

# 基于图神经网络的用户关联行为分析研究

林云婷<sup>1</sup> 唐乐红<sup>\*1</sup> 黄长泉<sup>2</sup> 阮煜鑫<sup>1</sup> 徐旭晖<sup>1</sup>

1. 阳光学院, 福建 福州 350015

2. 中科数创(厦门)智能科技研究院, 福建 厦门 361015

**[摘要]**针对传统用户行为分析方法难以刻画用户之间高阶关联、时序依赖与关系传播的问题,提出一种基于图神经网络的用户关联行为分析方法。首先,将用户属性、浏览点击、收藏购买及交互时间统一表示为多关系图,分别构建用户—行为二部图、用户关联图和时序行为图;其次,引入关系消息传播、时间衰减权重与注意力聚合机制,学习用户节点的结构化表示;最后,通过关联解码器计算用户间潜在关联概率,并支持群体发现、异常关系识别和个性化运营。在包含640个用户、2932条正关联边的可控仿真数据上进行验证,结果表明,所提方法的F1值和AUC分别达到0.755和0.814,较多层感知机分别提高3.5和5.8个百分点。

**[关键词]**图神经网络; 用户行为分析; 关联预测; 时序注意力; 图表示学习

DOI: 10.64635/ja.2026.1420

中图分类号: TP183

文献标识码: A

## Research on User Correlation Behavior Analysis Based on Graph Neural Networks

Lin Yunting<sup>1</sup>, Tang Lehong<sup>\*1</sup>, Huang Changquan<sup>2</sup>, Ruan Yuxin<sup>1</sup>, Xu Xuhui<sup>1</sup>

1. Yango University, Fuzhou, Fujian 350015, China

2. Zhongke Shuchuang (Xiamen) Intelligent Technology Research Institute, Xiamen, Fujian 361015, China

**Abstract:** To address the difficulty of traditional user behavior analysis methods in characterizing high-order correlations, temporal dependencies, and relationship propagation among users, this paper proposes a user correlation behavior analysis method based on graph neural networks. First, user attributes, browsing and clicking, favoriting and purchasing, and interaction time are uniformly represented as a multi-relational graph, and a user-behavior bipartite graph, a user correlation graph, and a temporal behavior graph are constructed respectively. Second, relation-based message propagation, time-decay weights, and attention aggregation mechanisms are introduced to learn structured representations of user nodes. Finally, a correlation decoder is used to calculate the potential association probability between users and support group discovery, abnormal relationship identification, and personalized operations. The method is validated on controllable simulation data containing 640 users and 2,932 positive association edges. The results show that the proposed method achieves F1 and AUC values of 0.755 and 0.814, respectively, representing improvements of 3.5 and 5.8 percentage points over the multilayer perceptron model.

**Keywords:** graph neural network; user behavior analysis; association prediction; temporal attention; graph representation learning

### 引言

互联网平台持续积累浏览、点击、搜索、收藏、评论和交易等行为数据。此类数据并非彼此独立:用户会因共同兴趣、社交关系、相似消费路径或同一事件形成隐性联系。传统聚类、协同过滤和序列模型通常将用户表示为相互独立的向量,能够分析个体偏好,却难以描述“用户—用户—行为对象”之间的高阶传播关系。图神经网络通过邻域消息聚合学习非欧式数据表示,已应用于社交网络、

电子商务和推荐系统等关系密集型场景<sup>[1]</sup>。

现有研究仍存在三方面不足:一是部分方法仅采用用户—项目交互图,忽略用户之间的显式或弱关系;二是静态图模型难以反映兴趣随时间变化的过程;三是深层传播可能引入噪声并产生过平滑,使不同用户的表示趋同。本文提出关联行为图神经网络(Relation-aware Behavior Graph Neural Network, RBGNN),将多源行为、关系结构与时间信息统一建模,重点解决多关系融合、时序权重

分配及关联预测问题。

## 1 相关研究

图神经网络推荐研究主要围绕用户-项目交互、会话序列、社交网络和知识图谱展开。融合社交信息的多图会话推荐方法表明,用户短期兴趣与好友影响需要联合建模,图注意机制能够动态区分不同关系的影响程度<sup>[2]</sup>。语义增强方法进一步将文本语义嵌入和邻居注意权重引入图传播过程,以改善实体特征质量<sup>[3]</sup>。这些研究为用户关联分析提供了基础,但其目标多为下一项目推荐,关联边本身常被视为辅助信息,而非直接预测对象。本文将用户间关联概率作为核心输出,并同时保留行为关系、时序关系和已知关联关系三个通道。

## 2 问题定义与图构建

设用户集合为  $U=\{u_1, u_2, \dots, u_n\}$ , 行为对象集合为  $I$ , 行为类型集合为  $R$ 。对任意交互记录  $e=(u, i, r, t)$ , 分别表示用户、对象、行为类型与时间戳。本文构建多关系图  $G=(V, E, X)$ , 其中  $V$  包含用户节点和行为对象节点,  $E$  由用户-对象交互边、用户关联边及时序邻接边组成,  $X$  为节点属性矩阵。

用户节点初始特征由三部分组成: ①各类行为频次及占比; ②按时间窗口统计的活跃分布; ③用户属性经脱敏编码后的低维向量。对交互边设置行为强度系数  $\omega_r$ , 并采用指数衰减函数  $d(\Delta t)=\exp(-\lambda \Delta t)$  降低历史行为权重。两个用户若在相近时间窗口对相同或相似对象产生交互, 则建立弱关联边, 其权重为行为相似度、时间接近度与共同对象比例的加权和。由此得到行为图  $A_b$ 、关联图  $A_r$  和时序图  $A_t$ 。

## 3 RBGNN 模型设计

### 3.1 多关系消息传播

对第  $l$  层用户节点  $u$ , 不同关系通道分别聚合邻域信息:  $h_{ru}(l+1)=\sigma(\sum_v \in N_r(u) \alpha_{ruv} W_r h_v(l))$ 。其中,  $N_r(u)$  表示关系  $r$  下的邻居集合,  $W_r$  为关系变换矩阵,  $\alpha_{ruv}$  由节点特征、边类型和时间衰减共同计算。为避免所有邻居等权传播, 引入注意力机制突出高相关邻居; 为缓解深层网络过平滑, 在两层传播后加入残差连接和层平均<sup>[4]</sup>。类似研究已证明, 跳跃连接与残差注意能够提升多层图注意网络在会话推荐中的稳定性。

### 3.2 时序注意与关系融合

用户兴趣具有阶段性。本文将历史记录划分为近期、中期和长期窗口, 分别生成时序表示  $h^t$ 。通过门控系数  $g=\text{softmax}(W_g[h_p||h_r||h^t])$  自适应融合三个通道, 得到最终用

户表示  $z_u=g_1h_p+g_2h_r+g_3h^t$ 。该机制既利用已知关系的拓扑传播, 也避免模型完全依赖显式社交边。针对稀疏用户, 模型通过结构邻居和语义邻居补充监督信号; 解耦结构邻居与语义邻居的图对比学习研究表明, 该思路有助于识别用户行为中的潜在意图<sup>[5]</sup>。

## 3.3 关联预测与优化目标

对用户对  $(u, v)$ , 关联解码器输入  $|z_u-z_v|$ 、 $z_u \odot z_v$  及共同邻域统计量, 输出关联概率  $p_{uv}=\text{Sigmoid}(\text{MLP}(\cdot))$ 。训练目标由二元交叉熵、嵌入正则项和关系一致性约束组成。正样本为已确认关联边, 负样本从未连接用户对中分层采样, 以降低热门用户带来的样本偏置。模型流程见图 1。



图 1 基于图神经网络的用户关联行为分析模型框架

## 4 实验与结果分析

### 4.1 实验设置

为避免将未授权平台数据作为论文证据, 本文构建可控仿真数据集: 包含 640 个用户、8 类行为兴趣、6 个时间窗口以及 2932 条正关联边, 并按 1:1 采样负关联边。数据按 60%、15% 和 25% 划分为训练集、验证集与测试集。对比方法包括逻辑回归 (LR)、多层感知机 (MLP)、基于行为相似图的标准 GCN 及本文 RBGNN。评价指标采用 Accuracy、F1 和 AUC, 分类阈值在验证集上确定。

### 4.2 结果分析

表 1 显示, RBGNN 在 F1 和 AUC 上均取得最优结果。与表现较好的非图模型 MLP 相比, F1 提高 3.5 个百分点, AUC 提高 5.8 个百分点。标准 GCN 仅在单一行为相似图上传播, 容易把偶然共同行为视为稳定关系; RBGNN 加入关联拓扑和时间权重后, 能够区分长期稳定联系与短期同步行为。Accuracy 提升幅度相对有限, 原因在于验证集阈值以 F1 最优为目标, 使模型更重视正关联用户的召回。

消融观察表明, 删除时间通道后, 近期兴趣突变用户更易被错误归入历史群体; 删除关联图后, 模型对低活跃用户的识别能力下降。说明行为内容与关系结构具有互补性。

表 1 不同方法的用户关联预测结果

| 方法         | Accuracy | F1    | AUC   |
|------------|----------|-------|-------|
| LR         | 0.589    | 0.684 | 0.697 |
| MLP        | 0.688    | 0.720 | 0.756 |
| GCN        | 0.647    | 0.708 | 0.744 |
| RBGNN (本文) | 0.731    | 0.755 | 0.814 |

注：分类阈值依据验证集 F1 值确定。

5 应用与安全讨论

模型可用于三类业务：第一，发现兴趣相近或传播路径相似的用户群体，为内容推荐和用户运营提供依据；第二，识别短时间内高度同步且关系异常密集的账户，为刷量、团伙注册等风险排查提供线索；第三，分析关键节点及其影响范围，为社群维护和信息扩散评估提供支持。

用户关联分析具有较高隐私敏感性，图结构可能泄露用户身份、社交圈和行为偏好。实际部署应坚持数据最小化、标识符匿名化、访问分级和结果用途限定。已有研究采用负数据库发布扰动图，在维持图分析能力的同时降低原始结构暴露风险。

6 结论

本文提出基于图神经网络的用户关联行为分析方法

RBGNN，通过多源行为构图、关系消息传播、时序注意融合和关联解码，实现用户间潜在关系识别。仿真结果表明，该方法在 F1 和 AUC 指标上优于对比模型，说明图结构能够补充独立向量模型难以利用的高阶关联。后续研究可引入真实公开数据、动态图增量训练和自监督学习，以增强模型在匿名会话、稀疏交互及兴趣快速迁移场景下的适应性。多视图图结构与自监督学习已被用于提取会话高阶特征并缓解数据稀疏问题，可作为模型进一步优化的方向。

[参考文献]

[1]李甜甜,张荣梅,张佳惠.图神经网络技术研究综述[J].河北省科学院学报,2022,39(2):1-13.

[2]雷景生,李冉,杨胜英,等.融合社交信息的多图神经网络会话推荐方法[J].计算机工程与应用,2023,59(15):264-273.

[3]闫阳,王雷全,李家瑞.融合语义信息与注意力的图神经网络推荐算法[J].计算机系统应用,2023,32(4):214-222.

[4]丁美荣,王雨航,曾碧卿.结合跳跃连接的多层图注意力网络会话推荐[J].计算机系统应用,2024,33(2):23-32.

[5]杨红伟,曹家晟,刘学军,等.融合结构邻居和语义邻居的解耦图对比学习推荐模型[J].计算机系统应用,2024,33(7):149-160.