

# 面向共享账户序列推荐的提示增强图注意力网络<sup>\*</sup>

赵中英<sup>1</sup>, 王莹<sup>1</sup>, 张劲羽<sup>1</sup>, 周慧<sup>1</sup>, 李超<sup>2</sup>, 曾庆田<sup>1</sup>

<sup>1</sup>(山东科技大学 计算机科学与工程学院, 山东 青岛 266590)

<sup>2</sup>(山东科技大学 电子信息工程学院, 山东 青岛 266590)

通信作者: 李超, E-mail: lichao@sdust.edu.cn; 曾庆田, E-mail: qtzeng@163.com



**摘 要:** 共享账户是指多个用户通常共享同一电子账户 (智能家居系统往往以一个家庭为一个账户), 因此, 同一账户下的行为序列往往包含多个用户的交互记录. 面向共享账户场景的序列推荐成为近年来的研究热点, 但仍然面临以下挑战: (1) 现有方法难以有效解耦同一账户下的不同潜在用户偏好, 导致用户表示学习出现偏差. (2) 现有方法通常采用固定的超参数来模拟共享账户中的潜在用户数量, 难以自适应地确定潜在用户数量, 易导致欠拟合或噪声干扰问题. 为了解决上述挑战, 提出一种面向共享账户序列推荐的提示增强图注意力网络 (PE-GAT). 具体而言, 该方法首先通过设计动态加权密度聚类算法来确定账户中的潜在用户数量, 构建用户级序列图, 以显式区分不同潜在用户的交互行为. 在此基础上, 提出基于图注意力网络的预训练模型, 通过解耦混合序列中的用户偏好来学习初始序列表示. 随后, 受预训练-提示调优范式启发, 设计了基于注意力机制的提示增强模块, 通过提示模板优化序列表示并指导模型再训练. 最终, PE-GAT 通过融合账户级表示与提示增强后的序列表示, 以生成个性化的推荐结果. 在 HVIDEO-E 和 HVIDEO-V 两个真实共享账户数据集上的实验表明, PE-GAT 优于 14 种基线方法, 在 *MRR* 和 *Recall* 指标上分别最高提升 4.73% 和 5.59%.

**关键词:** 推荐系统; 共享账户序列推荐; 提示增强; 图注意力网络

**中图法分类号:** TP18

中文引用格式: 赵中英, 王莹, 张劲羽, 周慧, 李超, 曾庆田. 面向共享账户序列推荐的提示增强图注意力网络. 软件学报. <http://www.jos.org.cn/1000-9825/7626.htm>

英文引用格式: Zhao ZY, Wang Y, Zhang JY, Zhou H, Li C, Zeng QT. Prompt-enhanced Graph Attention Network for Shared-account Sequential Recommendation. Ruan Jian Xue Bao/Journal of Software (in Chinese). <http://www.jos.org.cn/1000-9825/7626.htm>

## Prompt-enhanced Graph Attention Network for Shared-account Sequential Recommendation

ZHAO Zhong-Ying<sup>1</sup>, WANG Ying<sup>1</sup>, ZHANG Jin-Yu<sup>1</sup>, ZHOU Hui<sup>1</sup>, LI Chao<sup>2</sup>, ZENG Qing-Tian<sup>1</sup>

<sup>1</sup>(College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

<sup>2</sup>(College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

**Abstract:** In real-world scenarios such as e-commerce platforms and smart-home systems, multiple users often share the same account, resulting in behavioral sequences that mix interactions from different users. Sequential recommendation in shared-account settings has recently attracted increasing research attention but still faces two major challenges: (1) interactions from different latent users are difficult to disentangle accurately, leading to biased representation learning; (2) reliance on static hyperparameters to estimate the number of latent users prevents adaptive determination of user cardinality, which may cause underfitting or introduce noise. To address these challenges, this study proposes PE-GAT (prompt-enhanced graph attention network), an adaptive sequential recommendation framework for shared-account settings. PE-GAT first employs a dynamically-weighted, density-based clustering algorithm to infer the number of latent users and construct user-level sequential graphs that explicitly separate heterogeneous user behaviors. Based on these graphs, a graph attention

\* 基金项目: 国家自然科学基金 (62472263, 52374221, 62502288); 山东省泰山学者人才工程 (tstp20250506, tsqn202211154); 山东省自然科学基金项目 (ZR2024MF034, ZR2022MF268)

收稿时间: 2025-09-23; 修改时间: 2025-11-12; 采用时间: 2025-12-25; jos 在线出版时间: 2026-04-22

network pre-training module is employed to disentangle mixed user preferences and learn initial sequence representations. Inspired by the pre-training and prompt-tuning paradigm, a self-attention-based prompt-enhancement module is then designed to refine sequence embeddings via prompt templates during retraining. Finally, account-level representations are fused with prompt-enhanced sequence embeddings to generate personalized recommendations. Extensive experiments on two real-world shared-account datasets (HVIDEO-E and HVIDEO-V) show that PE-GAT outperforms 14 state-of-the-art baselines, achieving maximum improvements of 4.73% and 5.59% in *MRR* and *Recall* metrics, respectively.

**Key words:** recommender system; shared-account sequential recommendation (SSR); prompt enhanced; graph attention network (GAT)

序列推荐 (sequential recommendation, SR) 旨在通过分析用户的历史行为序列来推断用户偏好, 进而为用户推荐个性化商品或服务<sup>[1-5]</sup>. 经过长期探索, 序列推荐算法已在许多具体场景中得到应用, 例如基于会话的短序列推荐场景<sup>[6,7]</sup>、跨域序列推荐场景<sup>[8-10]</sup>等. 作为代表性应用场景之一, 共享账户序列推荐 (shared-account sequential recommendation, SSR) 同样引起研究者的广泛关注<sup>[11-14]</sup>. 具体而言, 共享账户是指某个账户中包含多名潜在用户, 且每位潜在用户均使用该账户产生有记录的交互行为<sup>[15]</sup>. 例如, 家庭成员常共享同一台智能终端设备 (IP 电视或智能音箱), 亲密关系中的用户通常共享同一电商账户以完成购物或消费. 图 1 展示了一个由多名家庭成员共享数字电视账户的推荐示例, 多名潜在用户与电视的历史交互被交错地记录在同一条混合行为序列中, 并且不同的潜在用户拥有不同的观看偏好. 由此可见, 在共享账户场景下, 需要同时实现对潜在用户的识别与混合序列的建模, 这对已有的序列推荐方法带来了新的挑战<sup>[11,16-18]</sup>.

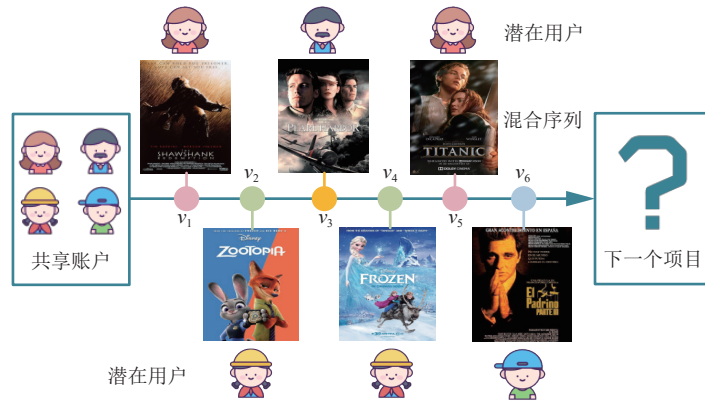


图 1 多名家庭成员共享数字电视账户的推荐示例

早期针对 SSR 的研究倾向于采用基于循环神经网络 (recurrent neural network, RNN) 的方法, 例如  $\pi$ -Net<sup>[19]</sup>与 PSJNet<sup>[12]</sup>.  $\pi$ -Net 提出了一种基于门控循环单元 (gated recurrent unit, GRU) 的并行信息共享网络. PSJNet 则采用一个更通用的框架, 即并行拆分连接网络, 来概括  $\pi$ -Net, 它引入“拆分”和“连接”的概念来解决共享账户序列推荐问题. 然而, 基于 RNN 的方法在长序列建模方面存在梯度消失的风险, 同时这类方法也难以捕捉用户与项目之间的复杂依赖关系, 限制了学习到的用户和项目表示的表达能力. 随后, 研究者尝试借助图卷积网络 (graph convolutional network, GCN) 来解决上述挑战<sup>[13,14,20]</sup>. 作为代表性方法之一, DA-GCN<sup>[20]</sup>通过构建序列图来表示共享账户混合序列中的复杂依赖关系, 并设计了一种基于注意力机制的消息传递机制, 旨在捕捉特定的潜在用户与项目间的关系. TiDA-GCN<sup>[13]</sup>则在 DA-GCN 的基础上改进序列建模组件, 进一步考虑项目间的时间间隔信息, 并采用自注意力机制捕捉项目的全局相关性信息. 上述方法在 SSR 任务上取得了优异的性能表现, 但依旧面临以下两项挑战.

(1) 难以有效解耦混合序列中不同潜在用户的偏好. 在共享账户场景下, 账户序列中混杂着多名潜在用户的历史交互行为<sup>[11,21,22]</sup>. 由于缺乏可靠的监督信息, 模型在训练过程中无法得知每个账户中存在多少潜在用户, 同时也难以判断混合序列中的每个交互具体属于哪位潜在用户<sup>[23-25]</sup>. 现有方法<sup>[5,13,20]</sup>通常缺乏明确区分用户身份的机制, 难以有效解耦混合序列中不同潜在用户的偏好特征. 此外, 在未解耦情况下对混合序列采用直接建模的策略<sup>[12,19,22,24,26]</sup>, 会导致不同潜在用户偏好之间产生相互干扰, 导致序列表示学习效果偏差.

(2) 难以自适应地确定潜在用户数量. 现有的共享账户序列推荐方法普遍采用固定超参数<sup>[12,13,19,20]</sup>来模拟每个账户内的潜在用户数量, 并通过实验在不同数据集上进行微调<sup>[21,23,25]</sup>. 然而, 不同共享账户中的潜在用户数量可能是不同的, 固定超参数的设定违背了客观规律, 导致方法学习到的表示存在偏差. 当超参数设定小于账户中实际潜在用户数量时, 模型将无法完整表征所有用户偏好, 导致信息丢失, 发生欠拟合; 反之, 模型会因此引入冗余, 不仅导致计算开销变高, 还产生了额外的噪声.

为了解决上述挑战, 本文提出了一种面向共享账户序列推荐的提示增强图注意力网络 (prompt-enhanced graph attention network, PE-GAT). 具体而言, 该方法首先设计动态加权密度聚类算法自适应地从混合序列中推断潜在用户数量, 从而将账户级别的混合交互序列转化为用户级序列图, 以便于后续显式地区分不同潜在用户的交互行为. 随后, 提出了一种基于图注意力网络 (graph attention network, GAT)<sup>[20]</sup>的预训练模型, 以解耦混合序列中不同潜在用户的偏好并学习初始的序列表示. 受预训练-提示调优训练方法<sup>[27-30]</sup>的启发, 本文设计了一种基于注意力机制<sup>[1]</sup>的提示增强模块. 该模块被设计用于优化预训练阶段学习到的序列表示, 并进一步指导模型进行再训练. 其中, 为进一步提升提示信息对真实用户分布的表征能力, 将动态加权密度聚类算法识别到的混合序列中的潜在用户作为后验提示信息. 随后借助注意力机制产生提示模板以加工原有的初始序列表示. 最终, 经过再训练后的 PE-GAT 通过融合账户级表示与提示增强后的序列表示, 为账户提供个性化推荐.

本工作的贡献总结如下.

(1) 提出一种面向共享账户序列推荐的提示增强图注意力网络 (PE-GAT). PE-GAT 的相关代码及数据集已公开发布至 GitHub 平台 (详见 <https://github.com/ZZY-GraphMiningLab/PE-GAT>).

(2) 设计一种自适应识别潜在用户的机制. 通过设计的动态加权密度聚类算法对序列数据进行自适应聚类, 准确识别共享账户中的潜在用户, 克服了现有方法因依赖固定超参数而产生的识别偏差问题.

(3) 提出一种基于注意力机制的提示增强模块. 通过引入注意力机制构建提示模板, 优化初始序列表示, 有效缓解了现有方法因缺乏显式用户区分机制而导致的偏好混淆问题.

(4) 在 2 个真实数据集上进行实验, 实验结果表明, PE-GAT 在 SSR 任务中实现了性能提升, 从而证实了本文方法的有效性.

本文第 1 节介绍共享账户序列推荐的相关方法和研究现状. 第 2 节介绍本文构建的面向共享账户序列推荐的提示增强图注意力网络. 第 3 节通过一系列实验验证所提方法的有效性. 最后总结全文.

## 1 相关工作

### 1.1 序列推荐方法

序列推荐方法通过分析用户历史行为, 预测其未来可能感兴趣的项目. 在 SR 的早期研究中, 基于马尔可夫链的方法<sup>[31,32]</sup>曾占主导地位, 该方法将推荐过程视为马尔可夫决策过程 (Markov decision process, MDP) 中的动作预测. 例如, Shani 等人<sup>[31]</sup>提出一个基于 MDP 的推荐器, 兼顾了推荐的长期效果和期望值. Chen 等人<sup>[32]</sup>则将播放列表生成问题转化为欧氏空间中马尔可夫链的嵌入问题, 并提出逻辑马尔可夫嵌入算法. 然而, 这些方法存在局限性. 首先, MDP 方法依赖于一个强大的基础模型, 该模型需快速求解且内存消耗低. 其次, MDP 在处理整个序列时, 会面临复杂的状态空间, 难以做出合理的序列化预测. 为提升模型捕捉高维序列信息的能力, 研究者转向基于深度神经网络的 SR 方法<sup>[1,6,7,33-36]</sup>. 例如, Hidasi 等人<sup>[6]</sup>将 RNN 应用于 SR 任务, 利用其门控机制记忆用户偏好的转移, 从而记录并传递序列的局部依赖关系, 显著改进了序列预测. 尽管如此, 基于 RNN 的 SR 方法在捕捉长序列依赖关系上存在困难, 且易出现梯度消失问题. 为应对这些挑战, Hsu 等人<sup>[37]</sup>提出一种时间关注图神经网络, 通过从用户-项目交互中提取局部子图并在其上训练带注意力的图神经网络 (graph neural network, GNN), 来学习更精确的序列表示. 该方法还改进了权重矩阵, 使其能够表示用户、项目和属性间的关系. 然而, 这些 SR 方法大多基于一个理想假设, 即每个推荐账户仅由一个用户使用. 因此, 在共享账户场景中, 由于未能充分考虑账户可能涉及的多用户, 这些方法的性能往往受限.

## 1.2 共享账户推荐方法

共享账户推荐方法通过分析账户共享者间的关系,为不同用户提供个性化服务.共享账户允许多个用户共同使用一个账户进行操作和访问,例如视频流媒体或在线购物账户.以往的研究将共享账户推荐分为两个步骤:用户识别和推荐<sup>[21,22]</sup>.例如,Zhang等人<sup>[24]</sup>将用户识别视为子空间聚类问题,他们提出的方法采用自表达矩阵将账户表示映射到潜在用户的子空间,并采用子空间聚类方法来获取用户的兴趣表示.Jiang等人<sup>[21]</sup>开发了一种无监督学习框架,通过会话异构图嵌入来识别潜在用户,区分和建模账户中用户的偏好,并对会话进行分组以精确识别用户.Wang等人<sup>[22]</sup>假设不同用户在不同时间活跃,通过时间切片划分用户,并使用消费日志进行用户建模.随后,他们采用K最近邻算法(K-nearest neighbor, KNN)生成账户表示并生成推荐.Yang等人<sup>[26]</sup>通过分析历史行为相似度来判断不连续行为是否属于同一用户,他们首先检测和定义账户中的时间间隔,随后计算间隔间的偏好相似度来学习潜在用户表示.然而,这些方法大多是两阶段模型,不利于端到端学习.为解决这一问题,Wen等人<sup>[25]</sup>提出了一种基于自注意力的多用户识别网络,该方法通过自注意力机制在序列中选择特定项目来学习潜在用户的兴趣表示.郭磊等人<sup>[8]</sup>提出一种基于自注意力网络的共享账户跨域序列推荐模型(SCRM),通过多头自注意力机制模拟共享账户中的多个用户,并结合跨域信息传输网络来提高推荐性能.尽管如此,基于注意力的方法在捕捉数据中的高阶信息和结构化知识方面仍存在困难. $\pi$ -Net<sup>[19]</sup>和PSJNet<sup>[12]</sup>提出了账户中包含 $H$ 个潜在用户的假设,并通过融合这些用户表示来学习共享账户的总体表示,并通过跨域信息来优化表示.继此之后,DA-GCN<sup>[20]</sup>和TiDA-GCN<sup>[13]</sup>沿用了这一假设,并进一步引入图卷积网络(GCN)来建模用户和项目之间的复杂关系.然而,这种固定潜在用户数量的假设与实际场景不符,限制了其在实际应用中的灵活性和准确性.

## 1.3 提示相关的推荐方法

提示技术最初在自然语言处理(natural language processing, NLP)领域被提出<sup>[38-41]</sup>.提示的最初概念是指与预训练模型的输入相结合的带前缀的纯文本,用于提高模型对任务语义的理解<sup>[39]</sup>.根据提示的表示方式,主要分为硬提示和软提示两类.硬提示由离散的真实单词组成<sup>[42,43]</sup>,但由于其依赖人工设计且优化成本较高,后续研究转向了自动搜索硬提示的方法<sup>[44,45]</sup>.相比之下,软提示由可学习的连续嵌入向量组成,如Prefix-tuning<sup>[38]</sup>和PPT<sup>[27]</sup>,这些方法在NLP生成任务中展现出更强的适应性.近年来,提示技术逐渐被引入推荐系统领域.例如,M6-Rec<sup>[28]</sup>和P5<sup>[29]</sup>等研究尝试将NLP中的提示调整策略迁移至推荐任务,如点击率预测和可解释推荐.这些方法将用户行为、描述、项目元数据和评论转换为自然语言序列作为提示,以有效利用预先训练的Transformer模型.此外,Wu等人<sup>[30]</sup>还将冷启动推荐建模为序列推荐的下游任务,并通过构建基于用户配置文件的前缀软提示优化模型性能.在共享账户场景下,由于多个用户可能共用同一账户,传统推荐方法难以准确区分个体偏好.为此,PE-GAT通过构建基于潜在用户的软提示来更好地学习预训练信息,从而有效提升推荐准确性.

## 2 面向共享账户序列推荐的提示增强图注意力网络

本节首先介绍PE-GAT的符号定义与SSR的任务定义,随后详细阐述PE-GAT模型的各个组件及结构设计.

### 2.1 符号与任务定义

令 $S = \{I_1, I_2, \dots, I_t\}$ 为行为序列, $I_i \in \mathcal{I}$  ( $1 \leq i \leq p$ )为项目索引, $\mathcal{I}$ 为项目集合. $U = \{U_1, U_2, \dots, U_k, \dots, U_{|U|}\}$ 为账户的子集,其中, $U_k \in \mathcal{U}$  ( $1 \leq k \leq n$ )表示账户索引, $\mathcal{U}$ 为整个账户集合.SSR的任务可形式化定义为,给定用户的历史行为序列 $S$ ,预测下一个项目 $I_{t+1}$ 的推荐概率:

$$P(I_{t+1} | S) \sim f(S) \quad (1)$$

其中, $f(S)$ 为学习到的概率估计函数,用于计算候选项目的推荐概率.

### 2.2 模型概述

本节设计了面向共享账户序列推荐的提示增强图注意力网络模型(PE-GAT),如图2所示,其核心思想是通过基于图注意力网络的预训练模型与自适应提示增强技术,解决共享账户场景下多用户偏好解耦与潜在用户数量自

适应的双重挑战. PE-GAT 包含 3 个关键模块: 模型输入、潜在用户建模和模型再训练, 实现从原始行为序列到推荐的端到端学习.

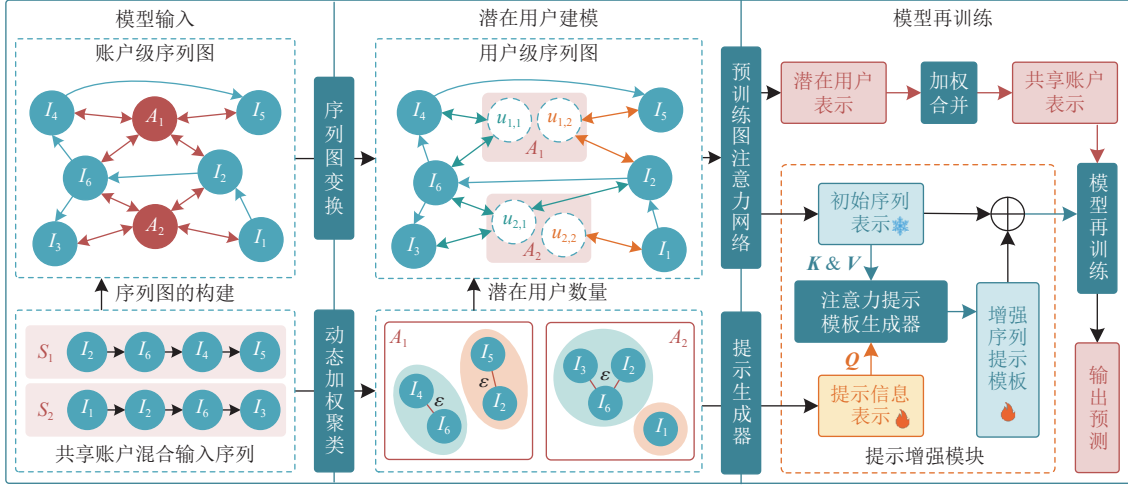


图2 面向共享账户序列推荐的提示增强图注意力网络模型图

(1) 模型输入. 针对共享账户中混杂的多用户行为特征, PE-GAT 首先构建账户级序列图并引入动态加权密度聚类算法获得潜在用户数量, 将原始账户级混合序列转化为用户级序列图. 该图通过两类节点 (用户节点和项目节点) 和两类边 (用户-项目交互边、项目-项目交互边) 的异构结构, 初步分离潜在用户行为模式. 与现有直接建模混合序列的方法相比, 这种设计能有效捕捉共享账户中的复杂交互模式, 并为后续显式地区分不同潜在用户的交互行为奠定基础.

(2) 潜在用户建模. 基于用户级序列图, 设计基于图注意力网络的预训练模型进行表示学习. 通过消息传递机制聚合相邻节点的信息, 生成初始序列表示, 并将预训练后的模型参数保存. 另一方面, 对原始输入采用动态加权密度聚类自适应地识别出共享账户中的不同使用者, 该方法无需预设用户数量, 避免了人工设定超参数引入的偏差. 能够根据数据分布自动划分用户群体. 将聚类识别结果输入提示生成器, 得到对应的提示嵌入表示. 通过这一过程, 模型不仅能够准确判别当前行为序列的实际使用者身份, 还能获取账户内潜在用户数量的后验分布信息, 从而增强模型.

(3) 模型再训练. 为了优化学习到的初始序列表示, PE-GAT 设计了提示增强模块. 该模块采用两阶段处理流程: 首先, 将提示生成器输出的提示信息表示作为查询向量 (Query), 预训练获取的初始序列表示作为键向量 (Key) 和值向量 (Value), 通过注意力机制生成增强序列提示模板; 其次, 将该模板与初始序列表示进行特征维度拼接, 形成融合后的增强表示. 这种设计相比简单的特征拼接, 能更好地在保留原始序列语义特征的同时引入用户区分信号. 最终通过再训练并联合优化共享账户表示与增强序列表示, 实现端到端推荐预测.

### 2.3 模型输入

为了建模复杂的用户-项目交互关系, 本文构建了账户级序列图  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , 其中,  $\mathcal{V}$  表示用户与项目节点的集合;  $\mathcal{E}$  为边的集合, 表示节点之间的联系, 考虑了两种类型的关联: 用户-项目交互和项目间的序列顺序. 本文采用矩阵的形式来描述序列图  $\mathcal{G}$ :

$$\mathcal{G} = \begin{bmatrix} \mathcal{M}_{I \rightarrow I} & \mathcal{M}_{I \rightarrow U} \\ \mathcal{M}_{U \rightarrow I} & 0 \end{bmatrix} \quad (2)$$

其中,  $\mathcal{M}_{I \rightarrow I} \in \mathbb{R}^{m \times m}$  表示项目间序列关系的拉普拉斯矩阵;  $\mathcal{M}_{I \rightarrow U} \in \mathbb{R}^{m \times n}$  表示项目-用户交互矩阵;  $\mathcal{M}_{U \rightarrow I} \in \mathbb{R}^{n \times m}$  为  $\mathcal{M}_{I \rightarrow U}$  矩阵的转置, 存储着用户与项目间的交互关系; 零矩阵表示用户节点间不存在直接连接, 符合实际应用场景中用户间无显式交互的特性.

为有效分离共享账户中的潜在用户行为模式, PE-GAT 将账户级序列图分解为用户级序列图. 具体而言, 每个共享账户  $U_k$  被建模为一组潜在用户节点  $\{U_{k,h}\}_{h=1}^{H_k}$ , 其中潜在用户数量  $H_k$  通过动态加权密度聚类算法 (详见第 2.4 节) 自适应确定. 每个潜在用户  $U_{k,h}$  的嵌入表示为  $\mathbf{e}_{U_{k,h}} \in \mathbb{R}^d$ , 模型通过聚合相邻节点信息来学习该表示. 这种基于动态用户数量推断的图分解方法, 为后续显式区分不同账户内的用户行为提供了基础表示空间.

## 2.4 潜在用户建模

本文设计了基于图注意力网络的预训练模型, 此类模型已被现有研究<sup>[20]</sup>证实能够有效建模共享账户下的多用户行为模式. 预训练模型采用单层 GAT 结构, 此设计基于以下考量: (1) 共享账户行为序列的局部关联特性, 单层注意力机制足以捕捉有效的直接依赖关系; (2) 符合模型的轻量化设计目标, 同时将复杂的模式挖掘交由后续的提示增强模块完成, 实现了架构上的功能解耦. 在信息聚合方面, PE-GAT 采用了神经图协同过滤 (neural graph collaborative filtering, NGCF)<sup>[46]</sup>的消息传递机制, 通过聚合直接相邻节点的信息来学习用户和项目的节点嵌入. 具体而言, 节点表示学习过程包含两个关键组成部分: 潜在用户表示学习和项目表示学习.

● 潜在用户表示学习用于捕捉共享账户下不同用户的偏好特征. 设  $I_i \in \mathcal{N}_{I_i}^{U_{k,h}}$  为  $U_{k,h}$  的项邻域, 其中  $\mathcal{N}_{I_i}^{U_{k,h}}$  为  $U_{k,h}$  的项邻居集, 则潜在用户表示学习的消息传递过程可以表示为:

$$\mathbf{m}_{U_{k,h} \leftarrow I_i} = \gamma_{I_i}^{U_{k,h}} \left( \mathbf{W}_1 \mathbf{e}_{I_i} + \mathbf{W}_2 (\mathbf{e}_{I_i} \odot \mathbf{e}_{U_{k,h}}) \right) \quad (3)$$

$$\mathbf{m}_{U_{k,h} \leftarrow U_{k,h}} = \gamma_{U_{k,h}}^{U_{k,h}} (\mathbf{W}_1 \mathbf{e}_{U_{k,h}}) \quad (4)$$

其中,  $\mathbf{m}_{U_{k,h} \leftarrow I_i}$  为从  $I_i$  到  $U_{k,h}$  的消息传递,  $\mathbf{W}_1, \mathbf{W}_2 \in \mathbb{R}^{d \times d}$  为可训练的权重矩阵.  $I_i$  与  $U_{k,h}$  之间的信息交互用元素积  $\mathbf{e}_{I_i} \odot \mathbf{e}_{U_{k,h}}$  表示, 其中,  $\mathbf{e}_{I_i}$  和  $\mathbf{e}_{U_{k,h}}$  分别为项目  $I_i$  和用户  $U_{k,h}$  的嵌入向量.  $\gamma_{I_i}^{U_{k,h}}$  是一个可学习的参数, 用于调控从项目  $I_i$  传递到用户  $U_{k,h}$  的信息量.  $\mathbf{m}_{U_{k,h} \leftarrow U_{k,h}}$  表示用户的自我连接, 旨在保留目标用户的固有信息.  $\gamma_{U_{k,h}}^{U_{k,h}}$  用于指导每条消息的传递强度, 该权重通过注意力机制计算得出.

采用项目感知的注意力机制, 以评估不同项目节点对目标用户的影响程度, 从而捕捉各个潜在用户对不同项目的个性化偏好. 定义项目  $I_i$  对  $U_{k,h}$  的重要性为:

$$s_{I_i}^{U_{k,h}} = f(\mathbf{e}_{U_{k,h}}, \mathbf{e}_{I_i}) \quad (5)$$

其中,  $f(\cdot, \cdot)$  是两两相似度度量, 采用余弦相似度计算. 进而通过归一化处理, 生成各项目的注意力权重值:

$$\gamma_{I_i}^{U_{k,h}} = \frac{\exp(s_{I_i}^{U_{k,h}})}{\sum_{I'_i \in \mathcal{N}_{I_i}^{U_{k,h}}} \exp(s_{I'_i}^{U_{k,h}}) + s_{U_{k,h}}^{U_{k,h}}} \quad (6)$$

其中,  $s_{U_{k,h}}^{U_{k,h}}$  为用户  $U_{k,h}$  内的自连接.  $\gamma_{U_{k,h}}^{U_{k,h}}$  可以通过替换相应的用户与项目嵌入以同样的方式得到.

随后, 通过以下方式更新用户  $U_{k,h}$  ( $\mathbf{e}_{U_{k,h}}$ ) 的嵌入, 对目标用户节点及其所有邻居节点的特征信息进行聚合:

$$\mathbf{e}_{U_{k,h}} = \text{LeakyReLU} \left( \mathbf{m}_{U_{k,h} \leftarrow U_{k,h}} + \sum_{I_i \in \mathcal{N}_{I_i}^{U_{k,h}}} \mathbf{m}_{U_{k,h} \leftarrow I_i} \right) \quad (7)$$

随后, 通过 LeakyReLU 激活函数对聚合后的特征表示进行处理. 最终, 模型将来自  $H$  个潜在用户表示空间的节点表示进行合并, 从而得到账户级表示  $\mathbf{e}_{U_k}$ .

$$\mathbf{e}_{U_k} = \frac{1}{H} \sum_{h=1}^H \mathbf{e}_{U_{k,h}} \quad (8)$$

● 项目表示学习旨在建模潜在用户与项目及项目间的交互模式. 设  $I_{i-1} \in \mathcal{N}_{I_{i-1}}^{I_i}$  和  $U_{k,h} \in \mathcal{N}_{U_{k,h}}^{I_i}$  分别表示连接的项目节点和用户节点, 其中,  $\mathcal{N}_{I_{i-1}}^{I_i}$  为  $I_i$  的项目邻居集,  $\mathcal{N}_{U_{k,h}}^{I_i}$  为  $I_i$  的用户邻居集. 项目表示学习的消息传递表示为:

$$\mathbf{m}_{I_i \leftarrow I_{i-1}} = \gamma_{I_{i-1}}^{I_i} (\mathbf{W}_1 \mathbf{e}_{I_{i-1}} + \mathbf{W}_2 (\mathbf{e}_{I_{i-1}} \odot \mathbf{e}_{I_i})) \quad (9)$$

$$\mathbf{m}_{I_i \leftarrow U_{k,h}} = \gamma_{U_{k,h}}^{I_i} (\mathbf{W}_1 \mathbf{e}_{U_{k,h}} + \mathbf{W}_2 (\mathbf{e}_{U_{k,h}} \odot \mathbf{e}_{I_i})) \quad (10)$$

$$\mathbf{m}_{I_i \leftarrow I_i} = \gamma_{I_i}^{I_i} (\mathbf{W}_i \mathbf{e}_{I_i}) \quad (11)$$

其中,  $\mathbf{m}_{I_i \leftarrow I_{i-1}}$  表示项目  $I_{i-1}$  传递给  $I_i$  的消息,  $\mathbf{m}_{I_i \leftarrow U_{k,h}}$  表示用户  $U_{k,h}$  传递给  $I_i$  的消息,  $\mathbf{m}_{I_i \leftarrow I_i}$  表示  $I_i$  的保留信息.  $\gamma_{I_{i-1}}^{I_i}$  为可学习的权重, 用于控制从  $I_{i-1}$  传递到  $I_i$  的信息量,  $\gamma_{U_{k,h}}^{I_i}$  和  $\gamma_{I_i}^{I_i}$  同理.

采用注意力机制量化目标用户与项目间关联性的重要程度. 项目  $I_{i-1}$  对  $I_i$  的重要性和用户  $U_{k,h}$  对项目  $I_i$  的重要性分别表示为:

$$s_{I_{i-1}}^{I_i} = f(\mathbf{e}_{I_i}, \mathbf{e}_{I_{i-1}}) \quad (12)$$

$$s_{U_{k,h}}^{I_i} = f(\mathbf{e}_{I_i}, \mathbf{e}_{U_{k,h}}) \quad (13)$$

随后, 对  $s_{I_{i-1}}^{I_i}$  进行归一化处理:

$$\gamma_{I_{i-1}}^{I_i} = \frac{\exp(s_{I_{i-1}}^{I_i})}{\sum_{I'_{i-1} \in \mathcal{N}_{I_{i-1}}^{I_i}} \exp(s_{I'_{i-1}}^{I_i}) + \sum_{U'_{k,h} \in \mathcal{N}_{U_{k,h}}^{I_i}} \exp(s_{U'_{k,h}}^{I_i}) + s_{I_i}^{I_i}} \quad (14)$$

其中,  $s_{I_i}^{I_i}$  表示项目  $I_i$  内的自连接的重要性,  $\gamma_{U_{k,h}}^{I_i}$  和  $\gamma_{I_i}^{I_i}$  可采用相同的计算方法获得.

项目节点通过聚合来自其邻居用户节点传递的所有消息来实现特征更新:

$$\mathbf{e}_{I_i} = \text{LeakyReLU} \left( \mathbf{m}_{I_i \leftarrow I_i} + \sum_{I_{i-1} \in \mathcal{N}_{I_{i-1}}^{I_i}} \mathbf{m}_{I_i \leftarrow I_{i-1}} + \sum_{U_{k,h} \in \mathcal{N}_{U_{k,h}}^{I_i}} \mathbf{m}_{I_i \leftarrow U_{k,h}} \right) \quad (15)$$

本文将用户嵌入划分为  $H$  个潜在空间, 并采用  $\mathbf{g}_{I_i}^h = \mathbf{e}_{I_i}$  来表示项目  $I_i$  在第  $h$  个潜在用户空间中的嵌入表示, 则项目  $I_i$  的全局表示  $\mathbf{g}_{I_i}$  可以定义为:

$$\mathbf{g}_{I_i} = \frac{1}{H} \sum_{h=1}^H \mathbf{g}_{I_i}^h \quad (16)$$

至此, 得到预训练后的用户和项目表示.

另一方面, 为了自适应地识别共享账户中的潜在用户, 本文设计了动态加权密度空间聚类算法, 通过特征加权与参数动态调节的双重机制改进传统 DBSCAN 框架<sup>[47]</sup>. 相较于传统 K-means 等需要预设聚类数量的算法, 我们的算法具有以下显著优势: (1) 通过邻域密度估计自主识别数据中的自然聚类结构, 无需预先设定潜在用户数量; (2) 通过特征加权提升混合行为数据的聚类质量; (3) 基于 K 近邻分布自适应调整邻域半径  $\varepsilon$ , 避免人工调参负担. 这些特性使得动态加权密度聚类算法特别适用于共享账户场景下的用户识别任务, 为后续的提示增强和模型再训练提供了可靠的聚类基础. 具体算法实现步骤如算法 1 所示.

---

**算法 1.** 基于动态加权密度聚类算法的共享账户自适应用户识别.

---

输入: 项目集合  $S = \{I_1, I_2, \dots, I_i\}$ , 特征矩阵  $\mathbf{X}$ , 邻域半径  $\varepsilon$ , 最小半径  $\varepsilon_{\min}$ , 最大半径  $\varepsilon_{\max}$ , 基础密度  $MinPts$ , 最小密度  $MinPts_0$ , 最大密度  $MaxPts$ , 特征权重矩阵  $\mathbf{W}_d$ ;

输出: 聚类结果  $C_i = \{C_{i1}, C_{i2}, \dots, C_{ik}\}$ , 噪声点集合  $\Omega$ .

---

**begin**

1. **for each**  $x \in \mathbf{X}$  **do**: /\* 特征加权处理 \*/
  2.    $x \leftarrow x \odot \mathbf{W}_d$  /\* 应用基于先验知识的特征权重矩阵 \*/
  3. **end for**
  4.  $\mathbf{X} \leftarrow (\mathbf{X} - \mu) / \sigma$  /\* 标准化特征矩阵 \*/
  5.  $\varepsilon \leftarrow \text{EstimateEps}(\mathbf{X})$  /\* 基于 KNN 距离估计的动态参数计算 \*/
  6.  $\varepsilon \leftarrow \max(\min(\varepsilon, \varepsilon_{\max}), \varepsilon_{\min})$  /\* 半径约束 \*/
  7.  $MinPts \leftarrow \text{CalcMinPts}(|S|)$  /\* 基于样本量计算 \*/
-

---

```

8.  $MinPts \leftarrow \min(\max(MinPts, MinPts_0), MaxPts)$  /* 密度约束 */
9.  $O \leftarrow \emptyset, \Gamma \leftarrow S$  /* 初始化核心对象集和未访问集 */
10. for each  $I_i \in S$  do: /* 核心对象识别 */
11.   if  $|N_\varepsilon(I_i)| \geq MinPts$  then
12.      $O \leftarrow O \cup \{I_i\}$  /*  $\varepsilon$ -邻域项目计数 */
13.   end if
14. end for
15. while  $O \neq \emptyset$  do: /* 密度聚类 */
16.    $I_o \leftarrow \text{random}(O)$  /* 随机选择核心对象 */
17.    $Q \leftarrow \{I_o\}, C_k \leftarrow \emptyset$  /* 初始化队列和聚类簇 */
18.   while  $O \neq \emptyset$  do:
19.      $I_q \leftarrow \text{dequeue}(Q)$ 
20.     if  $I_q \in O$ :
21.        $C_k \leftarrow C_k \cup \{I_q\}$  /* 加入当前簇 */
22.        $Q \leftarrow Q \cup |N_\varepsilon(I_q)|$  /* 扩展邻域 */
23.     end if
24.      $\Gamma \leftarrow \Gamma \setminus \{I_q\}$  /* 标记为已访问 */
25.   end while
26.    $k \leftarrow k + 1$ 
27.    $O \leftarrow O \setminus C_k$  /* 移除已聚类对象 */
28. end while
29.  $\Omega \leftarrow \Gamma$  /* 标记剩余项为噪声 */
30. output  $C, \Omega$  /* 输出聚类结果 */
end

```

---

为了自动有效地构建个性化的提示, 本文采用多层感知机 (multilayer perceptron, MLP) 来构建提示生成器. MLP 能够建模提示与用户之间的非线性交互关系, 并生成更具判别性的提示嵌入. 具体地, 对于项目  $I_i$ , 从其对应的提示中选取  $m$  个, 并将其嵌入拼接构成提示输入向量:

$$x^u = [a_1^u || a_2^u || \dots || a_m^u] \in \mathbb{R}^{m \cdot d_1} \quad (17)$$

其中,  $a_i^u \in \mathbb{R}^{d_1}$  为提示嵌入. 然后, 将  $x^u$  输入至一个两层的多层感知机中, 生成该项目对应的个性化提示表示  $P^u \in \mathbb{R}^{d_2}$ , 如公式 (18) 所示:

$$P^u = W_a \cdot \sigma(W_b x^u + b_1) + b_2 \quad (18)$$

其中,  $W_a \in \mathbb{R}^{(m \cdot d_1) \times d_2}$ ,  $W_b \in \mathbb{R}^{d_1 \times d}$ ,  $b_1 \in \mathbb{R}^{d_1}$ ,  $b_2 \in \mathbb{R}^{d_2}$  为可学习的参数;  $\sigma(\cdot)$  为非线性激活函数;  $d_1$  为隐藏层维度,  $d$  为最终输出维度, 与预训练嵌入维度一致. 该策略生成的提示简单高效, 也可采用其他神经网络架构作为替代构建提示生成器.

## 2.5 模型再训练

为了优化从预训练模型中学习到的初始序列表示, PE-GAT 设计了提示增强模块. 该模块采用注意力机制作为核心编码器, 具有显著优势: 首先, 注意力机制通过建模用户行为序列的上下文关联, 有效捕捉行为间的局部交互模式; 其次, 其特有的全局注意力计算方式能够量化序列中任意两个行为节点的相关性, 从而实现长期依赖关系和序列全局特征的精确建模.

在注意力机制中, 查询 ( $Q$ )、键 ( $K$ ) 和值 ( $V$ ) 的交互注意力通过点积运算实现:

$$Attention(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V} \quad (19)$$

具体而言,  $\mathbf{Q}$  与  $\mathbf{K}$  的转置矩阵相乘 ( $\mathbf{Q}\mathbf{K}^T$ ) 以计算查询与键之间的相似性, 随后通过缩放因子  $\sqrt{d}$  ( $d$  为向量维度) 对结果进行标准化, 避免点积值过大导致  $\text{Softmax}$  归一化后出现梯度饱和问题. 最终,  $\text{Softmax}$  函数将  $\mathbf{Q}\mathbf{K}^T / \sqrt{d}$  的输出转化为权重总和为 1 的概率分布, 从而动态分配不同行为间的注意力权重. 这一机制不仅能有效建模用户行为的局部相关性, 还能通过全局注意力捕获长期依赖关系, 提升序列表征的判别能力.

基于注意力机制的理论框架, 首先将初始序列表示  $\mathbf{Y} = \{\mathbf{g}_{l_1}, \mathbf{g}_{l_2}, \dots, \mathbf{g}_{l_i}\}$  和提示信息表示  $\mathbf{P} = \{\mathbf{e}_{p_1}, \mathbf{e}_{p_2}, \dots, \mathbf{e}_{p_m}\}$  进行变换, 分别生成查询矩阵  $\mathbf{Q} = \mathbf{Y}\mathbf{W}_q$ 、键矩阵  $\mathbf{K} = \mathbf{P}\mathbf{W}_k$  和值矩阵  $\mathbf{V} = \mathbf{P}\mathbf{W}_v$ . 其中,  $\mathbf{W}_q, \mathbf{W}_k, \mathbf{W}_v \in \mathbb{R}^{d \times d}$  为可学习的权重矩阵. 将这些矩阵输入到注意力网络中, 其计算过程可形式化表示为:

$$\mathbf{Z} = \text{Attention}(\mathbf{Y}\mathbf{W}_q, \mathbf{P}\mathbf{W}_k, \mathbf{P}\mathbf{W}_v) \quad (20)$$

输出的增强序列提示模板  $\mathbf{Z}$  与原始序列表示  $\mathbf{Y}$  再通过特征维度拼接操作进行融合, 形成增强表示:

$$\hat{\mathbf{Y}} = \text{Concat}([\mathbf{Y}, \mathbf{Z}]) \quad (21)$$

该提示增强模块使得模型能够同时捕获原始序列特征和注意力增强的全局依赖关系, 从而提升下游任务的性能表现. 得到增强表示之后, PE-GAT 需要构建最终的嵌入表示形式以进行推荐. 具体而言, 将提示增强模块得到的增强序列表示  $\hat{\mathbf{Y}}$  与预训练模型得到的共享账户表示整合后重新输入到预训练模型中进行再训练. 通过预训练模型的消息传递层更新节点表示:

$$\mathbf{h}^{(l)} = \text{GAT}(\mathbf{h}^{(l-1)}, \mathcal{G}) \quad (22)$$

其中,  $l$  表示图卷积层数,  $\mathbf{h}^{(l)}$  表示第  $l$  层图卷积后的节点表示,  $\mathbf{h}^{(l-1)}$  表示第  $l-1$  层图卷积后的节点表示. 基于最终层输出, 采用均值池化操作生成新的项目嵌入表示:

$$\mathbf{g}'_{l_i} = \frac{1}{H} \sum_{h=1}^H \mathbf{g}^h_{l_i} \quad (23)$$

再训练后得到更新后的提示序列表示  $\mathbf{h}_s$ , 对序列内的项嵌入  $(\mathbf{g}'_{l_1}, \mathbf{g}'_{l_2}, \dots, \mathbf{g}'_{l_i})$  进行最大池化 (max-pooling) 操作, 可以得到  $S_l$  的序列级表示  $\mathbf{h}'_s$ . 随后, 将其与新的账户嵌入  $\mathbf{e}'_{u_k}$  进行融合, 并将融合后的嵌入输入以下预测层:

$$P(I_{i+1} | S) \sim f(S) = \text{Softmax}\left(\mathbf{W}_l \cdot [\mathbf{h}'_s, \mathbf{e}'_{u_k}]^T + b_l\right) \quad (24)$$

其中,  $\mathbf{W}_l$  和  $b_l$  分别为可训练权重矩阵和偏置项. 模型通过最小化负对数似然损失函数进行优化:

$$L(\theta) = -\frac{1}{|\mathbb{S}|} \sum_{S \in \mathbb{S}} \sum_{I_i \in S} \log P(I_{i+1} | S) \quad (25)$$

其中,  $\theta$  表示本模型的所有参数,  $\mathbb{S}$  表示训练序列.

## 2.6 复杂度分析

动态加权密度聚类的时间复杂度为  $\mathcal{O}(A\bar{L}\log\bar{L})$ , 其中  $A$  为数据集中共享账户数量,  $\bar{L}$  为平均序列长度. 单层图注意力网络的计算复杂度为  $\mathcal{O}(H|\mathcal{E}|d)$ ,  $|\mathcal{E}|$  为序列图的边数,  $H$  为注意力头数,  $d$  为特征维度. 基于 MLP 的提示生成过程的时间复杂度为  $\mathcal{O}(d^2)$ , 其中  $d$  为提示表示维度. 提示融合阶段的计算开销为  $\mathcal{O}(d)$ . 综上, PE-GAT 的总体时间复杂度可表示为  $\mathcal{O}(A\bar{L}\log\bar{L}) + \mathcal{O}(H|\mathcal{E}|d)$ , 由于序列图具有稀疏性且共享账户中的交互序列通常较短, 模型的整体计算开销可控, 在实际场景中具备良好的可扩展性.

## 3 实验分析

本文在 2 个真实数据集上进行实验以验证 PE-GAT 的有效性, 重点回答以下 4 个研究问题.

RQ1: PE-GAT 模型在共享账户序列推荐场景中的表现如何? (第 3.4 节)

RQ2: PE-GAT 模型中不同成分有何影响? (第 3.5 节)

RQ3: 超参数如何影响 PE-GAT 模型的性能? (第 3.6 节)

RQ4: PE-GAT 模型的训练效率如何? (第 3.7 节)

### 3.1 实验数据

本文采用 HVIDEO-E 和 HVIDEO-V 两个真实场景下的智能电视数据集来评估 PE-GAT 模型的性能. HVIDEO-E 数据集包含面向各年级学生的教育视频和日常生活教学视频, 而 HVIDEO-V 数据集则涵盖了视频点播平台上的电影、动画和电视剧等多种视频内容. 这两个数据集均包含家庭账户在多个电视频道的观看历史数据, 具有较好的现实代表性. 应当指出, HAMZON 数据集虽属于共享账户场景, 但由于其为人工生成数据, 在预实验中自适应聚类效果不佳, 因此未被纳入正式实验.

实验中, 所有序列按照 85% 与 15% 的比例随机划分为训练集和测试集, 数据集的详细统计信息列于表 1.

表 1 实验数据集的统计信息

| 类别     | HVIDEO-E  | HVIDEO-V  |
|--------|-----------|-----------|
| 用户数量   | 8 367     | 11 404    |
| 交互数量   | 2 129 500 | 1 893 784 |
| 平均序列长度 | 15.85     | 14.09     |

### 3.2 评价指标及基线方法

● 评价指标. 本文采用每个数据集的每个序列中最后观看的项目作为基础事实, 并采用平均倒数排名 (*MRR*) 和召回率 (*Recall*) 两个评价指标进行系统性能的量化分析. 具体而言, 本文采用 *MRR@5*、*MRR@20*、*Recall@5* 以及 *Recall@20* 这 4 个指标来衡量实验结果.

(1) *MRR* 是衡量推荐系统准确度的常用指标, 用于量化检索结果列表中首个相关项目位置的倒数排名 (即平均倒数排名). 当用户查询的第 1 个相关文档在检索结果中的排名较高时, 相应的 *MRR* 值亦较高. *MRR@N* 的计算公式如下:

$$MRR@N = \frac{1}{|Q|} \sum_{q \in Q} \frac{1}{\sum_{k \in [1, N]} R_q(k)} \quad (26)$$

其中,  $Q$  为查询数量,  $q \in Q$  为一个具体的查询实例,  $R_q(k)$  则表示在执行查询  $q$  时, 召回结果集中第  $k$  个项目所对应的排名位置.

(2) *Recall* 是评估推荐系统性能的另一个关键指标, 用于量化推荐列表中正确推荐的项目数占测试集中所有真实相关项目总数的比例. 较高的召回率表明系统能够有效减少遗漏. *Recall* 的计算公式如下:

$$Recall = \frac{|TP|}{|TP| + |FN|} \quad (27)$$

其中, *TP* (true positive) 表示被系统正确推荐的相关项目, *FN* (false negative) 表示未被系统推荐的相关项目.

● 基线方法. 为了评估 PE-GAT 模型的性能, 本文将其与 3 类主流推荐方法进行对比实验. 这些对比方法涵盖了从传统推荐到针对特定场景优化的复杂模型.

#### (1) 传统推荐方法

BPR-MF<sup>[53]</sup>: 通过矩阵分解预测用户偏好, 并采用贝叶斯个性化排序优化推荐排序质量.

Item-KNN<sup>[54]</sup>: 通过计算目标项目与其他项目的特征相似度, 选择最相似的项目集合以预测用户偏好.

NCF<sup>[48]</sup>: 将用户历史行为序列化, 并采用基于多层感知机 (MLP) 的协同过滤技术, 学习用户嵌入表示.

LightGCN<sup>[49]</sup>: 一种基于图卷积网络的协同过滤算法, 通过直接聚合节点及其邻居的特征来学习用户和项目的深层嵌入.

#### (2) 序列推荐方法

GRU4REC<sup>[61]</sup>: 采用门控循环单元来处理序列数据, 并通过优化基于排名的损失函数, 成为实现序列推荐的一

种典范方法.

HGRU4REC<sup>[33]</sup>: 通过构建层次化的门控循环单元网络, 融合了用户会话中的短期与长期兴趣, 从而对 GRU4REC 模型进行增强与优化.

NAIS<sup>[34]</sup>: 一种基于项目的协同过滤算法, 通过采用注意力机制来区分不同历史交互的重要程度.

Time-LSTM<sup>[50]</sup>: 通过设计时间门控机制 (temporal gate mechanism) 来捕捉用户行为间的时间间隔特征, 同步优化了短期兴趣捕捉与长期兴趣建模.

TGSRec<sup>[51]</sup>: 这是另一种引入时间间隔的序列推荐方法, 该方法利用协作注意力网络显式地模拟项目间的协作信号与时间间隔, 并将时间跨度的表示转化为相应时间嵌入的核函数值.

### (3) 共享账户推荐方法

VUI-KNN<sup>[22]</sup>: 通过将用户观看日志划分为虚拟用户生成的片段, 进而计算并聚类这些虚拟用户的相似性, 构建潜在用户群体, 并利用用户最近邻算法进行推荐.

$\pi$ -Net<sup>[19]</sup>: 一种用于解决共享账户跨域序列推荐问题的并行信息共享网络, 通过共享账户过滤单元 (shared account filtering unit, SFU) 和跨域传输单元 (cross-domain transfer unit, CTU) 并行提取和共享信息, 以提升两个领域的推荐性能.

PSJNet<sup>[12]</sup>: 通过“分割”与“连接”机制, 并行地学习用户角色信息和跨领域信息, 对  $\pi$ -Net 算法进行了深入优化与改进.

DA-GCN<sup>[20]</sup>: 一种基于图卷积神经网络的方法, 该方法通过构建跨域序列图, 并采用领域感知的图卷积网络, 汇聚邻接节点的信息以学习用户与项目的嵌入表示, 取得了良好的效果.

TiDA-GCN<sup>[13]</sup>: 在 DA-GCN 模型的基础上, 融入时间间隔信息, 采用时间间隔感知的消息传递策略, 并结合账户感知的自注意力机制, 以学习项目的交互特征, 从而更有效地捕捉用户的行为模式.

LightGC<sup>2</sup>N<sup>[14]</sup>: 一种最新的共享账户序列推荐方法, 该模型通过轻量级图胶囊卷积网络实现了交互行为与潜在用户间的精细匹配, 并利用子空间对齐策略优化序列表示, 在提升精度的同时保证了模型的高效性.

## 3.3 实验方法

实验基于 NVIDIA RTX 3090 GPU 及 TensorFlow 框架实现. 模型参数与动态可学习权重矩阵均采用 Xavier 方法<sup>[52]</sup>随机初始化, 并采用 Adam 算法优化. 学习率设定为 0.001, 批量大小设定为 256. 在两个数据集的模型训练阶段, dropout 率设为 0.1. 关于模型超参数, 用户与项目的嵌入维度均被设置为 16. 在提示增强与再训练阶段, 本文在 50%–95% 的范围内, 以步长为 5% 的方式探索聚类空间中 K 距离分位数的最优值. 为确保对比实验的公正性, 本文将所有基线方法的嵌入维度统一设定为 16. 至于基线方法的其他超参数, 则均采用各方法原文献中的最佳设置.

## 3.4 实验结果与分析

表 2 展示了 PE-GAT 模型与其他 3 类基线方法在 HVIDEO-E 和 HVIDEO-V 数据集上的对比实验结果. 经过对实验数据的观察与分析, 本文提出的 PE-GAT 模型在多个指标上均表现出显著优势, 验证了其在处理共享账户序列推荐任务中的有效性. 实验结果表明, PE-GAT 模型在 HVIDEO-E 数据集上的所有指标均优于其他基线方法, 且相较于次优结果实现了显著提升. 在 HVIDEO-V 数据集上, 该模型同样展现出良好性能, 在 Recall@5、Recall@20 和 MRR@5 这 3 项指标上均达到最优. 这一结果证实了 PE-GAT 方法在处理共享账户序列推荐任务中的有效性.

相较于传统推荐方法 (BPR-MF、Item-KNN、NCF 和 LightGCN), PE-GAT 模型强调共享账户特征的重要性, 其通过整合共享账户信息, 更精确地捕捉多用户行为模式, 进而提升了推荐质量. PE-GAT 模型在混合行为用户推荐任务中, 相较于序列推荐方法 (GRU4REC、HGRU4REC、NAIS、Time-LSTM 和 TGSRec), 展现出基于 GCN 模型的优越性. 同时, 还揭示了图结构消息传递机制在识别用户复杂交互模式中的关键作用. 与基于 RNN 的方法 (例如 GRU4REC、HGRU4REC、 $\pi$ -Net 和 PSJNet) 相比, PE-GAT 模型证明了基于图神经网络的方法在建模用户与项目间复杂关系方面的有效性. 与基于共享账户推荐方法 (如 DA-GCN、TiDA-GCN 和 LightGC<sup>2</sup>N) 相比, PE-GAT 模型进一步验证了本文所提出的基于提示增强图注意力网络的共享账户序列推荐方法的有效性.

表2 HVIDEO-E 与 HVIDEO-V 数据集上的对比实验结果 (%)

| 方法                     | HVIDEO-E     |               |                 |                  | HVIDEO-V     |               |                 |                  |
|------------------------|--------------|---------------|-----------------|------------------|--------------|---------------|-----------------|------------------|
|                        | <i>MRR@5</i> | <i>MRR@20</i> | <i>Recall@5</i> | <i>Recall@20</i> | <i>MRR@5</i> | <i>MRR@20</i> | <i>Recall@5</i> | <i>Recall@20</i> |
| BPR-MF                 | 2.81         | 3.32          | 2.51            | 5.58             | 4.13         | 7.64          | 5.27            | 9.33             |
| Item-KNN               | 3.53         | 4.19          | 3.74            | 9.42             | 5.15         | 12.96         | 8.84            | 13.19            |
| NCF                    | 5.78         | 7.86          | 11.26           | 20.13            | 23.00        | 24.72         | 27.22           | 22.64            |
| LightGCN               | 11.56        | 13.57         | 20.69           | 39.93            | 53.99        | 56.73         | 58.46           | 63.56            |
| HGRU4REC               | 13.77        | 16.15         | 22.56           | 47.99            | 60.59        | 63.32         | 67.99           | 73.25            |
| GRU4REC                | 13.34        | 15.51         | 22.48           | 44.19            | 61.79        | 62.48         | 66.61           | 73.24            |
| NAIS                   | 11.46        | 13.25         | 19.79           | 40.18            | 51.53        | 54.48         | 59.29           | 67.42            |
| Time-LSTM              | 11.20        | 13.28         | 20.18           | 41.46            | 59.74        | 60.43         | 64.67           | 71.60            |
| TGSRec                 | 13.96        | 15.74         | 19.92           | 41.79            | 50.33        | 53.41         | 58.92           | 67.23            |
| VUI-KNN                | 3.22         | 4.09          | 3.45            | 6.93             | 4.33         | 8.29          | 7.86            | 9.09             |
| $\pi$ -Net             | 15.37        | 17.53         | 25.14           | 47.09            | 60.38        | 61.75         | 66.99           | 74.18            |
| PSJNet                 | 15.38        | 17.57         | 24.79           | 46.69            | 61.90        | 62.64         | 66.87           | 74.15            |
| DA-GCN                 | 35.64        | 37.28         | 51.36           | 66.94            | 59.79        | 60.54         | 75.40           | 82.38            |
| TiDA-GCN               | 38.66        | 40.23         | 54.11           | 68.98            | 63.58        | <u>65.37</u>  | 76.37           | 83.58            |
| LightGC <sup>2</sup> N | <u>42.13</u> | <u>43.62</u>  | <u>57.28</u>    | <u>69.24</u>     | <u>63.62</u> | <b>65.56</b>  | <u>78.01</u>    | <u>83.62</u>     |
| PE-GAT                 | <b>43.39</b> | <b>44.48</b>  | <b>59.70</b>    | <b>69.78</b>     | <b>63.71</b> | 64.22         | <b>79.17</b>    | <b>83.67</b>     |

注: 加粗数据表示最优结果, 加下划线数据表示次优结果

### 3.5 消融实验

本文在两个数据集上对 PE-GAT 模型进行消融实验, 以探究不同组件对模型性能的具体贡献. 如表 3 所示, 本文设计了 10 种模型变体: (1) PE w/o D: 此变体移除了动态加权密度聚类算法, 并以随机生成数据替代聚类结果; (2) PE-D: 此变体将动态加权密度聚类算法替换为 DBSCAN 聚类方法; (3) PE-K: 此变体将动态加权密度聚类算法替换为 K 均值聚类方法, 并指定聚类数量为 3; (4) PE-H: 此变体将动态加权密度聚类算法替换为层次聚类方法, 固定聚类数量为 3; (5) PE w/o N: 此变体移除了预训练图注意网络中基于动态加权密度聚类的自适应潜在用户数量预测模块, 改为采用统一预设的潜在用户数量, 设置为 3; (6) PE w/o AM: 此变体移除了注意力机制组件, 仅将提示嵌入与序列嵌入进行简单拼接; (7) PE-T: 此变体将原模型中基于 MLP 的提示生成器替换为 Transformer 结构, 并保持其他模块不变; (8) PE-A: 此变体将提示融合策略中的拼接融合结构替换为加性融合结构. (9) PE w/o PE: 此模型变体移除了提示增强与再训练组件, 并将预训练生成的输入作为最终预测结果; (10) PE w/o PT: 此变体移除了预训练过程, 直接采用端到端方式训练 PE-GAT 模型.

表3 PE-GAT 在 HVIDEO-E 与 HVIDEO-V 数据集上的消融实验结果 (%)

| 方法及变体     | HVIDEO-E     |               |                 |                  | HVIDEO-V     |               |                 |                  |
|-----------|--------------|---------------|-----------------|------------------|--------------|---------------|-----------------|------------------|
|           | <i>MRR@5</i> | <i>MRR@20</i> | <i>Recall@5</i> | <i>Recall@20</i> | <i>MRR@5</i> | <i>MRR@20</i> | <i>Recall@5</i> | <i>Recall@20</i> |
| PE w/o D  | 40.67        | 41.91         | 57.32           | 68.54            | 62.53        | 63.00         | 78.53           | 82.92            |
| PE-D      | 42.87        | 43.99         | 59.37           | 69.33            | 63.63        | 64.12         | 79.06           | 83.15            |
| PE-K      | 42.96        | 44.09         | 59.38           | 69.57            | 63.50        | 64.03         | 78.96           | 83.39            |
| PE-H      | 43.16        | 44.29         | 59.40           | 69.58            | 63.48        | 64.01         | 79.02           | 83.59            |
| PE w/o N  | 41.40        | 42.68         | 58.34           | 69.65            | 62.76        | 63.42         | 78.68           | 83.56            |
| PE w/o AM | 42.94        | 44.05         | 59.38           | 69.28            | 63.54        | 64.06         | 78.95           | 83.28            |
| PE-T      | 43.02        | 44.16         | 59.42           | 69.32            | 63.18        | 63.34         | 79.02           | 83.41            |
| PE-A      | 42.31        | 43.67         | 59.12           | 69.08            | 62.74        | 63.89         | 78.51           | 83.12            |
| PE w/o PE | 37.35        | 38.84         | 54.81           | 68.62            | 60.96        | 61.60         | 77.71           | 83.09            |
| PE w/o PT | 34.46        | 35.83         | 49.58           | 65.51            | 57.94        | 58.05         | 75.89           | 80.33            |
| PE-GAT    | <b>43.39</b> | <b>44.48</b>  | <b>59.70</b>    | <b>69.78</b>     | <b>63.71</b> | <b>64.22</b>  | <b>79.17</b>    | <b>83.67</b>     |

注: 加粗数据表示最优结果

针对表 3 所呈现的实验数据, 本文进行了以下分析: (1) PE-GAT 模型在两个数据集上的各项指标性能均优于 PE w/o D, 这一结果凸显了本文设计的动态加权密度聚类算法对潜在用户识别的有效性. (2) PE-GAT 模型相较于 PE-D 变体表现出一定的性能提升, 体现了本文设计的加权机制的有效性, 同时动态机制的设计减少了超参数调优的实验成本; (3) PE-GAT 模型相较于 PE-K 与 PE-H 变体表现出更优的性能. 尽管 K 均值聚类方法与层次聚类方法可以作为有效的替代, 但此类方法需要指定聚类数量, 而本文设计的动态加权密度聚类算法无需此限制. (4) PE-GAT 模型相较于 PE w/o N 变体性能更优, 这表明基于动态加权密度聚类的自适应用户数量预测机制比固定数量设定能更准确地识别潜在用户行为模式. (5) PE-GAT 模型相较于 PE w/o AM 展现出更佳的性能, 这体现了注意力机制模块在融合嵌入信息过程中所发挥的关键作用. (6) PE-GAT 模型相较于 PE-T 变体保持了一定的优势, 尽管 Transformer 结构具备有效的表征能力, 但本文设计的基于 MLP 的提示生成器在计算开销与性能之间取得了更优的平衡. (7) PE-GAT 模型相较于 PE-A 变体表现更佳, 证明了拼接融合结构在处理提示信息时效果更佳. (8) PE-GAT 模型相较于 PE w/o PE 表现出更优异的性能, 证实了提示增强模块对于提升模型性能的重要性. (9) PE-GAT 模型相较于 PE w/o PT 变体也展示出明显优势, 表明预训练阶段能够有效提升模型的初始化质量, 从而在提示增强与再训练阶段获得更优的性能.

### 3.6 参数影响分析

本节旨在探讨 K 距离分位数这一重要超参数对聚类效果的影响. 该参数用于估计密度聚类算法中的邻域半径  $\epsilon$ , 具体方法为: 基于样本点与其第 K 个最近邻点之间的距离分布, 选取某一特定百分位数作为  $\epsilon$  的取值阈值. 换言之, K 距离分位数决定了在 K 近邻距离分布中所采用的百分位位置, 进而直接影响聚类边界的划分. 为探索超参数的重要影响, 本文在 HVIDEO-E 和 HVIDEO-V 数据集上以 5% 的步长在 [50%, 95%] 区间内设计了 10 组 K 距离分位数参数实验, 以确定其最优取值. 图 3 展示了实验结果.

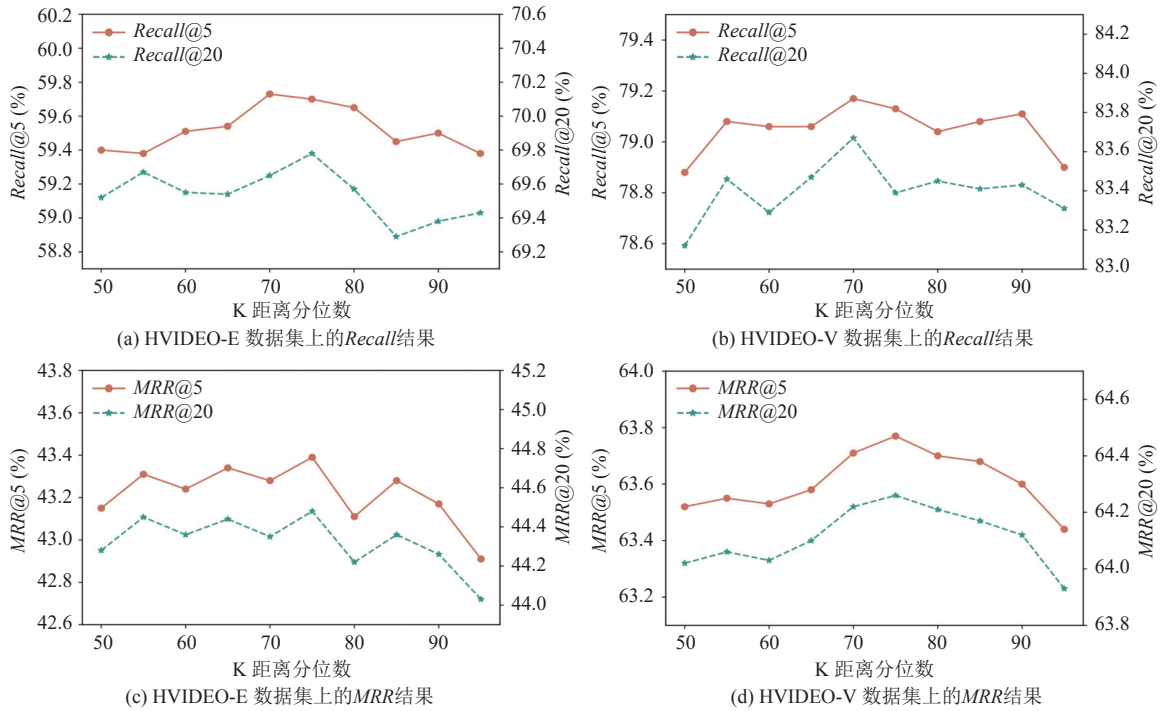


图 3 HVIDEO-E 与 HVIDEO-V 数据集上 K 距离分位数的影响

在 HVIDEO-E 数据集上, K 距离分位数设置为 75% 时, 模型性能达到最佳, 而对于 HVIDEO-V 数据集, 最佳性能出现在 70% 分位数. 这表明适中的邻域范围能够有效捕捉数据相关性, 分位数过低会导致邻域范围过小, 遗

漏潜在相关点. 例如, 当分位数设置为 50% 时, HVIDEO-E 数据集上的  $Recall@5$  从 59.70% 降至 59.40%; 反之, 分位数过高则会使邻域过度扩张, 引入噪声干扰, 如分位数为 90% 时,  $Recall@5$  下降至 59.38%.

总结而言, 实验结果揭示了  $K$  距离分位数对聚类算法性能的重要影响. 适当的  $K$  距离分位数是确保聚类结果准确性与有效性的关键. 不当的参数设置可能导致聚类性能降低, 表现为遗漏相关点或引入噪声干扰.

### 3.7 训练效率和可扩展性

为量化模型训练效率, 本节在 HVIDEO-E 和 HVIDEO-V 数据集上测试了不同数据比例 (0.2, 0.4, 0.6, 0.8, 1.0) 对训练耗时的影响, 如图 4 所示. 实验结果表明, 虽然 PE-GAT 的训练效率略低于 DA-GCN, 但相较其他对比模型仍具有明显优势. 具体而言, 随着训练数据规模增大, PE-GAT 与 DA-GAT 的训练时间增长速度平稳, 曲线位置始终低于其他对比模型, 展现出良好的可扩展性. 这种稳定的计算复杂度特征表明, PE-GAT 能够有效适应大规模数据场景.

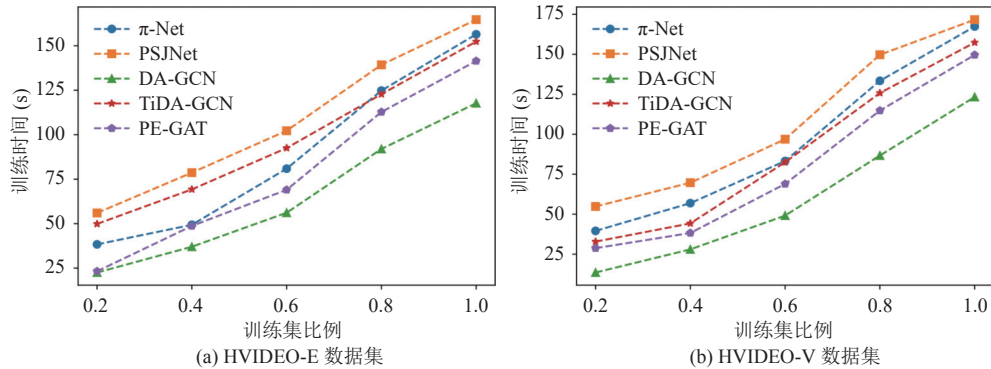


图 4 HVIDEO-E 与 HVIDEO-V 数据集上模型训练效率对比

此外, 尽管本文以典型的共享账户场景为主要研究对象, 所提出的潜在用户识别与偏好解耦框架在更复杂的实际系统中亦具有较强的适用性. 在智慧城市与物联网领域, 可通过解耦匿名使用记录以优化资源调度; 在金融科技领域, 可通过建立多用户交易基线提升异常检测精度; 在网联智能环境中, 可区分使用者偏好以实现深度个性化服务.

## 4 总 结

本文提出一种基于提示增强图注意力网络的自适应共享账户序列推荐 (PE-GAT) 方法, 以解决共享账户场景下潜在用户偏好解耦困难和潜在用户数量难以自适应确定的挑战. 该方法首先设计了一种动态加权密度聚类算法, 基于聚类结果构建用户级序列图, 并采用基于图注意力网络的模型进行预训练, 以学习用户及项目表示, 从而有效区分不同潜在用户的行为模式. 其次, PE-GAT 采用动态加权密度聚类算法自适应识别共享账户中的潜在用户作为后验提示信息, 并设计了基于注意力机制的提示增强模块, 通过生成提示模板优化初始序列表示并进行再训练, 最终通过账户级表示与提示增强表示的融合来完成个性化推荐任务. 一系列实验分析表明, 本文所提出的模型在共享账户序列推荐任务中取得了显著的性能提升, 验证了其在增强推荐系统个性化能力方面的有效性.

## References

- [1] Li J, Ren PJ, Chen ZM, Ren ZC, Lian T, Ma J. Neural attentive session-based recommendation. In: Proc. of the 26th ACM Int'l Conf. on Information and Knowledge Management. Singapore: ACM, 2017. 1419–1428. [doi: [10.1145/3132847.3132926](https://doi.org/10.1145/3132847.3132926)]
- [2] Sun F, Liu J, Wu J, Pei CH, Lin X, Ou WW, Jiang P. BERT4Rec: Sequential recommendation with bidirectional encoder representations from Transformer. In: Proc. of the 28th ACM Int'l Conf. on Information and Knowledge Management. Beijing: ACM, 2019. 1441–1450. [doi: [10.1145/3357384.3357895](https://doi.org/10.1145/3357384.3357895)]
- [3] Xie Z, Liu CX, Zhang YC, Lu HT, Wang D, Ding Y. Adversarial and contrastive variational autoencoder for sequential recommendation.

- In: Proc. of the 2021 Web Conf. Ljubljana: ACM, 2021. 449–459. [doi: [10.1145/3442381.3449873](https://doi.org/10.1145/3442381.3449873)]
- [4] Xie X, Sun F, Liu ZY, Wu SW, Gao JY, Zhang JD, Ding BL, Cui B. Contrastive learning for sequential recommendation. In: Proc. of the 38th IEEE Int'l Conf. on Data Engineering. Kuala Lumpur: IEEE, 2022. 1259–1273. [doi: [10.1109/ICDE53745.2022.00099](https://doi.org/10.1109/ICDE53745.2022.00099)]
  - [5] Qian ZS, Ye ZL, Yao CS, Zhang D, Huang H, Qin LY. Lightweight GCN recommendation combining adaptive period and interest factor. Ruan Jian Xue Bao/Journal of Software, 2024, 35(6): 2974–2998 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6915.htm> [doi: [10.13328/j.cnki.jos.006915](https://doi.org/10.13328/j.cnki.jos.006915)]
  - [6] Hidasi B, Karatzoglou A, Baltrunas L, Tikk D. Session-based recommendations with recurrent neural networks. In: Proc. of the 4th Int'l Conf. on Learning Representations. San Juan: OpenReview.net, 2016. 1–10.
  - [7] Qiu RH, Huang Z, Li JJ, Yin HZ. Exploiting cross-session information for session-based recommendation with graph neural networks. ACM Trans. on Information Systems, 2020, 38(3): 22. [doi: [10.1145/3382764](https://doi.org/10.1145/3382764)]
  - [8] Guo L, Li QJ, Liu FA, Wang XH. Shared-account cross-domain sequential recommendation with self-attention network. Journal of Computer Research and Development, 2021, 58(11): 2524–2537 (in Chinese with English abstract). [doi: [10.7544/issn1000-1239.2021.20200564](https://doi.org/10.7544/issn1000-1239.2021.20200564)]
  - [9] Zhang JY, Ma CX, Li C, Zhao ZY. Towards lightweight cross-domain sequential recommendation via tri-branches graph external attention network. Journal of Computer Research and Development, 2024, 61(8): 1930–1944 (in Chinese with English abstract). [doi: [10.7544/issn1000-1239.202440197](https://doi.org/10.7544/issn1000-1239.202440197)]
  - [10] Qian ZS, Wang YH, Yu QY, Fan FY, Fu TF. Multi-layer recurrent GCN cross-domain recommendation with pseudo-overlap detection mechanism. Ruan Jian Xue Bao/Journal of Software, 2025, 36(9): 4327–4348 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/7289.htm> [doi: [10.13328/j.cnki.jos.007289](https://doi.org/10.13328/j.cnki.jos.007289)]
  - [11] Wang XH, Yue HP, Guo L, Guo F, He C, Han XH. User identification network with contrastive clustering for shared-account recommendation. Information Processing & Management, 2025, 62(3): 104055. [doi: [10.1016/j.ipm.2024.104055](https://doi.org/10.1016/j.ipm.2024.104055)]
  - [12] Sun WC, Ma MY, Ren PJ, Lin YJ, Chen ZM, Ren ZC, Ma J, de Rijke M. Parallel split-join networks for shared account cross-domain sequential recommendations. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(4): 4106–4123. [doi: [10.1109/TKDE.2021.3130927](https://doi.org/10.1109/TKDE.2021.3130927)]
  - [13] Guo L, Zhang JY, Tang L, Chen T, Zhu L, Yin HZ. Time interval-enhanced graph neural network for shared-account cross-domain sequential recommendation. IEEE Trans. on Neural Networks and Learning Systems, 2024, 35(3): 4002–4016. [doi: [10.1109/TNNLS.2022.3201533](https://doi.org/10.1109/TNNLS.2022.3201533)]
  - [14] Zhang JY, Zhao ZY, Li C, Yu YW. Lightweight yet fine-grained: A graph capsule convolutional network with subspace alignment for shared-account sequential recommendation. In: Proc. of the 39th AAAI Conf. on Artificial Intelligence. Philadelphia: AAAI Press, 2025. 13242–13250. [doi: [10.1609/aaai.v39i12.33445](https://doi.org/10.1609/aaai.v39i12.33445)]
  - [15] Lu X, Zhu L, Cheng ZY, Nie LQ, Zhang HX. Online multi-modal hashing with dynamic query-adaption. In: Proc. of the 42nd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Paris: ACM, 2019. 715–724. [doi: [10.1145/3331184.3331217](https://doi.org/10.1145/3331184.3331217)]
  - [16] Yin HZ, Wang QY, Zheng K, Li ZX, Zhou XF. Overcoming data sparsity in group recommendation. IEEE Trans. on Knowledge and Data Engineering, 2022, 34(7): 3447–3460. [doi: [10.1109/TKDE.2020.3023787](https://doi.org/10.1109/TKDE.2020.3023787)]
  - [17] Yin HZ, Wang QY, Zheng K, Li ZX, Yang JL, Zhou XF. Social influence-based group representation learning for group recommendation. In: Proc. of the 35th IEEE Int'l Conf. on Data Engineering. Macao: IEEE, 2019. 566–577. [doi: [10.1109/ICDE.2019.00057](https://doi.org/10.1109/ICDE.2019.00057)]
  - [18] Guo L, Yin HZ, Wang QY, Cui B, Huang Z, Cui LZ. Group recommendation with latent voting mechanism. In: Proc. of the 36th IEEE Int'l Conf. on Data Engineering. Dallas: IEEE, 2020. 121–132. [doi: [10.1109/ICDE48307.2020.00018](https://doi.org/10.1109/ICDE48307.2020.00018)]
  - [19] Ma MY, Ren PJ, Lin YJ, Chen ZM, Ma J, de Rijke M.  $\pi$ -Net: A parallel information-sharing network for shared-account cross-domain sequential recommendations. In: Proc. of the 42nd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Paris: ACM, 2019. 685–694. [doi: [10.1145/3331184.3331200](https://doi.org/10.1145/3331184.3331200)]
  - [20] Guo L, Tang L, Chen T, Zhu L, Nguyen QVH, Yin HZ. DA-GCN: A domain-aware attentive graph convolution network for shared-account cross-domain sequential recommendation. In: Proc. of the 30th Int'l Joint Conf. on Artificial Intelligence. IJCAI, 2021. 2483–2489. [doi: [10.24963/ijcai.2021/342](https://doi.org/10.24963/ijcai.2021/342)]
  - [21] Jiang JY, Li CT, Chen YA, Wang W. Identifying users behind shared accounts in online streaming services. In: Proc. of the 41st Int'l ACM SIGIR Conf. on Research & Development in Information Retrieval. Ann Arbor: ACM, 2018. 65–74. [doi: [10.1145/3209978.3210054](https://doi.org/10.1145/3209978.3210054)]
  - [22] Wang ZJ, Yang Y, He L, Gu JZ. User identification within a shared account: Improving IP-TV recommender performance. In: Proc. of the 18th East European Conf. on Advances in Databases and Information Systems. Ohrid: Springer, 2014. 219–233. [doi: [10.1007/978-3-](https://doi.org/10.1007/978-3-)

- 319-10933-6\_17]
- [23] Zhao YF, Cao J, Tan YD. Passenger prediction in shared accounts for flight service recommendation. In: Proc. of the 10th Asia-Pacific Services Computing Conf. on Advances in Services Computing. Zhangjiajie: Springer, 2016. 159–172. [doi: [10.1007/978-3-319-49178-3\\_12](https://doi.org/10.1007/978-3-319-49178-3_12)]
  - [24] Zhang A, Fawaz N, Ioannidis S, Montanari A. Guess who rated this movie: Identifying users through subspace clustering. In: Proc. of the 28th Conf. on Uncertainty in Artificial Intelligence. Catalina: AUAI Press, 2012. 944–953.
  - [25] Wen XY, Peng ZH, Huang SH, Wang SZ, Yu PS. MISS: A multi-user identification network for shared-account session-aware recommendation. In: Proc. of the 26th Int'l Conf. on Database Systems for Advanced Applications. Taipei: Springer, 2021. 228–243. [doi: [10.1007/978-3-030-73200-4\\_15](https://doi.org/10.1007/978-3-030-73200-4_15)]
  - [26] Yang Y, Hu QM, He L, Ni MJ, Wang ZJ. Adaptive temporal model for IPTV recommendation. In: Proc. of the 16th Int'l Conf. on Web-age Information Management. Qingdao: Springer, 2015. 260–271. [doi: [10.1007/978-3-319-21042-1\\_21](https://doi.org/10.1007/978-3-319-21042-1_21)]
  - [27] Gu YX, Han X, Liu ZY, Huang ML. PPT: Pre-trained prompt tuning for few-shot learning. In: Proc. of the 60th Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers). Dublin: ACL, 2022. 8410–8423. [doi: [10.18653/v1/2022.acl-long.576](https://doi.org/10.18653/v1/2022.acl-long.576)]
  - [28] Cui ZY, Ma JX, Zhou C, Zhou JR, Yang HX. M6-Rec: Generative pretrained language models are open-ended recommender systems. arXiv:2205.08084, 2022.
  - [29] Geng SJ, Liu SC, Fu ZH, Ge YQ, Zhang YF. Recommendation as language processing (RLP): A unified pretrain, personalized prompt & predict paradigm (P5). In: Proc. of the 16th ACM Conf. on Recommender Systems. Seattle: ACM, 2022. 299–315. [doi: [10.1145/3523227.3546767](https://doi.org/10.1145/3523227.3546767)]
  - [30] Wu YQ, Xie RB, Zhu YC, Zhuang FZ, Zhang X, Lin LY, He Q. Personalized prompt for sequential recommendation. IEEE Trans. on Knowledge and Data Engineering, 2024, 36(7): 3376–3389. [doi: [10.1109/TKDE.2024.3357498](https://doi.org/10.1109/TKDE.2024.3357498)]
  - [31] Shani G, Heckerman D, Brafman RI. An MDP-based recommender system. Journal of Machine Learning Research, 2005, 6(43): 1265–1295.
  - [32] Chen S, Moore JL, Turnbull D, Joachims T. Playlist prediction via metric embedding. In: Proc. of the 18th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Beijing: ACM, 2012. 714–722. [doi: [10.1145/2339530.2339643](https://doi.org/10.1145/2339530.2339643)]
  - [33] Quadrana M, Karatzoglou A, Hidasi B, Cremonesi P. Personalizing session-based recommendations with hierarchical recurrent neural networks. In: Proc. of the 11th ACM Conf. on Recommender Systems. Como: ACM, 2017. 130–137. [doi: [10.1145/3109859.3109896](https://doi.org/10.1145/3109859.3109896)]
  - [34] He XN, He ZK, Song JK, Liu ZG, Jiang YG, Chua TS. NAIS: Neural attentive item similarity model for recommendation. IEEE Trans. on Knowledge and Data Engineering, 2018, 30(12): 2354–2366. [doi: [10.1109/TKDE.2018.2831682](https://doi.org/10.1109/TKDE.2018.2831682)]
  - [35] Xie T, Xu YJ, Chen L, Liu Y, Zheng ZB. Sequential recommendation on dynamic heterogeneous information network. In: Proc. of the 37th IEEE Int'l Conf. on Data Engineering. Chania: IEEE, 2021. 2105–2110. [doi: [10.1109/ICDE51399.2021.00208](https://doi.org/10.1109/ICDE51399.2021.00208)]
  - [36] Chen HX, Yin HZ, Chen T, Nguyen QVH, Peng WC, Li X. Exploiting centrality information with graph convolutions for network representation learning. In: Proc. of the 35th IEEE Int'l Conf. on Data Engineering. Macao: IEEE, 2019. 590–601. [doi: [10.1109/ICDE.2019.00059](https://doi.org/10.1109/ICDE.2019.00059)]
  - [37] Hsu C, Li CT. RetaGNN: Relational temporal attentive graph neural networks for holistic sequential recommendation. In: Proc. of the 2021 Web Conf. Ljubljana: ACM, 2021. 2968–2979. [doi: [10.1145/3442381.3449957](https://doi.org/10.1145/3442381.3449957)]
  - [38] Li XL, Liang P. Prefix-tuning: Optimizing continuous prompts for generation. In: Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int'l Joint Conf. on Natural Language Processing (Vol. 1: Long Papers). ACL, 2021. 4582–4597. [doi: [10.18653/v1/2021.acl-long.353](https://doi.org/10.18653/v1/2021.acl-long.353)]
  - [39] Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I. Language models are unsupervised multitask learners. OpenAI Blog, 2019, 1(8): 9.
  - [40] Liu X, Zheng YN, Du ZX, Ding M, Qian YJ, Yang ZL, Tang J. GPT understands, too. AI Open, 2024, 5: 208–215. [doi: [10.1016/j.aiopen.2023.08.012](https://doi.org/10.1016/j.aiopen.2023.08.012)]
  - [41] Liu X, Ji KX, Fu YC, Tam W, Du ZX, Yang ZL, Tang J. P-tuning: Prompt tuning can be comparable to fine-tuning across scales and tasks. In: Proc. of the 60th Annual Meeting of the Association for Computational Linguistics (Vol. 2: Short Papers). Dublin: ACL, 2022. 61–68. [doi: [10.18653/v1/2022.acl-short.8](https://doi.org/10.18653/v1/2022.acl-short.8)]
  - [42] Schick T, Schütze H. Exploiting cloze-questions for few-shot text classification and natural language inference. In: Proc. of the 16th Conf. of the European Chapter of the Association for Computational Linguistics: Main Volume. ACL, 2021. 255–269. [doi: [10.18653/v1/2021.eacl-main.20](https://doi.org/10.18653/v1/2021.eacl-main.20)]
  - [43] Brown TB, Mann B, Ryder N, *et al.* Language models are few-shot learners. In: Proc. of the 34th Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2020. 159.

- [44] Gao TY, Fisch A, Chen DQ. Making pre-trained language models better few-shot learners. In: Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int'l Joint Conf. on Natural Language Processing (Vol. 1: Long Papers). ACL, 2021. 3816–3830. [doi: [10.18653/v1/2021.acl-long.295](https://doi.org/10.18653/v1/2021.acl-long.295)]
- [45] Jiang ZB, Xu FF, Araki J, Neubig G. How can we know what language models know? Trans. of the Association for Computational Linguistics, 2020, 8: 423–438. [doi: [10.1162/tac1\\_a\\_00324](https://doi.org/10.1162/tac1_a_00324)]
- [46] Wang X, He XN, Wang M, Feng FL, Chua TS. Neural graph collaborative filtering. In: Proc. of the 42nd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Paris: ACM, 2019. 165–174. [doi: [10.1145/3331184.3331267](https://doi.org/10.1145/3331184.3331267)]
- [47] Duan WY, Liang DC. User disambiguation learning for precise shared-account marketing: A hierarchical self-attentive sequential recommendation method. Knowledge-based Systems, 2025, 315: 113328. [doi: [10.1016/j.knosys.2025.113328](https://doi.org/10.1016/j.knosys.2025.113328)]
- [48] He XN, Liao LZ, Zhang HW, Nie LQ, Hu X, Chua TS. Neural collaborative filtering. In: Proc. of the 26th Int'l Conf. on World Wide Web. Perth: Int'l World Wide Web Conferences Steering Committee, 2017. 173–182. [doi: [10.1145/3038912.3052569](https://doi.org/10.1145/3038912.3052569)]
- [49] He XN, Deng K, Wang X, Li Y, Zhang YD, Wang M. LightGCN: Simplifying and powering graph convolution network for recommendation. In: Proc. of the 43rd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2020. 639–648. [doi: [10.1145/3397271.3401063](https://doi.org/10.1145/3397271.3401063)]
- [50] Zhu Y, Li H, Liao YK, Wang BD, Guan ZY, Liu HF, Cai D. What to do next: Modeling user behaviors by time-LSTM. In: Proc. of the 26th Int'l Joint Conf. on Artificial Intelligence. Melbourne: IJCAI, 2017. 3602–3608. [doi: [10.24963/ijcai.2017/504](https://doi.org/10.24963/ijcai.2017/504)]
- [51] Fan ZW, Liu ZW, Zhang JW, Xiong Y, Zheng L, Yu PS. Continuous-time sequential recommendation with temporal graph collaborative Transformer. In: Proc. of the 30th ACM Int'l Conf. on Information & Knowledge Management. Queensland: ACM, 2021. 433–442. [doi: [10.1145/3459637.3482242](https://doi.org/10.1145/3459637.3482242)]
- [52] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: Proc. of the 13th Int'l Conf. on Artificial Intelligence and Statistics. Sardinia: PMLR, 2010. 249–256.
- [53] Rendle S, Freudenthaler C, Gantner Z, Schmidt-Thieme L. BPR: Bayesian personalized ranking from implicit feedback. In: Proc. of the 25th Conf. on Uncertainty in Artificial Intelligence. Montreal: AUAI Press, 2009. 452–461
- [54] Linden G, Smith B, York J. Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing, 2003, 7(1): 76–80.

## 附中文参考文献

- [5] 钱忠胜, 叶祖铎, 姚昌森, 张丁, 黄恒, 秦朗悦. 融合自适应周期与兴趣量因子的轻量级 GCN 推荐. 软件学报, 2024, 35(6): 2974–2998. <http://www.jos.org.cn/1000-9825/6915.htm> [doi: [10.13328/j.cnki.jos.006915](https://doi.org/10.13328/j.cnki.jos.006915)]
- [8] 郭磊, 李秋菊, 刘方爱, 王新华. 基于自注意力网络的共享账户跨域序列推荐. 计算机研究与发展, 2021, 58(11): 2524–2537. [doi: [10.7544/issn1000-1239.2021.20200564](https://doi.org/10.7544/issn1000-1239.2021.20200564)]
- [9] 张劲羽, 马晨曦, 李超, 赵中英. 基于三支图外部注意力网络的轻量化跨域序列推荐. 计算机研究与发展, 2024, 61(8): 1930–1944. [doi: [10.7544/issn1000-1239.202440197](https://doi.org/10.7544/issn1000-1239.202440197)]
- [10] 钱忠胜, 王亚惠, 俞情媛, 范赋宇, 付庭峰. 利用伪重叠判定机制的多层循环 GCN 跨域推荐. 软件学报, 2025, 36(9): 4327–4348. <http://www.jos.org.cn/1000-9825/7289.htm> [doi: [10.13328/j.cnki.jos.007289](https://doi.org/10.13328/j.cnki.jos.007289)]

## 作者简介

赵中英, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究领域为图数据挖掘, 推荐系统.

王莹, 硕士生, 主要研究领域为图神经网络, 推荐系统.

张劲羽, 博士生, CCF 学生会员, 主要研究领域为序列推荐, 轻量推荐.

周慧, 博士, 副教授, CCF 专业会员, 主要研究领域为图数据挖掘, 推荐系统.

李超, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为图神经网络, 推荐系统.

曾庆田, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为图数据挖掘, 推荐系统, 流程挖掘.