.12 | 16.72 | 20.65 | 11.21 | 11.73 |
| +CLSG | 27.34* | 38.87 | 15.48* | 17.00* | 10.50* | 14.22 | 6.86* | 7.35* | 21.02* | 26.11* | 13.97* | 14.65* |
| 提升比 | 2.45% | 1.22% | 3.39% | 2.89% | 1.88% | 1.07% | 3.52% | 3.17% | 25.70% | 26.46% | 24.59% | 24.89% |
| STAMP | 26.35 | 37.72 | 14.73 | 16.24 | 10.72 | 14.39 | 6.98 | 7.46 | 18.11 | 22.49 | 11.88 | 12.47 |
| +CLSG | 26.75* | 38.25 | 15.03* | 16.55* | 11.03* | 14.89* | 7.22* | 7.73* | 23.32* | 29.82* | 15.47* | 16.34* |
| 提升比 | 1.53% | 1.40% | 2.01% | 1.95% | 2.86% | 3.49% | 3.50% | 3.62% | 28.74% | 32.60% | 30.23% | 31.05% |
| SRGNN | 27.20 | 38.60 | 15.26 | 16.77 | 10.23 | 14.02 | 6.39 | 6.89 | 18.86 | 22.56 | 13.12 | 13.61 |
| +CLSG | 28.19* | 39.60* | 16.00* | 17.52* | 10.45 | 14.39* | 6.40 | 6.92 | 20.09* | 24.66* | 13.51* | 14.13* |
| 提升比 | 3.65% | 2.60% | 4.87% | 4.48% | 2.10% | 2.60% | 0.11% | 0.41% | 6.54% | 9.30% | 2.97% | 3.78% |
| CORE | 28.35 | 39.98 | 16.04 | 17.58 | 11.05 | 16.44 | 6.30 | 7.01 | 21.96 | 27.30 | 14.73 | 15.44 |
| +CLSG | 28.75* | 40.60* | 16.17* | 17.74* | 11.27* | 16.62* | 6.52* | 7.23* | 22.77* | 28.54* | 15.13* | 15.90* |
| 提升比 | 1.43% | 1.57% | 0.82% | 0.89% | 1.98% | 1.11% | 3.48% | 3.11% | 3.71% | 4.56% | 2.70% | 3.02% |
| ATTMIX | 27.59 | 39.11 | 15.52 | 17.04 | 11.83 | 15.62 | 7.97 | 8.47 | 25.57 | 31.11 | 17.34 | 18.08 |
| +CLSG | 28.10* | 39.27 | 15.96* | 17.45* | 11.99* | 15.82* | 7.99 | 8.50* | 26.72* | 31.95* | 18.14* | 18.85* |
| 提升比 | 1.86% | 0.42% | 2.86% | 2.39% | 1.33% | 1.25% | 0.20% | 0.30% | 4.49% | 2.68% | 4.64% | 4.22% |

(3)除Nowplaying上的SRGNN+CLSG方法之外,在其他情况下,CLSG均显著提升了基线模型在短会话上的性能表现,在HR和MRR指标上,分别获得了5.82%和4.63%

的平均性能提升.在电商数据集(Diginetica和Tmall)上,CLSG对SRGNN模型同样产生了积极影响.具体而言,在Diginetica数据集上,SRGNN+CLSG在HR和MRR指标上

分别实现了 2.66%、4.46% 的平均性能提升,在 Tmall 数据集上,这两个指标的平均性能提升分别为 2.63% 和 3.85%,可归因于电商平台中用户近邻行为之间的关联性相对较弱,因此相较于近邻关系扰动所造成的影响,CLSG 引入的性能增益更为显著. 在 Nowplaying 数据集上,CLSG 同样可以提升除 SRGNN 外其他基模型的性能,原因是这些基模型中的注意力机制可以通过权重灵活分配来聚焦用户核心兴趣,而 CLSG 生成的反事实长会话可以为其提供更丰富的交互行为支持,两者相辅相成,共同促进推荐系统性能的提升. 由此研究动机再次得以验证,即 CLSG 可以有效缓解短会话问题,从而对现有 SBR 模型实现有效赋能.

4.4 与现有 CIDM 方法比较

为回答研究问题 3,选取两款现有 CIDM 方法 (CASR 和 CauseRec) 进行对比,旨在深入探究 CLSG 在应对短会话挑战时的性能优势. 表 4 展示了 3 款方法在不同 SBR 基线模型上的性能表现,其中 H@K 和 M@K 分

别表示 HR@K 和 MRR@K,粗体表示最优性能.

实验结果显示,无论是在所有会话上还是在短会话上,CASR 和 CauseRec 均未能显著提升基模型的推荐性能,在部分情况下,甚至导致基模型性能的衰退(如表 4 中 ↓ 所示)或使基模型完全失效(例如 ATTMIX+ CASR). 原因在于上述方法未考虑短会话问题,通过替换序列内的物品生成反事实会话,不仅无法通过增加会话内交互数量以缓解短会话问题,而且面临丢失原本就稀缺的交互信息的风险,从而带来负面效果. 相比之下,CLSG 在几乎所有情况下均有效提升基模型性能,其主要优势在于采用频繁模式挖掘引导反事实长会话生成,不仅保留了观测会话中的完整交互信息,而且在先验知识库的支持下生成了更为可靠的长会话. 因此,CLSG 可以有效应对短会话挑战并显著提升推荐性能. 上述实验结果进一步验证了本文所提出的 CLSG 方法在应对短会话挑战时的优越性.

表 4 CLSG 与现有 CIDM 方法在基准数据集上的性能比较

| | Diginetica | | | | Nowplaying | | | | Tmall | | | |
|-----------|--------------|--------------|--------------|--------------|--------------|-------------|--------------|---------------|--------------|--------------|--------------|--------------|
| | 所有会话 | | 短会话 | | 所有会话 | | 短会话 | | 所有会话 | | 短会话 | |
| | H@10 | M@10 | H@10 | M@10 | H@10 | M@10 | H@10 | M@10 | H@10 | M@10 | H@10 | M@10 |
| NARM | 38.40 | 16.52 | 39.06 | 17.10 | 14.07 | 7.12 | 12.74 | 6.62 | 20.65 | 11.73 | 16.72 | 9.8 |
| +CASR | 38.09 ↓ | 16.61 ↓ | 38.86 ↓ | 17.04 ↓ | 13.70 ↓ | 7.20 | 13.08 | 6.72 | 24.63 | 14.60 | 18.78 | 11.32 |
| +CauseRec | 38.44 | 16.74 | 39.16 | 17.56 | 13.26 ↓ | 6.14 ↓ | 12.68 ↓ | 5.82 ↓ | 24.13 | 14.01 | 19.72 | 11.48 |
| +CLSG | 38.87 | 17.00 | 39.94 | 17.56 | 14.22 | 7.35 | 14.00 | 7.10 | 26.11 | 14.65 | 20.74 | 11.78 |
| STAMP | 37.72 | 16.24 | 38.72 | 16.74 | 14.39 | 7.46 | 11.75 | 7.10 | 22.49 | 12.47 | 19.74 | 11.74 |
| +CASR | 37.56 ↓ | 16.15 ↓ | 38.48 ↓ | 16.64 ↓ | 14.36 ↓ | 7.38 ↓ | 14.04 | 7.10 | 23.09 | 13.37 | 19.30 ↓ | 11.60 ↓ |
| +CauseRec | 36.87 ↓ | 15.80 ↓ | 37.60 ↓ | 16.10 ↓ | 14.50 | 6.66 ↓ | 14.30 | 6.50 ↓ | 23.38 | 13.78 | 18.94 ↓ | 11.32 ↓ |
| +CLSG | 38.25 | 16.55 | 39.10 | 17.06 | 14.89 | 7.73 | 14.66 | 7.48 | 29.82 | 16.34 | 22.86 | 12.68 |
| SRGNN | 38.60 | 16.77 | 39.28 | 17.18 | 14.02 | 6.89 | 13.80 | 6.76 | 22.56 | 13.61 | 18.26 | 11.22 |
| +CASR | 38.03 ↓ | 16.58 ↓ | 38.86 ↓ | 16.98 ↓ | 13.82 ↓ | 6.74 ↓ | 13.56 ↓ | 6.60 ↓ | 23.67 | 14.41 | 17.42 ↓ | 10.96 ↓ |
| +CauseRec | 34.89 ↓ | 14.90 ↓ | 35.68 ↓ | 15.28 ↓ | 12.43 ↓ | 5.49 ↓ | 12.14 ↓ | 5.42 ↓ | 23.97 | 14.81 | 15.82 ↓ | 10.12 ↓ |
| +CLSG | 39.60 | 17.52 | 40.12 | 17.92 | 14.39 | 6.92 | 13.82 | 6.56 ↓ | 24.66 | 14.13 | 19.22 | 11.18 ↓ |
| CORE | 39.98 | 17.58 | 40.90 | 17.96 | 16.44 | 7.01 | 16.04 | 6.88 | 27.30 | 15.44 | 21.74 | 12.36 |
| +CASR | 39.17 ↓ | 17.44 ↓ | 39.86 ↓ | 17.74 ↓ | 16.35 ↓ | 7.10 | 15.98 ↓ | 6.90 | 27.17 ↓ | 15.34 ↓ | 21.94 | 12.26 ↓ |
| +CauseRec | 38.44 ↓ | 17.14 ↓ | 39.88 ↓ | 17.80 ↓ | 14.97 ↓ | 6.28 ↓ | 15.46 ↓ | 6.48 ↓ | 26.61 ↓ | 15.13 ↓ | 21.38 ↓ | 12.05 ↓ |
| +CLSG | 40.60 | 17.74 | 41.28 | 18.06 | 16.62 | 7.23 | 16.20 | 7.00 | 28.54 | 15.90 | 22.42 | 12.58 |
| ATTMIX | 39.11 | 17.04 | 39.35 | 17.38 | 15.62 | 8.47 | 15.00 | 7.96 | 31.11 | 18.08 | 23.66 | 14.46 |
| +CASR | 8.31 ↓ | 2.74 ↓ | 8.44 ↓ | 2.74 ↓ | 2.51 ↓ | 0.77 ↓ | 2.48 ↓ | 0.78 ↓ | 3.26 ↓ | 1.07 ↓ | 2.78 ↓ | 0.96 ↓ |
| +CauseRec | 39.04 ↓ | 16.61 ↓ | 37.79 ↓ | 16.11 ↓ | 14.20 ↓ | 7.70 ↓ | 14.31 ↓ | 6.57 ↓ | 31.66 | 18.21 | 23.92 | 13.36 ↓ |
| +CLSG | 39.27 | 17.45 | 40.06 | 17.88 | 15.82 | 8.50 | 15.30 | 7.98 | 31.95 | 18.85 | 23.92 | 14.68 |

4.5 性能改进实例分析

为深入探究 CLSG 方法提升现有 SBR 性能的潜在因素,从 Tmall 数据集中选取若干实例,多角度剖析使用 CLSG 前后,会话内物品与下一次点击物品(即真实标签)之间的联系. 值得注意的是,真实标签在 CLSG 实施会话干预过程中始终不可见,本节仅将真实标签作为验证依据,旨

在探索 CLSG 是否增强了会话内物品与下一次点击物品之间的联系,进而提升模型的预测准确性. 如表 5 所示,红色加下划线部分突出显示了与真实标签增强联系的关键元素,在实例(a)和实例(b)中,干预前会话内的所有物品与下一次点击物品均不属于同一类目或品牌,通过 CLSG 引入了与下一次点击物品的类目或品牌一致的物品,从

表 5 使用 CLSG 前后会话内物品与真实标签之间的联系

| | | |
|-------|----------------------|--|
| 实例(a) | 真实标签 | 物品 ID: 30390 类目 ID: 63 |
| | 已观测会话 (使用 CLSG 前) | 物品 ID: 22104、16246、9169 类目 ID: 37、37、260 |
| | 反事实会话 (使用 CLSG 后) | 物品 ID: 3673、4114、22104、16246 类目 ID: 63、 63 、37、37 |
| 实例(b) | 真实标签 | 物品 ID: 16170 品牌 ID: 1988 |
| | 已观测会话 (使用 CLSG 前) | 物品 ID: 5770、22872、15035、27030 品牌 ID: 506、2821、506、3723 |
| | 反事实会话 (使用 CLSG 后) | 物品 ID: 5770、22872、15035、4969 牌品 ID: 506、2821、506、 1988 、3723 |
| 实例(c) | 真实标签 | 物品 ID: 4260 |
| | 已观测会话 (使用 CLSG 前) | 物品 ID: 23006、1431、5648 与真实标签共现次数: 0、0、0 |
| | 反事实会话 (使用 CLSG 后) | 物品 ID: 23006、1431、17129、5648 与真实标签共现次数: 0、0、 6 、0 |
| 实例(d) | 真实标签 | 物品 ID: 151 |
| | 已观测会话 (使用 CLSG 前) | 物品 ID: 8355、7530 与真实标签的欧式距离: 0.63、0.56 |
| | 反事实会话 (使用 CLSG 后) | 物品 ID: 8355、8354、7530 与真实标签欧式距离: 0.63、 0.39 、0.56 |

而在属性维度上增强了会话内物品与真实标签之间的联系;在实例(c)中,干预前会话内的物品与下一次点击物品从未在同一会话中共现,现有 SBR 模型难以捕捉两者的联系,通过 CLSG 引入的物品与下一次点击物品频繁共现,从而在共现维度上增强了会话内物品与真实标签的联系;在实例(d)中,干预前会话内的物品与下一次点击物品在隐空间(以基模型 CORE 学习到的隐空间为例)中的欧式距离相对较远,通过 CLSG 引入的物品更靠近下一次点击物品,从而拉近了会话嵌入与真实标签嵌入之间的距离.上述案例分析表明,CLSG 从物品属性、频繁共现和隐空间相似度等多维度上增强了会话内物品与下一次点击物品之间的联系,因此有助于提升现有 SBR 模型推荐准确性.

4.6 超参数实验

CLSG 方法涉及两个关键超参数:正则项系数 λ 和排序折扣因子 p ,本节探索不同超参数设置对推荐性能的影响.

4.6.1 正则项系数

正则项系数 λ 用于调控干预效应正则项在目标函数中的权重, λ 越大,干预效应正则项对目标函数贡献越大.对 λ 分别取值 $\{0, 0.1, 0.2, 0.3, 0.4\}$,以评估其对推荐性能的影响.实验结果如图 4 所示.模型性能随着 λ 的增大,呈现先上升后下降的趋势:在一定范围内,增大 λ 值有效强化了干预效应正则项对推荐主任务的正面促进作用,因此模型性能得以提升;然而,当 λ 超过一

定阈值后,优化过程过于偏重干预效应正则项,从而对推荐主任务的学习产生负面影响,导致模型性能下降.总体而言,在实验所设定的超参数取值范围内,复杂模型(CORE 和 ATTMIX)展现出了相对稳定的性能表现,而简单模型(NARM、STAMP 和 SRGNN)则对这一超参数的调整较为敏感,可通过在所设定的取值范围内进行精细搜寻,找到最佳超参数设置.

4.6.2 排序折扣因子 p

排序折扣因子 p 用于控制预测列表中不同排序位置在干预效应函数中的权重分配. p 较大时,不同排序位置的权重趋于一致,表明函数对列表排序一致性的要求较为宽松;反之, p 较小时,排序靠前位置权重更高,表明函数对列表排序一致性的要求更为严格.对 p 分别取值 $\{0.9, 0.8, 0.7, 0.6\}$,以评估 p 对模型性能的影响.实验结果如图 5 所示.在 Diginetica 和 Nowplaying 数据集上,最佳 p 值相对较大,原因是这两个数据集所包含的候选物品数量较多,导致排序任务的难度较大,因此可通过增大 p 值适当放松排序一致性要求;在 Tmall 数据集上,最佳 p 值相对较小,原因是该数据集包含的候选物品数量较少,排序任务相对简单,因此可通过减小 p 值提高列表排序一致性要求.在超参数取值范围内,复杂模型(CORE 和 ATTMIX)的性能相对稳定,而简单模型(NARM、STAMP 和 SRGNN)对 p 的调整较为敏感.综上所述,为获得最佳 p 值,应综合

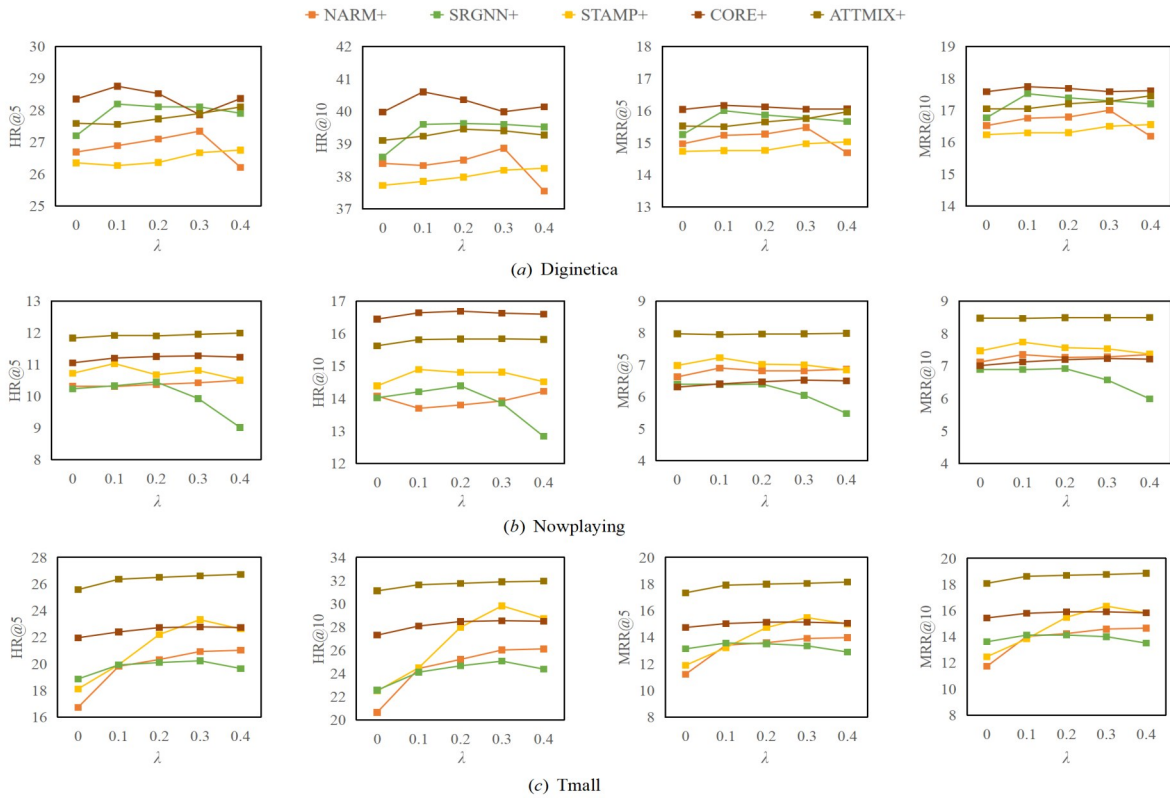


图4 正则项系数 λ 对性能的影响

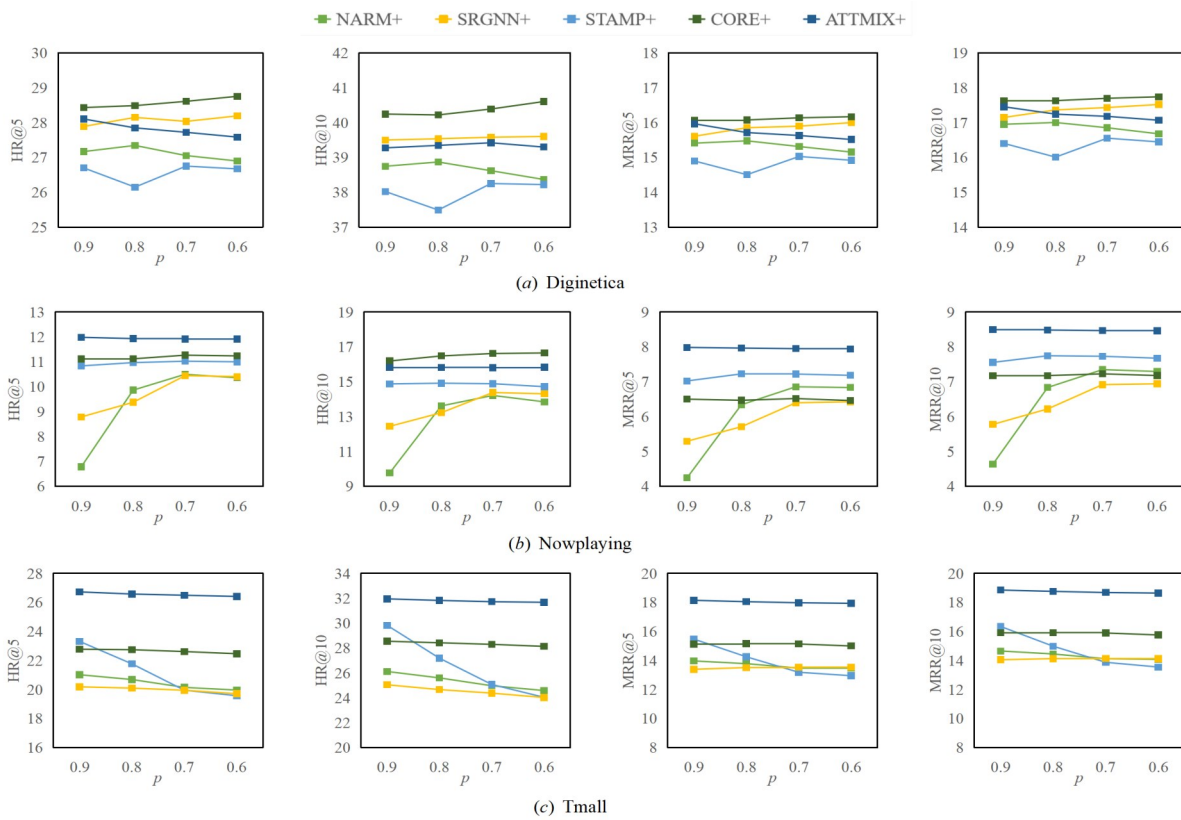


图5 排序折扣因子 p 对性能的影响

考虑数据集特点和模型特点,在所设定超参数取值范围内进行精细搜寻.

5 结论

针对 SBR 系统普遍面对的短会话挑战,提出一种通过频繁模式引导长会话生成的反事实推理方法 CLSG. 该方法遵循反事实理论的“归纳-行动-预测”经典三阶段推理流程,回答反事实问题:“如果会话内包含更丰富的交互物品,SBR 模型预测结果将会如何?”三个基准数据集上的实验显示,CLSG 显著提升了五款现有 SBR 模型的预测性能,表明该方法可有效缓解短会话问题,为现有形态各异的 SBR 模型提供普惠式赋能. 近年来,大语言模型备受瞩目,未来的研究将扩展 CLSG 方法至文本丰富的 SBR 场景,旨在利用大语言模型丰富的外部知识和卓越的推理能力,增强反事实长会话生成的可解释性,以提升建模透明度,从而为会话干预提供指导.

参考文献

- [1] ZHAO K, ZHENG Y K, ZHUANG T, et al. Joint learning of E-commerce search and recommendation with a unified graph neural network[C]//Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining. New York: ACM, 2022: 1461-1469.
- [2] 丁琪, 田萱, 孙国栋. 基于注意力增强的热点感知新闻推荐模型[J]. 电子学报, 2023, 51(1): 93-104.
DING Q, TIAN X, SUN G D. HAN: Hotspot-aware attention enhanced news recommendation[J]. Acta Electronica Sinica, 2023, 51(1): 93-104. (in Chinese)
- [3] 李琳, 唐守廉. 基于多层注意力表示的音乐推荐模型[J]. 电子学报, 2020, 48(9): 1672-1679.
LI L, TANG S L. Hierarchical attention representation model for music recommendation[J]. Acta Electronica Sinica, 2020, 48(9): 1672-1679. (in Chinese)
- [4] HE Z C, LIU W W, GUO W, et al. A survey on user behavior modeling in recommender systems[C]//Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence. California: International Joint Conferences on Artificial Intelligence Organization, 2023: 6656-6664.
- [5] PI Q, BIAN W J, ZHOU G R, et al. Practice on long sequential user behavior modeling for click-through rate prediction[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2019: 2671-2679.
- [6] WANG S J, CAO L B, WANG Y, et al. A survey on session-based recommender systems[J]. ACM Computing Surveys, 2022, 54(7): 1-38.
- [7] WANG S J, ZHANG Q, HU L, et al. Sequential/session-based recommendations: Challenges, approaches, applications and opportunities[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2022: 3425-3428.
- [8] HIDASI B, KARATZOGLOU A, BALTRUNAS L, et al. Session-based recommendations with recurrent neural networks[C]//Proceedings of the 4th International Conference on Learning Representations. New York: ACM: 2016: Poster.
- [9] LI J, REN P J, CHEN Z M, et al. Neural attentive session-based recommendation[C]//Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. New York: ACM, 2017: 1419-1428.
- [10] WU S, TANG Y Y, ZHU Y Q, et al. Session-based recommendation with graph neural networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 346-353.
- [11] CHEN T W, WONG R C W. Handling information loss of graph neural networks for session-based recommendation[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2020: 1172-1180.
- [12] 黄震华, 林小龙, 孙圣力, 等. 会话场景下基于特征增强的图神经推荐方法[J]. 计算机学报, 2022, 45(4): 766-780.
HUANG Z H, LIN X L, SUN S L, et al. Feature augmentation based graph neural recommendation method in session scenarios[J]. Chinese Journal of Computers, 2022, 45(4): 766-780. (in Chinese)
- [13] 闫昭, 项欣光, 李泽超. 基于交互序列商品相关性建模的图卷积会话推荐[J]. 中国科学: 信息科学, 2022, 52(6): 1069-1082.
YAN Z, XIANG X G, LI Z C. Item correlation modeling in interaction sequence for graph convolutional session recommendation[J]. Scientia Sinica (Informationis), 2022, 52(6): 1069-1082. (in Chinese)
- [14] LIU Q, ZENG Y F, MOKHOSI R, et al. STAMP: Short-term attention/memory priority model for session-based recommendation[C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2018: 1831-1839.
- [15] HOU Y P, HU B B, ZHANG Z Q, et al. CORE: Simple and effective session-based recommendation within consistent representation space[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2022: 1796-1801.

- [16] ZHANG P Y, GUO J Y, LI C Z, et al. Efficiently leveraging multi-level user intent for session-based recommendation via atten-mixer network[C]//Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. New York: ACM, 2023: 168-176.
- [17] LUO A J, ZHAO P P, LIU Y C, et al. Collaborative self-attention network for session-based recommendation[C]//Proceedings of the 29th International Joint Conference on Artificial Intelligence. New York: ACM, 2020: 2591-2597.
- [18] PAN Z Q, CAI F, CHEN W Y, et al. Star graph neural networks for session-based recommendation[C]//Proceedings of the 29th ACM International Conference on Information & Knowledge Management. New York: ACM, 2020: 1195-1204.
- [19] XIA X, YIN H Z, YU J L, et al. Self-supervised hypergraph convolutional networks for session-based recommendation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(5): 4503-4511.
- [20] TOUVRON H, LAVRIL T, IZACARD G, et al. LLaMA: Open and efficient foundation language models[EB/OL]. (2023-02-27)[2024-08-28]. <https://arxiv.org/abs/2302.13971>.
- [21] ACHIAM J, ADLER S, AGARWAL S, et al. Gpt-4 technical report[EB/OL]. (2023-03-15) [2024-08-28]. <https://arxiv.org/abs/2303.08774v5>.
- [22] QIAO S T, GAO C, WEN J H, et al. LLM4SBR: A lightweight and effective framework for integrating large language models in session-based recommendation[EB/OL]. (2024-02-21)[2024-08-28]. <https://arxiv.org/abs/2402.13840v1>.
- [23] GUO N C, CHENG H W, LIANG Q Q, et al. Integrating large language models with graphical session-based recommendation[EB/OL]. (2024-02-26) [2024-08-28]. <https://arxiv.org/abs/2402.16539v1>.
- [24] 邬俊, 刘林, 卢香葵, 等. 基于情感可控文本生成的可解释推荐系统[J]. 闽南师范大学学报(自然科学版), 2023, 36(4): 24-34.
WU J, LIU L, LU X K, et al. Explainable recommender systems via emotion controlling text generation[J]. Journal of Minnan Normal University (Natural Science), 2023, 36(4): 24-34. (in Chinese)
- [25] YU X, PENG Q L, XU L W, et al. A selective ensemble learning based two-sided cross-domain collaborative filtering algorithm[J]. Information Processing & Management, 2021, 58(6): 102691.
- [26] LIU Z W, FAN Z W, WANG Y, et al. Augmenting sequential recommendation with pseudo-prior items via reversely pre-training transformer[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2021: 1608-1612.
- [27] PEARL J, GLYMOUR M, JEWELL N P. Causal Inference in Statistics: A Primer[M]. United Kingdom: John Wiley & Sons, 2016.
- [28] GAO C, ZHENG Y, WANG W J, et al. Causal inference in recommender systems: A survey and future directions[J]. ACM Transactions on Information Systems, 2024, 42(4): 1-32.
- [29] WEI T X, FENG F L, CHEN J W, et al. Model-agnostic counterfactual reasoning for eliminating popularity bias in recommender system[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: ACM, 2021: 1791-1800.
- [30] WANG W J, FENG F L, HE X N, et al. Clicks can be cheating: Counterfactual recommendation for mitigating clickbait issue[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2021: 1288-1297.
- [31] SONG W Z, WANG S J, WANG Y, et al. A counterfactual collaborative session-based recommender system[C]//Proceedings of the ACM Web Conference 2023. New York: ACM, 2023: 971-982.
- [32] WANG Z L, ZHANG J S, XU H T, et al. Counterfactual data-augmented sequential recommendation[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2021: 347-356.
- [33] ZHANG S Y, YAO D, ZHAO Z, et al. CauseRec: Counterfactual user sequence synthesis for sequential recommendation[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2021: 367-377.
- [34] TRAN H X, LE T D, LI J Y, et al. Recommending the most effective intervention to improve employment for job seekers with disability[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: ACM, 2021: 3616-3626.
- [35] WANG W J, FENG F L, NIE L Q, et al. User-controllable recommendation against filter bubbles[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2022: 1251-1261.
- [36] ZHANG X, JIA H N, SU H J, et al. Counterfactual reward modification for streaming recommendation with delayed feedback[C]//Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Infor-

- mation Retrieval. New York: ACM, 2021: 41-50.
- [37] GOETHALS B. Survey on frequent pattern mining[J]. University of Helsinki, 2003, 19: 840-852.
- [38] FOURNIER-VIGER P, GUENICHE T, ZIDA S, et al. ER-Miner: Sequential rule mining using equivalence classes[M]// Lecture Notes in Computer Science. Cham: Springer International Publishing, 2014: 108-119.
- [39] OH S, USTUN B, MCAULEY J, et al. Rank list sensitivity of recommender systems to interaction perturbations[C]//Proceedings of the 31st ACM International Conference on Information & Knowledge Management. New York: ACM, 2022: 1584-1594.
- [40] ZANGERLE E, PICHL M, GASSLER W, et al. Nowplaying music dataset: Extracting listening behavior from twitter[C]// Proceedings of the First International Workshop on Internet-Scale Multimedia Management. New York: ACM, 2014: 21-26.
- [41] ZHOU P L, HUANG Y L, XIE Y Q, et al. Is contrastive learning necessary? A study of data augmentation vs contrastive learning in sequential recommendation[C]//Proceedings of the ACM Web Conference 2024. New York: ACM, 2024: 3854-3863.
- [42] KANG W C, MCAULEY J. Self-attentive sequential recommendation[C]//2018 IEEE International Conference on Data Mining (ICDM). Piscataway: IEEE, 2018: 197-206.

作者简介



卢香葵 女, 1998年12月出生于广西壮族自治区河池市. 现为北京交通大学计算机科学与技术学院博士研究生. 主要研究方向为机器学习基础理论及其在信息检索与推荐系统中的应用等.

E-mail: luxkui@bjtu.edu.cn



邬俊 男, 1981年11月出生于辽宁省鞍山市. 2010年毕业于大连海事大学计算机应用技术专业. 现为北京交通大学计算机科学与技术学院副教授、博士生导师. 主要研究方向为机器学习基础理论及其在信息检索、推荐系统、数字医疗、智慧交通、计算甲骨学等领域中的应用等.

E-mail: wuj@bjtu.edu.cn