



基于深度学习的群组推荐方法研究综述

郑楠 章颂 刘玉桥 王雨桐 王飞跃

A Comprehensive Review of Group Recommendation Methods Based on Deep Learning

ZHENG Nan, ZHANG Song, LIU Yu-Qiao, WANG Yu-Tong, WANG Fei-Yue

在线阅读 View online: <https://doi.org/10.16383/j.aas.c230781>

您可能感兴趣的其他文章

深度对比学习综述

Deep Contrastive Learning: A Survey

自动化学报. 2023, 49(1): 15–39 <https://doi.org/10.16383/j.aas.c220421>

基于深度学习的表面缺陷检测方法综述

A Survey of Surface Defect Detection Methods Based on Deep Learning

自动化学报. 2021, 47(5): 1017–1034 <https://doi.org/10.16383/j.aas.c190811>

融合深度学习的贝叶斯滤波综述

A Survey on Bayesian Filtering With Deep Learning

自动化学报. 2024, 50(8): 1502–1516 <https://doi.org/10.16383/j.aas.c230457>

多Agent深度强化学习综述

Deep Multi-Agent Reinforcement Learning: A Survey

自动化学报. 2020, 46(12): 2537–2557 <https://doi.org/10.16383/j.aas.c180372>

基于深度学习的单幅图像超分辨率重建算法综述

A Review of Single Image Super-resolution Reconstruction Algorithms Based on Deep Learning

自动化学报. 2021, 47(10): 2341–2363 <https://doi.org/10.16383/j.aas.c190859>

解耦表征学习综述

A Review of Disentangled Representation Learning

自动化学报. 2022, 48(2): 351–374 <https://doi.org/10.16383/j.aas.c210096>

基于深度学习的群组推荐方法研究综述

郑楠^{1,2} 章颂^{1,2} 刘玉桥^{1,2} 王雨桐³ 王飞跃^{2,3}

摘要 群组推荐在信息检索与数据挖掘领域近年来备受关注, 旨在从海量候选集中挑选出一组用户可能感兴趣的项目. 随着深度学习技术的不断发展, 基于深度学习的群组推荐方法大量涌现. 鉴于此, 首先介绍群组推荐问题的背景知识, 然后系统综述数据获取方法, 全面评述近年来基于深度学习的群组推荐算法, 并进行系统分类与深入分析. 此外, 还归纳了适用于深度学习方法的群组推荐数据集和评价方法, 对各类推荐算法进行对比实验分析与讨论. 最后, 针对本领域的研究难点进行深入探讨, 并提出未来有待深入研究的方向.

关键词 群组推荐, 推荐系统综述, 深度学习, 群组表示学习

引用格式 郑楠, 章颂, 刘玉桥, 王雨桐, 王飞跃. 基于深度学习的群组推荐方法研究综述. 自动化学报, 2024, 50(12): 2301-2324

DOI 10.16383/j.aas.c230781 **CSTR** 32138.14.j.aas.c230781

A Comprehensive Review of Group Recommendation Methods Based on Deep Learning

ZHENG Nan^{1,2} ZHANG Song^{1,2} LIU Yu-Qiao^{1,2} WANG Yu-Tong³ WANG Fei-Yue^{2,3}

Abstract Group recommendation has emerged as a highly active research topic in the fields of information retrieval and data mining in recent years. Its objective is to select a group of items from a large candidate set that is likely to be of interest to a set of users. With the advancement of deep learning, numerous group recommendation methods based on deep learning have been proposed. This paper provides a brief introduction to the background knowledge of this problem. It reviews the methods of data acquisition and conducts a comprehensive review, systematic classification, and in-depth analysis of group recommendation algorithms based on deep learning. In addition, this paper outlines some group recommendation datasets and evaluation methods suitable for deep methods, and conducts comparative experimental analysis and discussion on various recommendation algorithms. Finally, the research challenges in this field were analyzed, and valuable future research directions were discussed.

Key words Group recommendation, recommender system overview, deep learning, group representation learning

Citation Zheng Nan, Zhang Song, Liu Yu-Qiao, Wang Yu-Tong, Wang Fei-Yue. A comprehensive review of group recommendation methods based on deep learning. *Acta Automatica Sinica*, 2024, 50(12): 2301-2324

随着 Web2.0 时代的兴起, 互联网用户的上网自由度得到显著提升, 其可随时随地以多种接入形式传播信息. 在此背景下, 越来越多的人群开始积极参与群组活动. 群组是由具有相同特征或共同兴

趣的个体组成的社群或社区 (Group/Community). 其成员涵盖现实生活中的亲友关系, 也包括因共同兴趣而聚集的陌生人. 在群组中, 成员之间可以保持联系, 分享各类资讯、产品, 以及交流思想. 例如, 拼多多的社交拼团业务以“社交 + 电商”的模式为用户提供共享式购物体验; 豆瓣小组致力于将志趣相投的个体聚集在一起, 共同讨论感兴趣的话题并分享经验; Meetup 平台上, 用户可与兴趣相投的其他用户共同参与线上或线下活动等.

截至 2022 年 11 月, 美团 APP 的月活跃用户 (Monthly active user, MAU) 已达到 1.7 亿人, 而在 2021 年 12 月至 2022 年 11 月的 12 个月间, 该应用的平均 MAU 稳定维持在 1.6 亿人. 然而, 随着大量用户的参与和网络资源的增加^[1], 信息过载 (Information overload) 问题日益凸显^[2]. 在此背景下, 如何有效帮助群组在海量信息空间中发现和获取知识成为亟待解决的问题. 对于群组而言, 降低

收稿日期 2023-12-18 录用日期 2024-05-12

Manuscript received December 18, 2023; accepted May 12, 2024

国家重点研发计划 (2023YFC3304104), 国家自然科学基金 (U1811463) 资助

Supported by National Key Research and Development Program of China (2023YFC3304104) and National Natural Science Foundation of China (U1811463)

本文责任编辑 张敏灵

Recommended by Associate Editor ZHANG Min-Ling

1. 中国科学院自动化研究所多模态人工智能系统全国重点实验室 北京 100190 2. 中国科学院大学人工智能学院 北京 100049 3. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190

1. State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190 2. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049 3. State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190

获取信息的成本可以提升用户体验;对于网站而言,通过提供群组信息服务,吸引用户参与,可增强用户粘性;在商业领域,对群组行为进行收集和分析可帮助企业更好地理解群组需求,定位潜在客户群,制定市场策略,从而提高市场占有率;而在公共管理领域,群组偏好反映了社会某些问题,通过监测热点话题、跟踪群组的偏好转变,可以辅助决策者采取有效措施,正确引导舆论.推荐系统作为提供信息筛选与服务的重要途径之一,可以与多种互联网应用相结合,为国家和民众的各个领域提供服务.

群组推荐(Group recommendation)旨在从庞大的候选集中挑选出一组可能引起群组用户兴趣的项目,并将其推送给群组成员.相较于个性化推荐^[3-4],群组推荐系统更为复杂,因为其除了考虑到用户个性化偏好外,还需考虑群组内成员之间复杂的交互关系^[5].因此,如何挖掘群组共性,实现有效的群组推荐,成为学术界和工业界日益关注的焦点.早在1998年,美国学者McCarthy等^[6]在世界计算机支持的协同工作会议上首次提出一种用于音乐推荐的群组推荐系统MusicFX,该工作成为该领域的里程碑.随着社交网络和移动设备的迅速发展,人们越来越频繁地参与群体活动,因此对群组推荐系统的需求也日益增长.

深度学习的迅速发展,尤其是注意力神经网络的引入,对深度群组推荐的发展起到重要推动作用.深度群组推荐方法基于历史交互信息,利用神经网络强大的建模能力学习更加鲁棒的群组偏好表示,从而实现更为准确的群组推荐.本文所探讨的研究聚焦于对2016年—2023年间利用深度学习技术进行群组推荐算法的系统综述.相较于既有文献^[7-11],本文的独特之处在于从方法论的角度对算法进行分类梳理.现有文献主要着眼于偏好融合策略或从不同应用的角度审视群组推荐方法,而本文将深度学习作为研究焦点,包括启发式聚合策略、概率模型、注意力机制、图神经网络以及预训练模型等.本文提出的文献分类框架能够为深度群组推荐领域的进展提供更深入的方法学探讨.

本文第1节对群组推荐的类型及其交互关系进行介绍,并提供群组推荐的形式化描述,同时对涉及到的相关术语进行详细说明.第2节详细地阐述基于深度学习的群组推荐系统的基本流程,包括数据输入层、表示学习层和模型层.特别是,在表示学习层,按照不同的群组类型和模型技术对方法进行系统分类.第3节对群组推荐的数据集和评价指标进行系统梳理,并对各类推荐算法进行对比实验分

析与讨论.在第4节中,分析现有方法在数据、可解释性和公平性等方面所面临的挑战.最后,在第5节中,结合群决策理论和大语言模型等技术,展望群组推荐未来可能的研究方向和发展趋势.

1 背景知识

在群组推荐领域,系统地收集群组相关数据并运用算法和模型进行处理是实现有效推荐的关键步骤,图1展示了群组推荐过程的示意图.系统首先收集群组相关的数据,包括交互关系、描述信息(如成员偏好、年龄、性别等,项目属性、类别等),以及其他辅助信息(如标签、好友关系等).接下来,通过算法和模型对这些数据进行分析和处理,例如表示学习、相似度计算等.最后,系统根据项目的评分,生成最终的群组推荐结果,并将其呈现给用户.

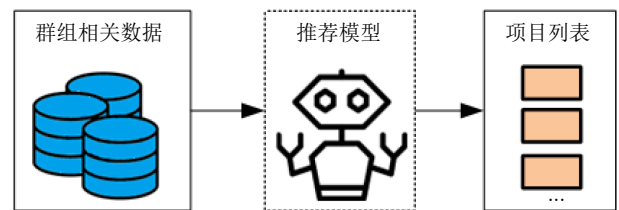


图1 群组推荐示意图

Fig.1 Diagram of group recommendation

1.1 群组的类型

O'Connor等^[12]根据群组成员的固定性将群组划分为持续性群组和临时性群组.持续性群组指的是长期存在的群组,其成员通常固定,存在一定数量的群组与项目的历史交互记录.例如,数据集CAMRa2011中的家庭组和豆瓣小组等属于这一类别.而临时性群组则是因一次活动而临时组建的用户群体,活动结束后即解散,因此群组成员不固定.由于临时性群组缺乏历史交互记录,因此也被视为冷启动群组.临时性群组通常只有极少甚至几乎没有历史交互记录,因此无法直接对群组偏好进行建模.例如,Meetup活动组和Weeplaces/Yelp中基于位置信息的活动组都属于临时性群组.

1.2 交互关系概述

图2描述典型的群组交互过程.用户可以加入一个或多个群组,其加入群组的行为构成〈群组-用户〉交互关系.用户通过浏览、购买、收藏、评论、参与等行为与项目产生交互,这类行为构成〈用户-项目〉交互关系.以群组为单位,群组内所有成员共同参与某类活动或购买某类产品,形成〈群组-项

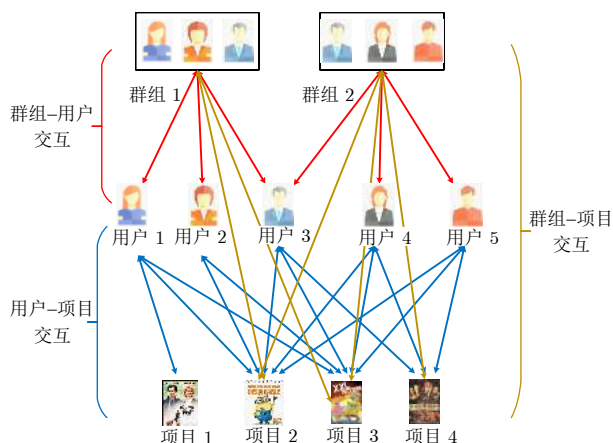


图 2 群组交互关系图

Fig.2 Diagram of group interaction

目)交互关系.可以看出,群组交互过程中主要包含三类交互关系,分别是〈群组-用户〉、〈用户-项目〉和〈群组-项目〉交互关系.如何利用和挖掘这些丰富的交互关系,实现更准确、更多样化的群组推荐是业界和学术界研究的热点.

1.3 群组推荐的形式化描述

本文中所涉及的数学符号如表 1 所示.

表 1 数学符号说明

Table 1 Explanation of mathematical symbols

符号	说明
$U = \{u_1, u_2, \dots, u_n\}$	用户集合
$V = \{v_1, v_2, \dots, v_m\}$	项目集合
$G = \{g_1, g_2, \dots, g_s\}$	群组集合
$A = [a_{ij}]_{s \times m}$	〈群组-项目〉交互矩阵
$B = [b_{ij}]_{n \times m}$	〈用户-项目〉交互矩阵
$C = [c_{ij}]_{s \times n}$	〈群组-用户〉交互矩阵
\mathcal{G}_{UV}	〈用户-项目〉交互图
$R_V(u)$	与用户 u 有过交互的所有项目集合
p	元路径
$\mathcal{N}^p(u)$	节点 u 基于元路径 p 找到的近邻集合
u_t^g	群组 g 中的第 t 个成员, $u_t^g \in U$
$f: V \rightarrow \mathbf{R}$	由项目集到实数域的函数(映射) f
e_u, e_v, e_g	用户 u , 项目 v 和群组 g 的 ID 嵌入向量
h_u, h_v, h_g	e_u, e_v, e_g 经过编码后的向量表示

群组推荐任务可形式化定义如下:

输入: 给定用户集合 U 、群组集合 G 和项目集合 V , 〈群组-项目〉交互矩阵 A 、〈用户-项目〉交互矩阵 B 和 〈群组-用户〉交互矩阵 C , 以及其他描述或辅助信息.

输出: 评分函数 $f: V \rightarrow \mathbf{R}$.

1.4 相关术语

本节对文中涉及的与群组推荐相关的概念术语进行说明.

1) 用户: 用户是指推荐系统中的个体对象, 通常具有属性描述信息, 例如年龄、性别、职业、爱好等. 一些系统会收集用户的历史行为和社交关系等数据, 这些数据和用户的属性信息构成了描述用户画像的重要参考.

2) 群组: 群组是由多个用户组成的组织, 是群组推荐系统的服务对象. 加入群组的用户成为群组的成员. 群组的形成可以是基于家庭或朋友关系形成的好友群组, 也可以是基于共同兴趣爱好形成的主题群组.

3) 项目: 本文将被推荐的对象统称为项目. 这些对象可以是实际的商品, 如衣服、玩具、电子产品等; 也可以是非实物产品, 如新闻、电影、音乐等; 或者是为用户提供某种服务或参与某类活动. 一些系统对项目有描述信息, 例如产品描述、商家信息、适合年龄等, 这些描述信息构成了挖掘项目特征的重要参考.

2 基于深度学习的群组推荐方法

深度学习利用深层次非线性网络结构, 将低层的特征组合起来形成更加抽象的高层表示属性或特征, 以发现数据的潜在特征表示. 目前, 深度学习已广泛应用于推荐系统^[13-14]. 相较于传统的推荐模式, 深度学习能够更好地捕获用户与项目之间的非线性复杂关系, 从而更好地获得用户和项目的深层次特征表示, 并揭示用户与项目之间的高阶交互关系. 本文主要介绍利用深度学习技术进行群组推荐的相关研究工作.

基于深度学习的群组推荐系统框架如图 3 所示, 主要包括数据输入层、表示学习层、模型层和输出层. 数据输入层涵盖了交互信息、描述信息、辅助信息等; 表示学习层根据输入层的数据, 学习用户的个体偏好表示、群组的群组偏好表示, 以及项目的特征表示等; 模型层的主流方法包括基于内积的方法和神经协同过滤 (Neural collaborative filtering, NCF) 模型等; 输出层则给出推荐列表.

2.1 数据输入层

数据输入层包括交互信息、描述信息、辅助信息等. 其中, 交互信息指用户或群组对项目的显式或隐式反馈数据; 描述信息包括用户的年龄、性别、职业、兴趣描述等用户画像信息, 以及项目的类型、

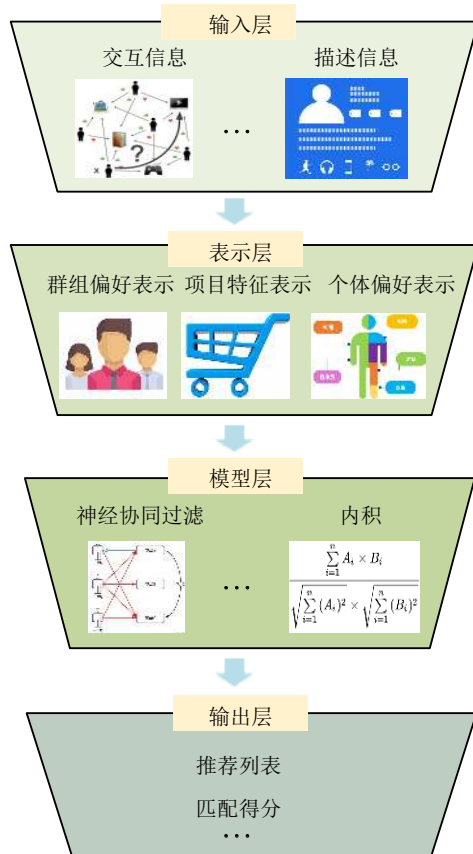


图 3 基于深度学习的群组推荐系统框架

Fig.3 Framework of a deep learning based group recommendation system

类别、内容描述等属性信息；而辅助信息是指可以用于更好地建模用户或项目的附加信息，例如标签、评论、社交关系等。交互信息作为实现群组推荐的基础，根据用户的参与方式，可以大致分为显式获取和隐式获取两类。

2.1.1 显式获取

早期的研究主要基于成员个体的显式评分数据，利用简单的平均或加权求和等方式获得群组评分数据，进而使用个性化推荐方法来生成推荐项目。例如，美国学者 McCarthy 等^[6]根据用户对音乐不同流派的评分数据，将用户的评分数据加和以获取群组的评分数据，并以此为基础为群组推荐成员共同感兴趣的音乐流派的歌曲。另外，美国西北大学的 Crossen 等^[15]提出了音乐推荐系统 Flytrap，该系统要求用户通过投票打分的形式显式表达自己的偏好，例如，喜欢哪个歌手、哪个流派的音乐等。系统将所有成员的打分信息加和并进行归一化，以得到所有音乐的推荐概率，最后按照此概率进行音乐推荐。上述方法是通过用户的显式评分来获取个体偏好的。尽管这种方式得到的信息能够更准确地反

映用户的偏好，但由于该方式需要用户付出额外的精力来维护自己的偏好信息，容易降低用户使用的积极性。研究表明，即使用户知道主动表达自身的偏好信息会提升体验感，但仍不愿意花费额外的精力提供自身的兴趣偏好^[16]。因而，通过显式获取方式获取用户的偏好信息在实践中较为困难，在一定程度上限制了群组推荐系统的发展。

2.1.2 隐式获取

为了更全面地获取用户的偏好信息，学者们开始研究隐式的兴趣获取方法。隐式获取方法通过分析用户长期的交互历史或检索日志来自动挖掘用户的兴趣。例如，Zheng 等^[17]基于用户参与 Flickr 群组的交互记录，利用用户参与的群组、用户给项目标注的标签，以及群组内项目的标签等信息，抽取〈用户-标签-群组〉三元交互关系，并统计标签的频率构建用户的兴趣模型。在共同主题空间的基础上，计算用户与群组之间的潜在关联，从而实现群组推荐。Liu 等^[18]提出了个人影响主题 (Personal impact topic, PIT) 模型，通过分析群组决策过程，考虑一群用户参与活动的行为和用户的打卡行为等因素，抽取〈群组-项目〉和〈用户-项目〉交互，并基于群体成员的个体偏好和个人的影响力来确定特定群体的偏好概况，以实现群组推荐。此外，Yuan 等^[19]提出了群组推荐的共识模型 (Consensus model, COM)，该模型同样考虑了一群用户参与活动的行为和用户的打卡行为等因素，基于〈群组-项目〉和〈用户-项目〉交互，构建群组偏好主题和用户偏好影响，并基于群组决策过程同时受到群组偏好主题和成员个人喜好影响的假设，计算群组与项目的潜在关联，提升群组推荐的质量。隐式偏好获取方法的优势在于：一方面，交互记录数据量大且易于获取；另一方面，这些交互记录是用户在与系统主动交互的过程中产生的，因此无需用户额外的精力维护，从而在获取用户偏好属性的同时提高用户体验。目前的研究也越来越侧重于使用该方法挖掘用户或群组的偏好。

2.2 表示学习层

表示学习，又称嵌入方法，是一种将对象向量化的技术，以供模型输入使用^[20-21]。在群组推荐任务中，表示学习层接收数据输入层的数据，将特征嵌入到低维向量中，并结合具体任务学习优化该向量，从而构建群组、用户和项目的向量化表示。其数学原理是通过最优化的方式使得存在交互关系的对象在向量空间中距离更接近，为模型层的推荐计算提供特征表示。基于深度学习的群组推荐任务的关键

在于如何建模群组偏好。根据群组的不同类型, 可以将群组偏好表示方法分为面向持续性群组的表示学习和面向临时性群组的表示学习方法两类。

持续性群组具有丰富的历史交互信息, 这为实现群组的偏好表示学习提供了宝贵的数据基础。传统方法通常采用基于深度学习的个性化推荐和简单的聚合策略来实现群组的偏好表示学习, 然而, 这些方法较为简单, 未能充分考虑成员对群组选择项目的影响力。随着机器学习的发展, 研究者开始探索将概率模型用于建模群组的生成过程, 以潜在变量表示用户对群组或其他成员的影响力。尽管这些方法在一定程度上依赖于先验分布的假设, 但它们为更精细的群组偏好表示学习提供了新的思路。随着深度学习技术的不断进步, 注意力机制被引入到交互特征的建模中, 以对用户、群组和项目的表示进行更精细的处理。通过动态调整注意力权重, 注意力机制能够有效地捕捉到用户在不同上下文中的偏好, 从而提升群组偏好表示的表达能力。为了进一步建模用户、群组和项目之间的高阶交互关系, 图神经网络成为一种主流方法。图神经网络通过建模信息沿着交互关系传播的权重, 实现更复杂的偏好表示学习, 从而进一步提升了群组偏好表示的精度和效果。本文从方法的发展角度对持续性群组的偏好表示学习进行了清晰的梳理。

对于临时性群组的表示学习, 本文针对不同的策略和技术途径对方法进行了分类梳理。由于临时性群组缺少(群组-项目)交互历史, 传统的群组推荐方法则不再适用。在这种情况下, 临时性群组面临冷启动和稀疏性等关键问题。为了应对这些问题, 目前的主流方法采用聚合成员偏好的方式来表示群组偏好, 从而提升推荐质量。本文将现有研究方法归纳为两种不同的策略: 增加约束和引入外部信息。增加约束策略是通过额外的约束条件来降低模型解空间的规模, 从而提升成员偏好优化的结果, 包括多任务联合优化、互信息最大化和预训练等方法。引入外部信息策略是通过引入外部信息来学习泛化性更好的成员偏好向量表示, 从而实现临时性群组的偏好表示学习。其中, 一种常见的方式是引入社交网络等外部信息, 以丰富成员偏好的表示。通过上述这两种策略的组合, 研究者可以更好地解决临时性群组推荐中的冷启动和稀疏性问题, 并提高推荐的准确性和质量。

2.2.1 基于启发式聚合策略的偏好表示方法

传统的群组推荐系统通常依赖于启发式聚合策略获取群组的偏好表示^[22]。这种方法简单高效, 因而得到了广泛的应用。其中, 启发式聚合策略是指

基于简单规则或者启发式方法表示群组偏好, 例如将群组成员的偏好取平均或加权平均。启发式聚合策略可以分为偏好聚合和评分聚合^[23]。偏好聚合策略将所有用户的偏好组合成一个群组偏好, 从而实现群组推荐; 评分聚合策略则通过计算每个成员的推荐列表, 将成员的列表组合生成群组的推荐列表。图 4 展示了偏好聚合策略和评分聚合策略的过程对比。

经典的聚合策略包括平均策略 (Average, AVG)^[24]、最小痛苦策略 (Least misery, LM)^[12] 和最大满意度策略 (Maximum satisfaction, MS)^[25] 等。平均策略将群组中成员的平均得分作为群组得分, 从而最大化整体的群组满意度; 最小痛苦策略首先让所有成员对项目单独评分, 然后将最低评分视为群组对项目的评分, 从而不断优化群组得分的下限^[26]; 与之相反, 最大满意度策略将最高分视为群组对项目的评分。Baltrunas 等^[27] 的研究指出, 上述启发式聚合策略的性能受群组规模和组内成员偏好相似度等因素的影响。对于规模较大且组内成员偏好相似度高的群组, 最大满意度策略表现更佳; 对于组内成员偏好相似度低或中等的群组, 平均策略表现最佳; 对于规模小但组内成员偏好相似度高的群组, 平均策略和最大满意度策略优于最小痛苦策略^[23]。

为了更准确地表示群组偏好, 一些研究者提出不同的策略和技术途径。其中, 一种方法是通过词嵌入技术表示用户特征, 并利用优化方法来调整群组偏好向量, 使其接近大多数成员向量^[28]。然而, 此类研究默认群组成员对群组偏好的影响相同, 而实际情况往往并非如此。因此, 一些学者提出加权融合策略, 通过计算成员的影响力并根据影响力为成员分配权重, 最终通过加权求和来获得群组偏好。这种策略可分为静态模型和交互模型两种, 其中, 静态模型使用预定义的权重^[29-30], 而交互模型根据用户和项目的交互行为来确定权重^[31-32]。

为进一步提高群组偏好的鲁棒性, 一些学者提出混合融合策略, 即将多种融合策略组合形成新的规则表示群组偏好。例如, Jameson^[33] 提出一种中值策略, 通过与均值策略结合, 有效规避恶意评分和篡改信息的影响, 提升了群组偏好的鲁棒性。Du 等^[34] 提出基于纳什均衡的混合融合策略, 通过多层感知器获得用户和项目之间的潜在特征向量后, 考虑成员推荐分数, 以确保群体用户的平均满意度。

尽管这些启发式方法简单直观, 但由于基于预定义规则或启发式方法计算出的用户权重可能无法完全反映其在群组中的真实影响力, 因此其提升效果有限。这主要是因为上述策略与群组的历史交互

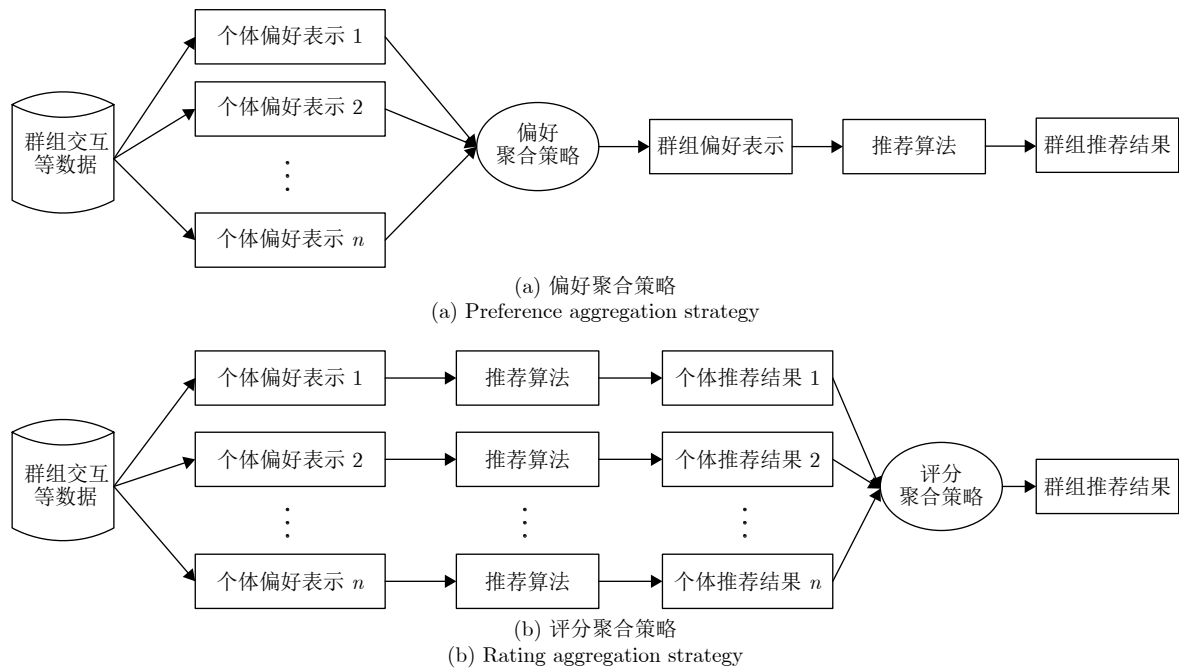


图 4 偏好聚合策略与评分聚合策略的过程对比

Fig. 4 Comparison of the process of preference aggregation strategy and rating aggregation strategy

数据是相互独立的,无法根据交互数据自身的模式来调整成员的权重.为此,一些研究者提出基于概率模型的群组偏好表示方法,以有效地捕捉和反映成员在群组中的真实影响力.

2.2.2 基于概率模型的群组偏好表示方法

成员对群组偏好的影响力存在差异,例如,在面对冲突或分歧时,具有竞争型处理方式的人可能在最终选择上具有较大的影响力,而顺从型处理方式的人则可能影响较小.此外,在不同的情景下,用户选择处理冲突的方式也会不同.为了描述这种影响力的差异,可以利用概率模型对群组成员的偏好分布进行建模.

Quintarelli 等^[35]提出基于上下文影响力的聚合函数,通过考虑上下文依赖影响聚合个体成员偏好来获得群体偏好,实现群组推荐.另外,西班牙马德里大学的 Quijano-Sanchez 等^[36]结合个性化偏好信息和用户社交信任关系,提出 HappyMovie 系统.该系统不仅考虑了个性化偏好信息,还基于遗传算法计算成员信任水平.信任水平通过一系列社会因素来衡量,如社交网络中的距离、共同朋友数量、关系强度和持续时间、社会距离,以及同时出现在照片中的比例等^[37-38].Hu 等^[39]将从历史交互数据中提取的成员特征分为个性化特征和群体特征.所有成员的群体特征可以表征群组的偏好特征,并提出一种深度学习群组推荐模型 (Deep learning group recommendation, DLGR).

此外, Liu 等^[18]提出一种基于主题模型的群组偏好聚合方法 PIT,在计算群组的主题分布时考虑用户的个人影响力.该研究假设用户的影响力越大,对群组的主题分布就会产生越大的影响.类似地, Yuan 等^[19]提出一种基于隐性狄利克雷的模型 COM 来建模群组活动的生成过程.该模型认为,群组最终选择某个项目的过程应该包含成员对候选项目的观点表达和成员观点的加权以达成群组共识.文献^[19]指出某个用户在选择某个候选项目时,如果该项目符合群组的主题,且用户在该主题上具有非常强的专业性知识,那么该用户在此时会对群组有较大的影响力.

PIT 和 COM 提供了不同的策略来聚合群组成员的偏好,但这些方法只考虑〈群组-项目〉和〈用户-项目〉之间的交互,忽略了群组内用户之间的交互.在实际的群组决策过程中,成员往往会在做出决定之前进行讨论,因此,如何有效地动态建模群组内用户之间的交互是一个亟待解决的挑战.

2.2.3 基于注意力机制的群组偏好表示方法

Bahdanau 等^[40]于 2015 年在国际表征学习大会 (International Conference on Learning Representations, ICLR) 上首次提出用于神经机器翻译的注意力模型.目前,注意力模型已在各个领域得到广泛应用,包括自然语言处理、语音识别、推荐系统和计算机视觉等^[41-45].该方法能从大量交互数据中学习数据集本身的模式,因此,能够有效缓解基于

预定义聚合策略和基于概率模型的群组偏好表示方法的缺陷. 注意力机制能够根据不同的上下文动态调整群组成员的重要性, 从而更加精确地识别用户在群组中的影响力, 进而更好地表示群组的偏好.

Cao 等^[46] 首次将注意力模型引入群组推荐领域, 提出注意力群组推荐模型 (Attentive group recommendation, AGREE). 该模型通过学习成员在特定项目下的影响力来动态地加权成员的偏好, 从而得到群组的偏好表示. 如图 5 所示, AGREE 首先获取成员、项目和群组的低维向量表示 $e_{u_t^g}$, e_v 和 e_g . 然后, 通过注意力神经网络获取组内成员 u_t^g 的相对权重 $\alpha(v, t)$, 形式化描述为

$$o(v, t) = \mathbf{h}^T \text{ReLU}(\mathbf{P}_v e_v + \mathbf{P}_u e_{u_t^g} + b) \quad (1)$$

$$\alpha(v, t) = \text{softmax}(o(v, t)) = \frac{\exp\{o(v, t)\}}{\sum_{1 \leq t' \leq |g|} \exp\{o(v, t')\}} \quad (2)$$

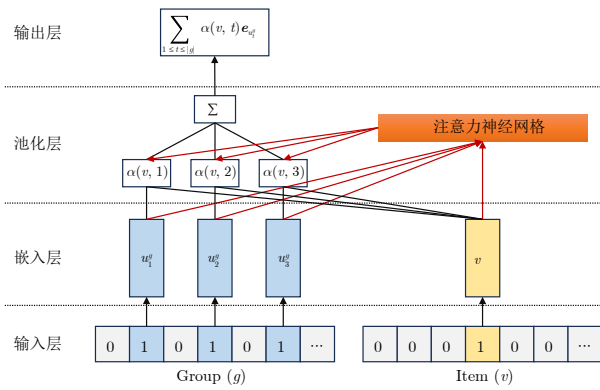


图 5 AGREE 模型结构图

Fig. 5 Diagram of AGREE model structure

式中, \mathbf{P}_v , \mathbf{P}_u , \mathbf{h} 和 b 均为模型参数, $|g|$ 是群组 g 中所含成员的数量. 值得注意的是, 相对权重的取值不仅与成员自身的偏好有关, 同时也与候选项目 v 有关. 因此, 该算法中的相对权重 $\alpha(v, t)$ 可以有效地建模成员在群组决定是否选择候选项目 v 时的影响力. 最后, 基于相对权重聚合成员的偏好向量, 并加上群组自身的偏好向量后得到群组偏好的最终表示, 如式 (3) 所示:

$$\mathbf{g}(j) = \underbrace{\sum_{1 \leq t \leq |g|} \alpha(j, t) e_{u_t^g}}_{\text{成员偏好聚合}} + \underbrace{e_g}_{\text{群组偏好向量}} \quad (3)$$

继 AGREE 算法之后, 基于注意力网络的群组推荐模型不断涌现. 例如, Tran 等^[47] 提出混合子注意力网络群组推荐模型 MoSAN (Medley of sub-attention network), 该模型通过建模用户之间的交互关系来学习用户对群组的影响力. 如图 6 所示, 该模型首先采用一个双层的全连接网络建模用户之间的交互, 从而获取给定用户上下文 c_g 的情况下, 群组的偏好表示 \mathbf{g}_l , 如式 (4) 和式 (5) 所示:

$$\alpha(g, l, m) = \mathbf{w}^T \sigma(\mathbf{W}_c c_l + \mathbf{W}_u e_{u_{m, m \neq l}} + b) + d \quad (4)$$

$$\mathbf{g}_l = \sum_{m=1, m \neq l}^{|g|} \alpha(g, l, m) e_{u_m} \quad (5)$$

式中, \mathbf{W}_c , \mathbf{W}_u 和 \mathbf{w} 是模型的参数, c_l 与 e_{u_m} 分别表示第 l 个用户的上下文和第 m 个用户的嵌入向量, b 和 d 是偏置, $\sigma(\cdot)$ 表示激活函数. 由此可见, $\alpha(g, l, m)$ 可以表示在群组 g 中, 成员 u_m 对成员 u_l 的影响力. \mathbf{g}_l 通过聚合除成员 u_l 之外所有其他成员偏好得到. 最后, 对 \mathbf{g}_l 求和得到群组的最终偏

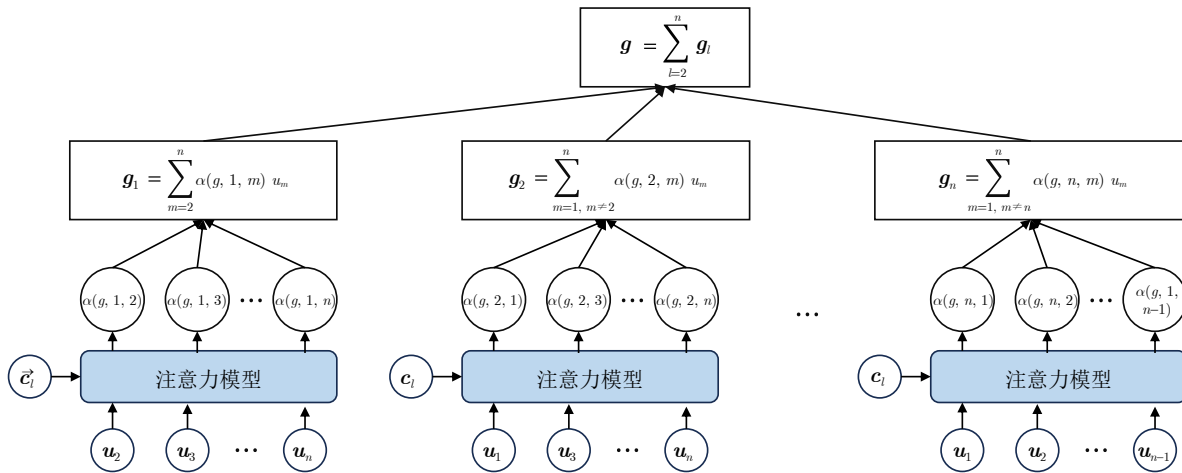


图 6 MoSAN 算法结构图

Fig. 6 Diagram of MoSAN algorithm structure

好表示: $\mathbf{g} = \sum_{l=1}^{|\mathcal{g}|} \mathbf{g}_l$. 值得注意的是, MoSAN 算法中尽管使用了注意力机制, 但注意力机制仅用于计算两两用户之间的影响力, 在聚合成员偏好时仍使用简单快捷的求和方法.

AGREE 和 MoSAN 算法分别考虑了用户在决策时的影响力和用户之间的影响力. 然而, 现实中的群组往往对应着一些主题, 这种主题属性会自上而下地影响每一位成员, 使得成员在群组内表现出的偏好会趋近群组的主题. He 等^[48] 提出基于双层注意力网络的群组推荐 (Group recommendation using attentive dual influences, GRADI) 模型, 该模型同时考虑群组对成员的影响和成员对群组的影响. 如图 7 所示, GRADI 在进行群组偏好表示时主要包含三个模块: 输入层、嵌入层和聚合层. 其中, 输入层包括群组 g 的嵌入向量 \mathbf{e}_g 以及所包含的成员嵌入向量 $\mathbf{e}_{u_i^g}$. 嵌入层采用多层感知器建模群组 g 对成员 u_i^g 偏好的影响 $\mathbf{c}_i(g)$, 从而获取成员 u_i^g 在特定群组 g 中的表示 $\mathbf{h}_{u_i^g}$, 如式 (6) 和式 (7) 所示:

$$\mathbf{c}_i(g) = \text{MLP}([\mathbf{g}_l, \mathbf{u}_i^l]) \quad (6)$$

$$\mathbf{h}_{u_i^g} = \mathbf{e}_{u_i^g} + \mathbf{c}_i(g) \odot \mathbf{e}_g \quad (7)$$

式中, $[\cdot, \cdot]$ 表示拼接操作, \odot 表示逐元素相乘操作. 在聚合层中, GRADI 将项目的表示作为关键向量, 成员的表示作为查询向量和值向量, 通过多头注意力机制计算用户在特定候选项目下对群组的影响力 α_{tv}^d , 如式 (8) 所示:

$$\alpha_{tv}^d = \text{softmax}_t \left(\frac{\mathbf{h}_{u_i^g}^{d\text{T}} \mathbf{h}_v^d}{\sqrt{d_h}} \right) \quad (8)$$

式中, d 表示第 d 个头; d_h 是向量的维度; $\mathbf{h}_{u_i^g}^d$ 和 \mathbf{h}_v^d 分别由成员表示 $\mathbf{h}_{u_i^g}$ 和项目表示 \mathbf{h}_v 通过线性变换得到: $\mathbf{h}_{u_i^g}^d = \mathbf{h}_{u_i^g} \mathbf{W}_d^U$, $\mathbf{h}_v^d = \mathbf{h}_v \mathbf{W}_d^V$, \mathbf{W}_d^U 和 \mathbf{W}_d^V 分别表示用户和项目第 d 个头的变换矩阵. 然后, 根据

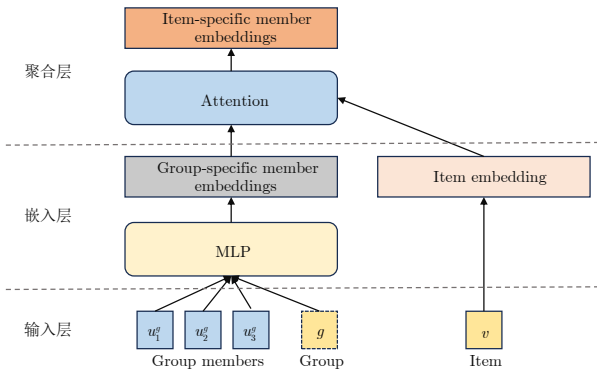


图 7 GRADI 算法结构图

Fig. 7 Diagram of GRADI algorithm structure

该影响力聚合所有成员偏好得到群组在不同头下的偏好表示 \mathbf{h}_g^d , 最后拼接所有头中的表示并进行线性变换后得到群组最终的偏好表示 \mathbf{h}_g :

$$\mathbf{h}_g^d = \sum_t \alpha_{tv}^d \mathbf{h}_{u_i^g}^d \quad (9)$$

$$\mathbf{h}_g = [\mathbf{h}_g^1, \dots, \mathbf{h}_g^k] \mathbf{W}^o \quad (10)$$

式中, \mathbf{W}^o 是线性变换矩阵, $[\cdot, \cdot]$ 是拼接操作.

尽管 GRADI 考虑了自上而下的影响力, 但是该算法与 AGREE 和 MoSAN 都存在一个共性问题, 即没有充分考虑用户、群组和项目之间的交互关系. 为了解决这一问题, Zhang 等^[49] 提出基于同质性效应的局部与全局信息融合的深度群组推荐模型 GLIF (Global and local information fusion neural network). 该模型通过多种交互关系和同质性理论, 融合相似群组作为全局信息, 以增强群组和项目的表示, 进一步提升群组推荐的效果. 如图 8 所示, GLIF 的表示层主要包含四个模块: 项目表示模块、群组局部信息表示模块、群组全局信息表示模块和融合模块. 项目表示模块采用注意力算子聚合与当前项目存在交互的用户集合, 从而获取项目表示; 群组局部信息表示模块通过聚合与当前群组存在交互的用户集合和项目集合获取群组的局部表示; 群组全局信息表示模块基于同质化理论^[50], 采用相似的群组来获取群组的全局信息表示; 融合模块采用注意力算子将群组的全局信息和局部信息融合得到群组偏好的最终表示.

Liang 等^[51] 提出了一种层级注意力网络 HAN-CDGR (Hierarchical attention network for cross-domain group recommendation), 旨在建模成员和群组的偏好表示. 该方法首先利用自注意力网络建模用户在不同群组下的偏好, 并结合用户自身的嵌入向量, 得到用户在特定群组中的偏好表示 \mathbf{h}_u^g . 随后, 采用线性注意力网络聚合用户在不同群组中的偏好表示, 得到用户的最终表示 \mathbf{h}_u :

$$\mathbf{h}_u^g = \mathbf{m}^T \mathbf{O} + \mathbf{e}_u \quad (11)$$

$$\mathbf{h}_u = \sum_{g \in \mathcal{H}_u} \alpha(g, u) \mathbf{h}_u^g \quad (12)$$

式中, \mathbf{O} 是通过自注意力算子计算得到的群组偏好矩阵, \mathbf{m} 是用户 u 的指示向量. 在获得用户在特定群组下的偏好 \mathbf{h}_u^g 后, HAN-CDGR 采用 AGREE 中的注意力方法来得到群组的偏好表示 \mathbf{h}_g , 具体如式 (1) ~ (3) 所示. 从算法实现的角度来看, HAN-CDGR 采用层次注意力网络, 分别对用户与用户、用户与群组之间的影响力进行建模, 以提升模型的

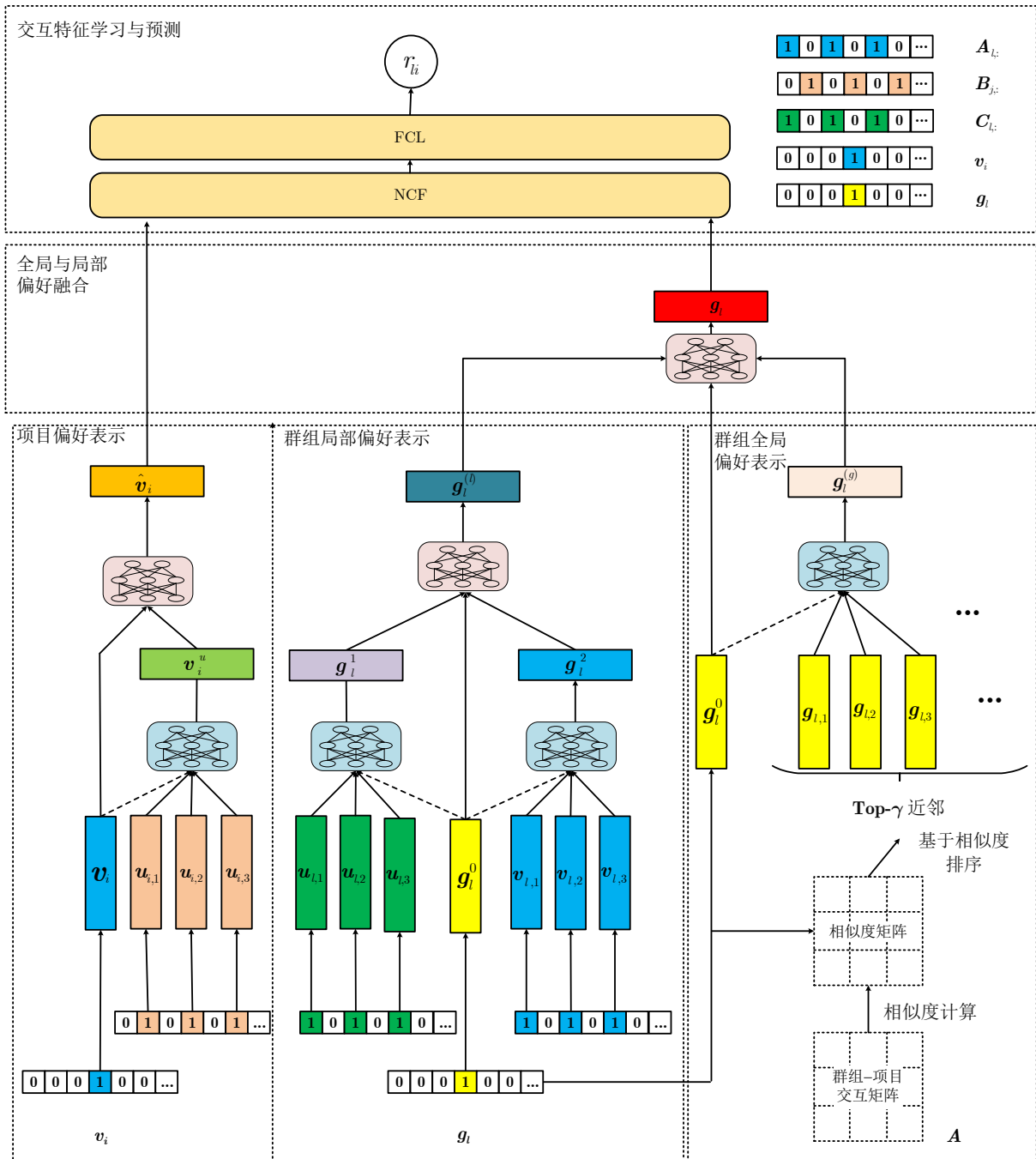


图 8 GLIF 算法框图

Fig.8 Block diagram of GLIF algorithm

表达能力. 尽管如此, 数据稀疏性的问题依然存在. 为解决这一问题, HAN-CDGR 进一步提出基于对抗学习的跨域群组推荐方法. 该方法通过利用相同或相似的用户作为桥梁, 借助域对抗学习将源域中丰富的个性化知识迁移到目标域的群组推荐任务中. 损失函数设计如下:

$$\mathcal{L}_{\text{user}}(\theta_p^s, \theta_p^u, \theta_c) = \mathcal{L}_p^t(\theta_p^u) + \beta \mathcal{L}_p^s(\theta_p^s) - \gamma \mathcal{L}_c(\theta_c) \quad (13)$$

式中, θ_p^s 和 θ_p^u 分别表示源域和目标域中个性化预测模块的参数, θ_c 是域判别器中的参数, β 和 γ 是超参数, $\mathcal{L}_p^s(\theta_p^s)$ 是源域中个性化预测的损失函数, $\mathcal{L}_p^t(\theta_p^u)$ 是目标域中个性化预测的损失函数, $\mathcal{L}_c(\theta_c)$ 是域判别器的损失函数. 通过域对抗学习, 源域中用户的偏好知识能够指导目标域中相同或相似用户的偏好表示, 从而有效地缓解了数据稀疏导致的群组推荐任务中用户偏好学习不充分的问题.

2.2.4 基于图神经网络的群组偏好表示方法

基于注意力网络的群组推荐方法旨在利用神经网络强大的建模能力,主动地从交互数据中学习用户在群组中的影响力,进而聚合群组成员的偏好(或成员的推荐结果),以获得群组的偏好(或群组的推荐结果).尽管这些方法在区分用户在群组中的角色信息方面表现出色,但大多基于群组、用户或物品的一阶交互关系进行建模,未充分利用历史数据中的多种高阶交互关系.为此,研究人员采用图神经网络(Graph neural network, GNN)来挖掘高阶交互关系,实现群组偏好的表示学习.这类方法利用图神经网络对群组的交互结构和成员之间的关系进行建模,并将这些信息用于表示群组的偏好.

根据节点和交互关系的异同,基于图神经网络的方法可以进一步细分为基于同构图神经网络和基于异构图神经网络的方法.基于同构图神经网络的方法,通常采用统一的注意力算子来聚合邻居节点;而基于异构图神经网络的方法考虑到交互图中节点和关系的异构性,通常采用不同的注意力算子来聚合不同类型的节点和关系.通过图神经网络,可以更好地捕捉群组成员之间的复杂关系和群组的整体特征,从而实现更准确的群组偏好表示.

He等^[52]提出了一个注意力多视图嵌入模型GAME(Graphical and attentive multi-view embedding)来学习群组、用户和项目的表示.GAME算法首先通过注意力算子GASE(Graph-aware single-view embeddings)聚合用户(项目)节点不同类型的邻居节点,获取不同视角下的用户(项目)表示,然后通过拼接这些表示来获取更加鲁棒的表示向量.以用户为例,通过注意力算子获取项目视角下的用户偏好表示可以形式化描述为

$$\begin{aligned} \mathbf{h}_u^{(v)} &= \text{GASE}(\mathcal{G}_{UV}, u) = \sum_{j \in R_V(u)} \alpha_j^{(v)} \mathbf{e}_{v_j} \quad (14) \\ \alpha_j^{(v)} &= \text{ATT}(\mathbf{e}_u, \mathbf{e}_{v_j}) = \\ & \underset{j}{\text{softmax}} \left(\mathbf{w}^{(v)\top} \sigma \left(\mathbf{W}^{(v)} [\mathbf{e}_u, \mathbf{e}_{v_j}] \right) + b^{(v)} \right) \quad (15) \end{aligned}$$

式中, \mathbf{e}_{v_j} 是第 j 个项目的嵌入向量, $R_V(u)$ 表示与用户 u 产生过历史交互的项目集合, $\mathbf{w}^{(v)}$, $\mathbf{W}^{(v)}$ 和 $b^{(v)}$ 均为网络参数.同理,可以分别获得群组视角下的用户表示 $\mathbf{h}_u^{(g)}$ 、用户视角下的项目表示 $\mathbf{h}_v^{(u)}$ 和群组视角下的项目表示 $\mathbf{h}_v^{(g)}$.在得到用户和项目在不同视角下的表示后, GAME 算法采用拼接操作并将其输入到多层感知器中,从而获得更加鲁棒的用户偏好表示和项目特征表示.值得注意的是, GAME

算法需要学习每个群组的嵌入向量,因此在遇到冷启动群组时,通常需要重新训练整个网络.

Deng等^[53]提出一种基于知识图谱的注意力群体推荐模型KGAG(Knowledge graph-based attentive group recommendation),通过引入外部知识建立项目与项目之间的高阶交互关系,在一定程度上缓解了因交互稀疏导致的表示学习不充分问题.如图9(a)所示,KGAG算法主要包含信息传播和偏好聚合两个模块.信息传播模块利用知识图谱和〈用户-项目〉交互记录构建协同图谱(如图9(a)左侧所示),并通过信息传播方式更新节点的表示:

$$\mathbf{e}_{N_e} = \sum_{(e, r, e_t)} \tilde{\pi}(e, r, e_t) \mathbf{e}_t \quad (16)$$

式中, \mathbf{e}_{N_e} 表示从实体 e 所有一阶近邻传播过来的信息, $\tilde{\pi}(e, r, e_t)$ 表示 \mathbf{e}_t 沿着关系 r 向 e 传播时的权重,通过式(17)和式(18)获取:

$$\pi(e, r, e_t) = \mathbf{i}_e \cdot \mathbf{r} \quad (17)$$

$$\tilde{\pi}(e, r, e_t) = \frac{\exp(\pi(e, r, e_t))}{\sum_{\tilde{\pi}(e, r_i, e_i) \in \mathcal{G}} \exp(\pi(e, r_i, e_i))} \quad (18)$$

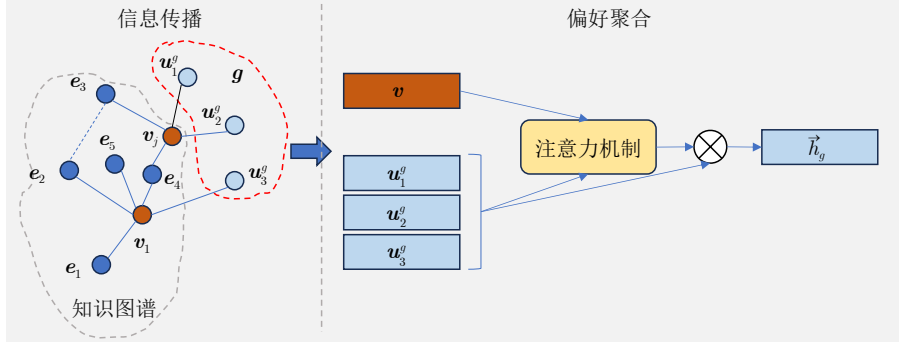
式中,如果 e 代表的是成员,那么 \mathbf{i}_e 表示与成员 e 有过交互的项目嵌入;如果 e 代表的是项目,那么 \mathbf{i}_e 表示与项目 e 有过交互的群组内所有成员嵌入向量的均值; \mathbf{r} 是关系 r 的嵌入向量.通过上述操作获取来自所有邻居节点的信息 \mathbf{e}_{N_e} 后,可以通过图卷积网络(Graph convolutional network, GCN)或者图采样与聚合器(Graph sample and aggregate, GraphSage)将 \mathbf{e}_{N_e} 和 \mathbf{e} 聚合得到成员和项目的表示,分别记作 \mathbf{h}_u 和 \mathbf{h}_v .偏好聚合模块考虑了群组内两种交互影响:自我坚持(Self persistence, SP)和同辈影响(Peer influence, PI),如图9(b)所示.自我坚持是成员个体对候选项目的偏好程度,可以一定程度上影响群组整体的选择,因此可以视为成员个体对群组选择的影响,通过式(19)获取,即

$$\alpha_{SP}(g, t, v) = \mathbf{h}_{u_t^g} \cdot \mathbf{h}_v \quad (19)$$

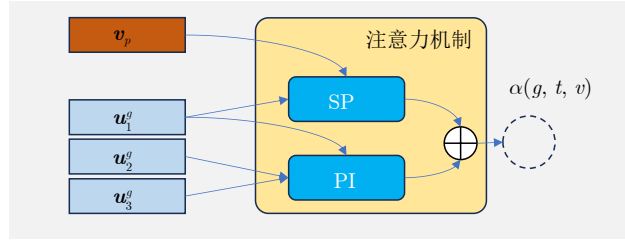
同辈影响表示成员个体的偏好会受到同组中其他成员的影响,即

$$\begin{aligned} \alpha_{PI}(g, t) &= \\ & \mathbf{w}_c^T \sigma \left(\mathbf{W}_{c1} \mathbf{h}_{u_t^g} + \mathbf{W}_{c2} [\mathbf{h}_u, \dots]_{u \in S_{g,t}^P} + b \right) \quad (20) \end{aligned}$$

式中, \mathbf{w}_c , \mathbf{W}_{c1} , \mathbf{W}_{c2} 和 b 都是模型参数, $[\cdot, \cdot]$ 表示拼接操作, $S_{g,t}^P$ 表示在群组 g 中除去 u_t 之外的所有成员的集合.最后通过求和的方式获取用户综合的影响力权重,并对成员偏好进行加权求和得到群组



(a) 信息传播模块和偏好聚合模块
(a) Information propagation and preference aggregation module



(b) 群内的交互影响
(b) Interactive influence across members within group

图 9 KGAG 模型示意图

Fig.9 Schematic diagram of KGAG model

的偏好表示 \mathbf{h}_g .

Zhang 等^[54] 提出了基于群组交互异构图注意力网络模型 HetGRec (Heterogeneous graph attention network for group recommendation). 该模型通过基于元路径的随机游走重启策略获取与目标节点强相关的邻居集合, 并使用双层注意力子网络聚合这些邻居节点, 从而获取更丰富的群组和项目表示. 最后, 基于神经协同过滤网络实现群组推荐. 如图 10 所示, HetGRec 游走序列遵循马尔科夫过程, 为了找到高度相关的高阶近邻, 在选择下一节点之前会先根据概率 $1/(1 + \varepsilon_p)$ 选择是否重新回到中心节点, 通过调整超参数 ε_p 控制近邻与中心节点的平均距离, ε_p 值越小, 平均距离则越小, 采样的近邻则越可能相关. 如果不返回, 则依据概率 $\Pr(n_{t+1} = x | n_t = v, p)$ 选择下一个节点, 其中, $p \in \mathcal{P}$ 是元路径. 设 $\mathcal{N}^p(u)$ 是节点 u 基于元路径 p 找到的近邻集合. HetGRec 通过节点级注意力算子聚合 $\mathcal{N}^p(u)$ 得到节点 u 基于元路径 p 的近邻信息, 形式化描述为

$$\mathbf{h}_u^p = f_{\text{node}}(\mathcal{N}^p(u)) = LN \left(\sum_{j \in \mathcal{N}^p(u)} \alpha_{uj}^p \mathbf{h}_j' \right) \quad (21)$$

$$\alpha_{uj}^p = \text{softmax}_j \left(\exp \left(\text{LeakyReLU} \left(\partial_p^T [\mathbf{h}_u', \mathbf{h}_j'] \right) \right) \right) \quad (22)$$

式中, \mathbf{h}_j' 是由节点 (用户、群组或项目) 的嵌入向量经过一阶线性变换得到, $LN(\cdot)$ 是层归一化函数, ∂_p 是模型参数. 在获取到 \mathbf{h}_u^p 后, HetGRec 进一步使用语义级注意力算子获取群组和项目的最终表示 \mathbf{h}_g 和 \mathbf{h}_v , 形式化描述为

$$\mathbf{h}_g = f_{\text{sem}}(\mathcal{H}(g), \tau) \quad (23)$$

$$\mathbf{h}_v = f_{\text{sem}}(\mathcal{H}(v), \tau) \quad (24)$$

其中,

$$f_{\text{sem}}(\mathcal{H}(v), \tau) = LN \left(\sum_{\mathbf{h}_v^k \in \mathcal{H}(v)} \beta_k \mathbf{h}_v^k \right) \quad (25)$$

$$\beta_k = \text{softmax}_k \left(\exp \left(\mathbf{q}_{\tau(v)}^T \mathbf{W}_{\tau(v)} \mathbf{h}_v^k \right) \right) \quad (26)$$

式中, τ 是将节点映射到对应的节点类型, $\mathbf{q}_{\tau(v)}$ 和 $\mathbf{W}_{\tau(v)}$ 均为模型参数, $\mathcal{H}(g)$ 表示节点 g 所有不同元路径下的近邻信息. 通过双层注意力子网络的聚合操作, HetGRec 可以计算邻居节点和元路径的重要性, 分别从节点级别和语义级别融合目标节点的邻居节点, 得到语义更加丰富的群组和项目的表示.

2.2.5 面向临时性群组的偏好表示方法

目前, 群组推荐领域的研究主要聚焦于具有丰富历史活动的持续性群组, 而对于临时性群组, 其

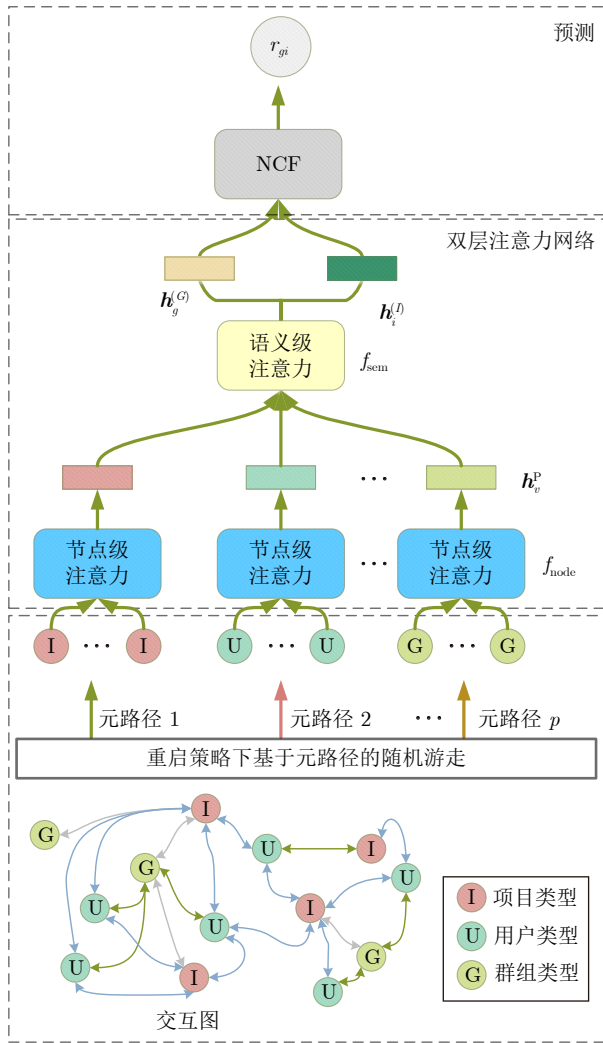


图 10 HetGRec 算法结构图

Fig. 10 Diagram of HetGRec algorithm structure

缺乏足够的历史交互数据, 因此, 临时性群组推荐也被称为冷启动群组推荐, 传统的群组推荐方法不适用于这类冷启动推荐场景. 这一问题的根源在于临时性群组往往缺乏充足的历史记录, 无法直接建模群组偏好. 本文将现有的方法归纳为两类: 增加约束类方法和引入外部信息类方法.

增加约束类方法旨在通过加强用户、项目和群组之间的交互关系来对群组表示施加额外约束, 例如互信息最大化、预训练和多任务联合优化等. 这类方法通过获得更鲁棒的模型参数, 以提升表示学习的效果. 而引入外部信息类方法则试图通过引入社交等外部信息来增强用户表示, 并将这些信息传递到群组表示中, 以提升临时性群组推荐的性能.

以 GroupIM (Group information maximization) 为例^[55], 该模型结合了互信息最大化方法和注意力网络, 设计了两种约束: 互信息最大化约束和上

下文约束. 其中, 互信息最大化约束旨在衡量正样本对之间的互信息, 以此作为监督信号进行优化; 而上下文约束则利用成员的历史交互信息作为监督标签, 通过分类器预测结果对成员的损失进行加权求和, 从而提供更多监督信息, 优化群组表示. GroupIM 互信息最大化约束的形式化描述为

$$L_{MI} = -\frac{1}{|G|} \sum_{g \in G} \frac{1}{\alpha_g} \left[\log_2 D_{ug} + \sum_{j=1}^{M_g} E_{\tilde{u} \sim P_N} \log_2 (1 - D_{\tilde{u}g}) \right] \quad (27)$$

式中, $D_{ug} = D(\mathbf{e}_u, \mathbf{e}_g) = \sigma(\mathbf{e}_u^T \mathbf{W} \mathbf{e}_g)$ 是分类器; $\sigma(\cdot)$ 是 logistic 函数, 该函数基于正样本和负样本之间的 JS (Jensen-Shannon) 散度^[56] 来衡量正样本对 $(\mathbf{e}_u, \mathbf{e}_g)$ 之间的互信息, \mathbf{W} 是权重矩阵; M_g 是为群组 g 采样的非成员用户数量; $\alpha_g = |g| + M_g$; \tilde{u} 是负样本用户. 上下文约束的形式化描述为

$$L_{UC} = -\sum_{g \in G} \frac{1}{|x_g|} \sum_{i \in I} \sum_{1 \leq t \leq |g|} D(\mathbf{e}_{u_t^g}, \mathbf{e}_g) x_{ui} \log \pi_i(\mathbf{e}_g) \quad (28)$$

其中, $\pi(\mathbf{e}_g) = \text{softmax}(\mathbf{W}_I \mathbf{e}_g)$ 用于预测群组或用户在全体项目集 I 上的偏好分布. 该目标函数以成员的历史交互信息作为监督标签, 同时, 以分类器 D_{ug} 预测的结果作为权重, 对所有成员的损失值进行加权求和. 这样做的好处是在模型优化时, 能够为优化 \mathbf{e}_g 提供更多的监督信息, 使隐空间中的群组表示能够更靠近互信息较大的成员.

随着预训练语言模型在自然语言处理中的成功应用, 例如 BERT (Bidirectional encoder representations from transformers)^[57] 和 GPT (Generative pre-trained transformers)^[58-59], 其在各个领域的研究也受到越来越多的关注, 包括群组推荐系统. Zhang 等^[60] 提出预训练-微调群组推荐模型 GBERT (Group recommendation based-on BERT) 实现临时性群组推荐任务. 在预训练阶段, GBERT 设计了三个预训练任务, 包括掩码成员预测、偏好一致性检测和个性化偏好预测, 旨在学习用户在不同群组中的偏好以及偏好之间的关系. 这种方法通过联合优化群组推荐任务与用户分类任务, 获得更鲁棒的群组偏好表示, 可视为一种多任务联合优化的临时性群组偏好表示方法^[61]. 如图 11 所示, 掩码成员预测任务基于 BERT 学习用户在不同群组的偏好偏移, 捕获群组内成员之间的偏好依赖关系, 目标函数 \mathcal{L}_{MMP} 形式化表达式为

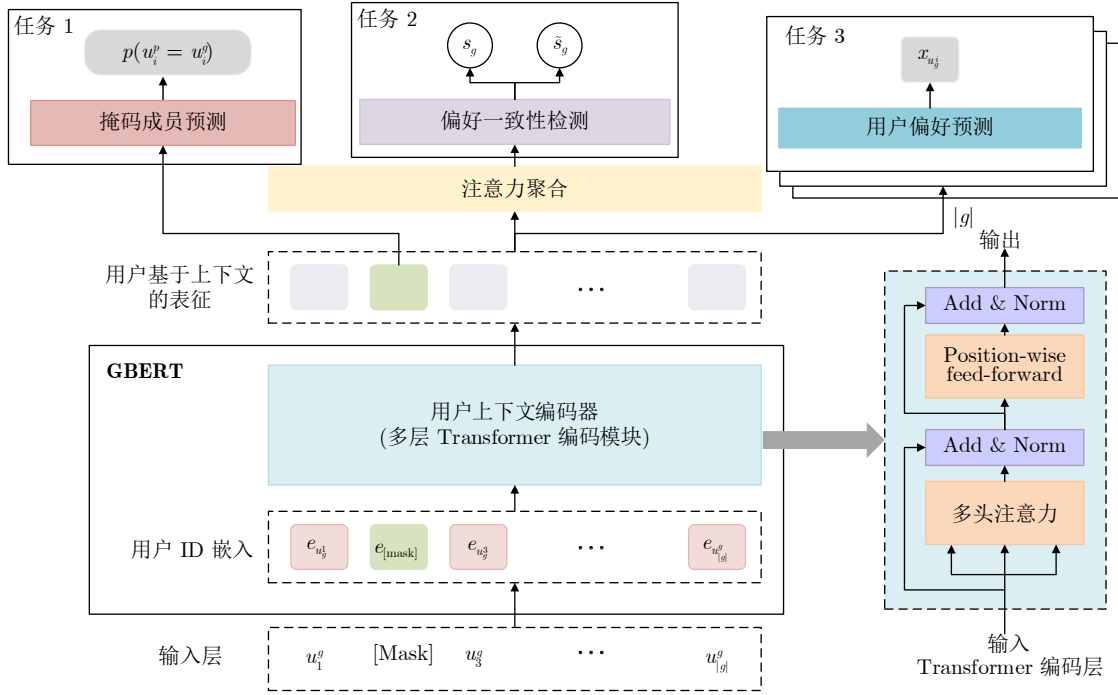


图 11 GBERT 算法预训练阶段流程框图

Fig. 11 GBERT algorithm pre-training phase flowchart

$$\mathcal{L}_{\text{MMP}} = - \sum_{g \in G} \log_2 p(u_i^p = u_i^g) \quad (29)$$

式中, $P(u_i^p = u_i^g)$ 为模型预测第 i 个被随机掩码 (Mask) 屏蔽掉的用户是 u_i^g 的概率. 偏好一致性检测任务通过对比学习抽取高共识水平的偏好特征, 目标函数 \mathcal{L}_{PCD} 形式化表达式为

$$\mathcal{L}_{\text{PCD}} = \sum_{g \in G} (1 - s_g + \tilde{s}_g) \quad (30)$$

式中, $s_g = f(\{u_i^g\})$ 是正样本群组的共识水平; \tilde{s}_g 是负样本群组的共识水平, 负样本群组由使用非成员用户随机替换掉其中一个成员组成. 个性化偏好预测则在潜在空间中对所学到的偏好表示进行约束, 提升模型训练的效率, 目标函数 \mathcal{L}_{IPP} 形式化表达式为

$$\mathcal{L}_{\text{IPP}} = - \sum_{g \in G} \frac{1}{|B_u|} \sum_{1 \leq i \leq |g|} \alpha_i^g p(u_i^g) B_u \quad (31)$$

式中, α_i^g 是成员 u_i^g 的影响力, B_u 是用户 u 在交互矩阵中对应的行, $p(u_i^g)$ 是预测得到的成员在全体项目集 V 上的偏好分布.

在引入外部信息类方法中, 例如, SoAGREE (Social-enhanced attentive group recommendation)^[62] 和 SIGR (Social influence-based group representation learning for group recommendation)^[32] 利用社交等外部信息增强用户表示. So-

AGREE 利用注意力算子聚合所有关注目标用户的用户表示并通过残差连接得到目标用户的最终表示. SIGR 采用一个两阶段的方法, 利用自编码器抽取用户的社交特征, 并将其作为推荐模型的输入来预测用户的影响力. 该算法借助用户所处的社交网络, 首先计算网络的各种特征属性, 包括 PageRank 中心性、接近中心性、介数中心性、特征向量中心性等全局结构特征以及每个节点的局部结构特征, 从而得到用户的社交特征 X_L , 然后采用随机掩码的方式将部分特征属性进行遮挡以获得残缺的用户社交特征 X_0 , 并使用堆叠的降噪自编码器 (Stacked denoising auto encoder, SDAE) 重构用户的社交特征 X_L . 以第 l 层为例, 重构损失为

$$O_{\text{SDAE}} = \frac{1}{2} \sum_l \left(\frac{\|\sigma(\mathbf{X}_{l-1} \mathbf{W}_l + b_l) - \mathbf{X}_l\|_2^2}{\rho_X^2} + \frac{\|\mathbf{W}_l\|_2^2}{\rho_W^2} + \frac{\|b_l\|_2^2}{\rho_b^2} \right) \quad (32)$$

式中, $\mathbf{X}_l \sim N(\sigma(\mathbf{X}_{l-1} \mathbf{W}_l + b_l), \rho_X^2 I)$ 是 SDAE 第 l 层输出的特征向量, $\mathbf{W}_l \sim N(0, \rho_W^2 I)$ 是 SDAE 第 l 层网络的参数矩阵, $b_l \sim N(0, \rho_b^2 I)$ 是第 l 层网络的偏置向量. SIGR 将第 $L/2$ 层输出的特征向量视为用户社交信息的高阶特征 \mathbf{h}_i . 然后基于 \mathbf{h}_i 计算成员的影响力变量 γ_i , 通过注意力算子聚合成员得到群组的偏好表示 \mathbf{h}_g , 如式 (33) 和式 (34) 所示:

$$\lambda_{ig} = \frac{\exp(\gamma_i)}{\sum_{1 \leq j \leq |g|} \exp(\gamma_j)} \quad (33)$$

$$\mathbf{h}_g = \sum_{1 \leq i \leq |g|} \lambda_{ig} \mathbf{h}_i \quad (34)$$

然而, SIGR 存在两个阶段独立优化的问题, 导致学习到的社交特征与最终推荐任务之间的不匹配. 为此, Yin 等^[63] 提出一种集中感知群组推荐算法 CAGR (Centrality-aware group recommender), 通过计算用户的多种中心性特征, 并结合注意力机制来聚合邻居节点信息, 以获取更准确的用户偏好表示. CAGR 首先基于用户的社交信息计算多种中心性特征 c_k , 包括 PageRank 中心性、接近中心性、介数中心性等. 然后, 将当前用户的一阶邻居节点依据不同的中心性特征值进行排序, 选取目标用户在不同中心性特征值下的 Top- K 个邻居节点, 采用池化方法^[64] 将同一个特征值下的 K 个邻居节点聚合, 得到该特征值下的邻居节点信息表示 $\mathbf{e}_u^{c_k}$, 最后采用注意力算子聚合这些邻居节点得到用户的偏好表示 \mathbf{h}'_u , 形式化表达如式 (35) 和式 (36) 所示:

$$\mathbf{h}'_u = \sum_{k=1}^{|C|} \alpha_u^{c_k} \mathbf{e}_u^{c_k} \quad (35)$$

$$\alpha_u^{c_k} = \frac{\exp(\mathbf{z}_k \cdot \mathbf{e}_u^C)}{\sum_j \exp(\mathbf{z}_j \cdot \mathbf{e}_u^C)} \quad (36)$$

式中, \mathbf{e}_u^C 由所有不同视角下邻居信息表示 $\mathbf{e}_u^{c_k}$ 拼接得到; \mathbf{z}_k 是参数向量, 表示特定的网络中心性类型.

为了进一步建模成员之间的交互, CAGR 采用多头注意力模型作为成员偏好的编码器, 从而获取成员的最终偏好表示 \mathbf{h}_u , 即

$$\mathbf{h}_u = \tanh(\mathbf{W}_s \mathbf{O}_u + \mathbf{b}_s) \quad (37)$$

$$\mathbf{O} = [\mathbf{M}_1, \dots, \mathbf{M}_h] \quad (38)$$

$$\mathbf{M}_i = \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \quad (39)$$

式中, \mathbf{W}_i^Q , \mathbf{W}_i^K 和 \mathbf{W}_i^V 是第 i 个头的变换矩阵, \mathbf{Q} , \mathbf{K} 和 \mathbf{V} 是均由同一个组的成员嵌入向量变换得到的矩阵, $\text{Attention}(\cdot)$ 是自注意力算子. 得到成员的嵌入后, 通过注意力算子计算不同成员的影响力, 并加权求和得到群组的偏好向量 \mathbf{h}_g , 即

$$\lambda_s = \frac{\exp(\mathbf{h}_{u_i^g}^\top \cdot \mathbf{a}_s)}{\sum_j \exp(\mathbf{h}_{u_j^g}^\top \cdot \mathbf{a}_s)} \quad (40)$$

$$\mathbf{h}_g = f'_a(\mathbf{h}_{u_i^g} | 1 \leq i \leq |g|) = \sum_{1 \leq i \leq |g|} \lambda_i \mathbf{h}_{u_i^g}^\top \quad (41)$$

式中, \mathbf{a}_s 是注意力向量.

面向临时性群组推荐的表示学习方法所面临的挑战是如何在稀疏的交互数据下学习到鲁棒的群组表示. 增加约束类方法 (GroupIM 和 GBERT) 通过加强交互关系以获得更鲁棒的模型参数; 而引入外部信息类方法 (SoAGREE、SIGR 和 CAGR) 则通过增强用户表示来提升推荐性能.

2.2.6 表示层方法综合对比分析

综上所述, 针对持续性群组, 研究者通过设计聚合策略, 探索使用概率模型模拟群组的生成过程, 采用注意力机制和图神经网络等技术建模用户、群组和项目之间的交互关系, 以实现更精细的群组偏好表示学习. 对于临时性群组, 因为缺乏历史交互数据, 导致冷启动和稀疏性等问题. 为了解决这些挑战, 研究者提出增加约束和引入外部信息的策略, 以提升群组的偏好表示学习效果. 图 12 对基于深度学习的群组推荐表示学习层的方法进行了系统分类和总结. 这些方法为群组偏好的表示学习提供了新的思路和技术手段, 对于解决持续性和临时性群组推荐问题具有重要意义, 表 2 汇总了每类方法的特点与不足之处. 通过综合利用不同方法和策略, 可以提高群组推荐系统的准确性和效果, 促进推荐技术在实际应用中的进一步发展.

2.3 模型层

模型层的核心任务是计算群组偏好与项目特征之间的匹配得分, 以生成推荐列表. 这一任务主要分为基于内积的方法和基于神经协同过滤的方法.

2.3.1 基于内积的方法

基于内积的方法利用嵌入层将稀疏向量转换为稠密向量, 将群组的偏好向量与项目的特征向量进行内积运算, 得到群组对项目的偏好得分. 具体而言, Tran 等^[47] 通过计算群组偏好向量与项目特征向量的内积来获得群组对项目的偏好得分:

$$r_{ij} = \mathbf{g}_i^\top \mathbf{v}_j \quad (42)$$

式中, \mathbf{g}_i 是第 i 个群组的偏好向量, \mathbf{v}_j 是第 j 个项目的特征向量. Sankar 等^[55] 通过内积的方式获取单个群组对所有候选项目的评分, 然后基于 softmax 函数将所有评分映射到 $[0, 1]$ 区间, 最后排序评分得到推荐列表. 基于内积的方法简单直接, 但仅考虑了群组偏好与项目特征之间的线性关系.

2.3.2 基于神经协同过滤的方法

神经协同过滤方法 (NCF) 采用多层神经网络来替代矩阵分解的操作, 以更好地捕捉群组与项目之间的高阶交互关系^[63]. 首先, 通过群组偏好向量

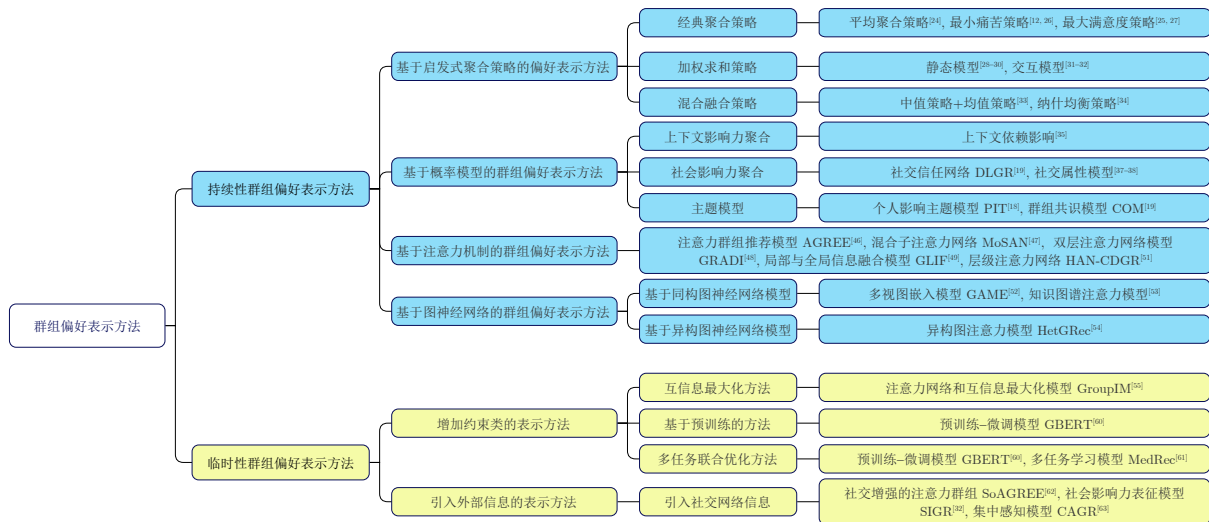


图 12 群组偏好表示方法分类总结

Fig.12 Classification summary of group preference representation methods

表 2 群组偏好表示学习方法对比

Table 2 Comparison of learning methods for group preference representation

表示学习层技术	特点	不足
基于启发式聚合策略的群组偏好表示方法	结合个性化推荐方法和预定义的聚合策略完成群组推荐任务,方法简单高效	无法根据交互数据自身的模式来学习成员之间、成员与群组之间的影响力
基于概率模型的群组偏好表示方法	建模群组的生成过程,采用潜变量表示用户对群组或其他成员的影响力	较依赖于先验分布的假设,无法动态地建模成员用户的影响力
基于注意力机制的群组偏好表示方法	采用注意力机制主动从用户交互记录等信息中学习成员的影响力	数据稀疏性可能导致模型训练低效,使得学习到的影响力不准确
基于图神经网络的群组偏好表示方法	采用图神经网络建模用户、群组和项目之间的高阶交互关系,并结合注意力算子计算信息沿着关系传播的权重,有效缓解因数据稀疏导致推荐效果不佳的问题	可能需要用户的社交信息来构建网络,较难实现;针对冷启动群组,需要重新训练网络
增加约束类的群组偏好表示方法	采用增加约束的方式降低解空间的规模,基于多任务之间的共性特征,提升模型优化的效果	较依赖于预训练数据集的质量;较依赖于任务之间的关联强度
引入外部信息的群组偏好表示方法	通过引入外部信息的方式,增强群组偏好的表示学习,如社交网络信息、项目描述信息和用户评论信息等	外部信息较难获取等问题

和项目特征向量获取群组与项目的交互特征向量。之后,通过多层感知机计算群组偏好与项目特征之间的高阶交互关系,最终计算匹配得分:

$$r_{ij} = MLP(F(\mathbf{g}_i, \mathbf{v}_j)) \quad (43)$$

式中, $MLP(\cdot)$ 是多层感知器, $F(\cdot)$ 是用于将 \mathbf{g}_i 和 \mathbf{v}_j 转成交互特征向量的函数,实现方式为

$$F(\mathbf{g}_i, \mathbf{v}_j) = [(\mathbf{g}_i \otimes \mathbf{v}_j) \parallel \mathbf{g}_i \parallel \mathbf{v}_j] \quad (44)$$

其中, \parallel 表示拼接操作, \otimes 表示哈达玛积。NCF 的优势在于利用多层神经网络使得群组和项目的特征能够得到充分的交叉,进而获得更多的组合特征。

总体而言, NCF 实际上是一个灵活的推荐框架,能够将不同操作得到的特征进行有效组合,利

用神经网络强大的拟合能力进行优化。因此,大多数群组推荐模型都或多或少地受到了 NCF 框架的启发,并采用了类似的方法来计算群组偏好与项目特征之间的匹配得分。

3 群组推荐数据集及评价指标

下面分别对本综述的群组推荐文献中涉及的数据集和评价指标进行简要介绍。

3.1 群组推荐的数据集

群组推荐研究所涉及的数据集主要有 CAMRa-2011、MovieLens 1M、Weeplaces、Yelp、Douban、Meetup、BookCrossing、Jester Joke、Netflix 等。CAMRa2011 是一个电影数据集,包含用户和家庭

对电影的评分记录, 其中一个家庭的成员形成一个群组. MovieLens 1M 是一个包含数百万条用户对电影的评分记录数据集. Weeplaces 是一种基于位置的可视化地图服务, 记录了用户在基于地理位置的社交网络中的签到历史. Yelp 是美国最大的点评网站, 用户可以在网站中对商户进行评分、提交评论, 以及分享购物体验等. Douban 是国内的在线社交网络平台, 用户可以发布或参加各种社交活动, 数据集中记录了这些活动的时间、地点和参与者信息. Meetup 是典型的群组社交平台, 通过收集用户的兴趣爱好和住址信息, 将可能参与同一活动的潜在用户聚集在一起, 形成临时性的线上或线下活动群组. BookCrossing 是一个图书共享阅读网站, 通过跟踪每本书籍的流通情况以及连接相关读者, 实现阅读和社交的目的. Jester Joke 是一个在线推荐和分享笑话的网站, 数据集记录了用户对笑话的评分情况. Netflix 是一家提供会员订阅制流媒体服务的平台, 其数据集包含用户对影片的评分记录以及评分日期等信息. 表 3 提供了这些数据集的主要信息. 这些数据集的选择涵盖了不同领域和应用场景, 从电影评分到位置签到、社交活动以及图书分享等. 通过使用这些多样化的数据集, 研究人员能够在不同的情境下测试和评估群组推荐算法的性能和有效性, 为进一步的研究提供了丰富的实验基础.

3.2 群组推荐评价指标

当前评价群组推荐质量的指标多为 Top- K 推荐指标^[73], 包括命中率 (Hit-rate, $H@K$)、召回率

(Recall, $R@K$) 以及排序指标 (归一化折损累计增益 (Normalized discounted cumulative gain, $N@K$)) 等. 除此之外, 部分推荐算法采用平均绝对误差 (Mean absolute error, MAE) 来衡量个体在群体中的满意度.

3.2.1 命中率

命中率 ($H@K$) 是衡量推荐结果准确性的重要指标, 它通过 hits 数来衡量推荐结果的准确性, hits 指的是群组推荐的项目与其测试集中项目的交集, 即符合推荐的项目恰好是群组实际交互的项目这种条件的项目集合. $H@K$ 的计算方法如式 (45) 所示:

$$H@K = \frac{\#hit@K}{N} \quad (45)$$

式中, $\#hit@K$ 表示在 Top- K 命中测试集中项目的个数, N 是测试集中实例的总数, 可以看出 $H@K$ 值越大表示推荐准确率越高.

3.2.2 召回率

精确率指标无法衡量推荐模型应对长尾分布的能力. 为此, 现实场景中, 除了使用精确率来衡量推荐系统准确度外, 还会使用查全率指标, 即召回率 ($R@K$) 指标来衡量推荐模型在处理长尾项目时的能力. 该指标是正确推荐的项目数与真实项目数之间的比值. 在群组推荐任务中, 较高的召回率表明推荐系统可以尽可能地不遗漏正确的项目, 其计算方法如式 (46) 所示:

$$R@K = \frac{\#hit@K}{TP + FN} \quad (46)$$

式中, $TP + FN$ 是数据集中所有正例的数量.

表 3 群组推荐数据集信息
Table 3 Information of group recommendation dataset

数据集	类别	内容
CAMRa2011 ^[46, 48-49, 54, 62, 66]	电影	包含 602 个用户组成的 290 个群组对 7 710 部电影的评分.
MovieLens 1M ^[31, 47, 49, 53-54, 67]	电影	包含上百万评分记录的电影数据集, 由于该数据集不存在显式的群组, 通常根据用户相似度构建群组.
Weeplaces ^[55, 60]	签到	由于该数据集中不存在显式的群组, 通常将 15 min 以内在同一地点签到且存在朋友关系的用户视为一个群组, 形成包含 8 643 个用户打卡 25 081 个商户的 22 733 个群组.
Yelp ^[32, 53, 55, 60, 68-69]	点评	包含 34 504 个用户对 22 611 个餐厅的点评. 由于该数据集不存在显式的群组, 将在同一时间段内打卡同一个餐厅且存在社交关系的用户视为一个群组, 形成 24 103 个群组.
Douban ^[32, 55, 60, 63, 68-69]	活动	包含 70 743 个用户对 60 028 个活动的评分. 由于该数据集不存在显式的群组, 将参加同一活动的用户视为一个群组, 形成 109 538 个群组.
Meetup ^[31, 47-48, 52, 63, 70]	活动	按照事件的地点, 该数据集分为 Meetup-NYC (纽约市) 和 Meetup-Cal (加利福尼亚). 这两个数据集均没有显式的群组, 通常将参加同一个事件的人视为一个群组. 其中, Meetup-NYC 包含 46 619 个用户、9 428 个群组、2 326 个项目. Meetup-Cal 包含 59 486 个用户、15 207 个群组、4 472 个项目.
BookCrossing ^[67]	书籍	包含 278 858 个用户, 提供 271 379 本书的 1 149 780 个评分. 该数据集不包含显式的群组, 通常通过寻找相似用户构建群组.
Jester Joke ^[71]	笑话	包含 73 421 个用户对 100 个笑话进行的 410 万次评分, 评分范围是 $-10 \sim 10$ 的连续实数. 不包含分组信息, 通过计算用户相似度来进行分组.
Netflix ^[72]	电影	包含 480 507 个用户对 17 770 部电影的 100 480 507 条评价数据, 其中评分以 5 分制为基准. 不包含分组信息, 利用用户的偏好相似信息构造群组.

3.2.3 排序指标

命中率对推荐结果排序不敏感, 即对推荐列表中的项目重新排序是不会影响 $H@K$ 的取值的. 然而, 实际生活中, 推荐列表的顺序至关重要, 群组往往更关注推荐列表前几位的结果. 推荐系统将正确的项目尽可能地排在推荐列表的前面会让用户拥有更好的体验. 为此, 可以用归一化折损累计增益 ($N@K$) 作为推荐结果排序问题的评价指标, 其计算方法如式 (47) 和式 (48) 所示:

$$DCG@K = \sum_{i=1}^K \frac{2^{rel_i} - 1}{\log_2(i+1)} \quad (47)$$

$$N@K = \frac{DCG@K}{IDCG@K} \quad (48)$$

式中, rel_i 取值为 0 或 1, 如果排在第 i 个项目在测试集中, 那么 $rel_i = 1$, 否则 $rel_i = 0$; $IDCG@K$ 表示最佳排序情况下 DCG 的值. 可以看出, 目标项目排序位置越靠前, $N@K$ 值越大.

3.2.4 平均绝对误差

上述评价可以从群组的维度来衡量群组推荐的准确性, 但无法衡量个体在群体中的满意度. 因此, Garcia 等^[24] 基于平均绝对误差提出一种用于衡量组内用户偏好与群组推荐结果偏离程度的指标 MAE^u , 计算方式如式 (49) 所示:

$$MAE^u = \frac{\sum_{i=1}^N |d^{ui} - d^{Gi}|}{N} \quad (49)$$

式中, d^{ui} 是通过个性化推荐系统得到的用户 u 对项目 i 的偏好得分, d^{Gi} 是通过群组推荐系统得到的群组 G 对项目 i 的偏好得分, N 是用户总数. 通过计算个性化推荐系统和群组推荐系统得到的偏好得分之间的差异, 评估群组推荐结果与用户个性化偏好之间的一致性. 较低的 MAE 值表示群组推荐系统能够更好地满足用户的个性化需求.

这些评价指标综合考虑了推荐结果的准确性、全面性、排序质量以及个性化满意度, 为评估群组推荐算法的性能提供了全面的参考依据.

3.3 实验对比分析

本节旨在比较和分析不同算法在常用数据集上的性能表现, 对这些方法的性能进行对比分析, 有助于更好地理解不同算法的优势和局限性, 从而为群组推荐领域的研究和应用提供有益的参考. 表 3 介绍了 MovieLens 1M 数据集, 其中包含上百万条电影评分记录. 考虑到该数据集缺乏明确的群组信息, 本文采用文献 [47] 的方法, 在此基础上生成两

个群组推荐的数据集, 分别为数据集 MS 和数据集 MR. 在 MS 中, 群组的成员具有较高的相似度; 而在 MR 中, 群组的成员是随机挑选的.

以下是对相关比较方法的简要介绍:

1) NCF-Based^[69]: 这类方法以 NCF 为基本框架, 首先计算每个成员对项目的偏好表示, 然后根据预定义的聚合策略计算得到群组对项目的偏好表示, 例如平均策略 (记为 NCF-AVG)、最小痛苦策略 (记为 NCF-LM) 和最大满足策略 (记为 NCF-MS).

2) AGREE^[46]: 该方法利用注意力机制计算群组成员的权重, 然后通过加权求和的方式聚合成员的偏好表示以得到群组的偏好表示.

3) MoSAN^[47]: 该方法首先利用注意力机制建模群组内成员之间的影响力, 然后通过加权求和的方式融合其他群组成员的偏好表示得到成员在特定群组下的偏好表示, 最后采用直接加和的方式聚合成员偏好表示以得到群组的偏好表示.

4) GAME^[52]: 采用 GCN 从多视图角度学习用户和项目的特征表示, 通过聚合目标节点的一阶邻居节点得到群组的偏好表示和项目的特征表示.

5) GLIF^[49]: 该方法基于同质性理论, 采用注意力网络融合多种交互关系和全局相似群组, 得到群组的偏好表示和项目的特征表示.

6) KGAG^[53]: 该方法采用 GCN 学习用户和项目的特征表示, 从群组内其他成员和候选项目两个方面聚合得到群组的偏好表示.

7) HetGRec^[54]: 该方法采用异构图注意力网络学习群组和项目的表示, 并基于相似的群组, 挖掘出目标群组的全局偏好表示.

8) GroupIM^[55]: 该方法基于个性化交互预训练模型, 利用两个额外的互信息最大化的任务来学习用户和群组的偏好表示.

9) GBERT^[60]: 该方法利用 BERT 作为基本框架, 采用两阶段学习的方法学习用户的偏好表示, 通过设计基于影响力的正则化方法, 利用成员的交互记录指导群组偏好的表示学习.

表 4 提供了不同表示层算法在三个常见的持续性群组数据集上的推荐效果. 观察结果显示, 相比于基于 NCF 的三种预定义聚合策略方法 (NCF-AVG、NCF-LM 和 NCF-MS), 基于注意力机制 (如 AGREE、MoSAN 和 GLIF) 和基于图神经网络 (如 GAME、KGAG 和 HetGRec) 的方法在三个数据集上表现更为优秀. 这表明启发式方法无法有效地捕捉成员在群组中的影响力, 因而其表现较差. 而注意力机制表现较好的原因可能在于其能够学习每个成员的影响力, 并更好地区分不同成员的作用. 另外, 基于图神经网络的方法能够取得最佳结果,

表 4 不同表示层算法在三个常见的持续性群组数据集上的推荐效果 (%)
Table 4 The recommendation performance of different presentation layer algorithms on three common persistent group datasets (%)

方法	数据集											
	CAMRa2011				MS				MR			
	H@5	H@10	N@5	N@10	H@5	H@10	N@5	N@10	H@5	H@10	N@5	N@10
NCF-AVG	58.33	77.65	39.69	46.25	59.19	83.15	47.35	52.21	63.52	78.42	45.32	50.29
NCF-LM	57.14	77.13	39.63	45.81	63.31	81.07	45.92	51.19	63.32	78.46	45.18	50.03
NCF-MS	57.19	75.12	38.50	44.41	64.43	82.25	46.62	51.98	62.35	77.85	44.43	49.02
AGREE	58.50	77.93	40.25	46.62	65.96	83.23	47.33	52.94	64.10	79.01	45.76	50.69
MoSAN	58.73	77.51	40.24	46.31	66.41	81.77	47.02	51.63	65.21	79.75	45.23	50.54
GAME	59.09	78.64	40.23	46.70	65.97	83.22	48.38	53.25	65.55	79.32	46.41	50.10
GLIF	59.18	78.93	40.30	46.73	66.43	83.55	48.20	53.44	65.61	79.93	46.43	51.07
KGAG	59.83	79.83	40.35	47.01	66.41	83.55	49.03	54.01	65.80	79.99	46.63	51.39
HetGRec	62.31	81.95	42.33	48.90	68.32	86.15	50.24	55.39	68.01	82.20	48.32	53.39

这一结果在一定程度上验证了图注意力算子在群组推荐任务中的有效性. 通过图神经网络进行表示学习不仅能够自适应地学习成员的影响力, 还能够有效地抽取高阶交互关系中更丰富的语义信息, 从而获取更为鲁棒的群组和项目表示.

表 5 列出了不同表示层算法在三个常见的临时性群组数据集上的推荐效果. 观察发现, 针对临时性群组, 持续性群组推荐算法 (如 NCF-AVG、NCF-LM、NCF-MS、AGREE 和 MoSAN) 与临时性群组推荐算法 (如 GroupIM 和 GBERT) 相比, 表现差距明显. 这一结果可能是由于临时性群组交互的高度稀疏性导致上述算法失效. 值得关注的是, 尽管基于图神经网络的方法 (如 GAME 和 KGAG) 并非针对临时性群组推荐任务设计, 但其表现与专门用于临时性群组的推荐算法 GroupIM 非常接近. 这说明通过图神经网络建模高阶交互关系的方式在一定程度上也适用于临时性群组推荐. 在此背景下,

GroupIM 表现良好的原因在于: 1) 采用互信息最大化任务以确保群组成员偏好的高一致性; 2) 在训练时引入个人标签数据, 将群组偏好约束在个人偏好附近, 提高学习效率. 尽管 GroupIM 已取得相当大的进展, 但 GBERT 在三个数据集上的表现最佳, 原因在于: 1) 采用表现更强大的 BERT 作为用户编码器, 能够捕获用户在不同群组中的偏好偏移; 2) 设计了三个用于群组推荐的任务, 能够学习到更好的成员表示; 3) 与 GroupIM 类似, GBERT 引入了个人标签, 约束成员和群组的表示. 不同的是, GroupIM 中的约束因子仅与用户相关, 而 GBERT 的约束因子涉及用户和候选项目, 因此更为有效.

4 群组推荐的研究难点

4.1 数据的稀疏性

推荐系统通常通过 (用户-项目) 交互数据进

表 5 不同算法在三个常见的临时性群组数据集上的推荐效果 (%)
Table 5 The recommendation performance of different algorithms on three common temporary group datasets (%)

方法	数据集											
	Weeplaces				Yelp				Douban			
	R@5	R@10	N@5	N@10	R@5	R@10	N@5	N@10	R@5	R@10	N@5	N@10
NCF-AVG	20.91	29.56	11.06	12.90	21.84	29.14	15.08	16.43	35.33	43.23	22.98	24.70
NCF-LM	20.32	28.33	10.49	12.19	23.22	31.44	16.04	17.20	44.29	49.56	31.91	33.10
NCF-MS	19.75	28.72	10.74	12.65	21.38	28.22	14.50	15.08	35.36	42.10	23.04	24.51
AGREE	20.53	29.09	11.40	13.22	24.16	30.98	16.80	17.63	45.95	51.22	33.39	34.57
MoSAN	31.81	37.71	26.25	28.15	46.57	50.61	34.66	36.18	47.10	52.22	36.12	37.24
GAME	41.97	48.53	28.90	30.35	46.44	51.94	35.32	36.52	58.76	77.52	40.29	46.33
KGAG	41.50	48.42	28.96	30.54	46.35	51.87	35.23	36.47	58.64	77.49	40.25	46.29
GroupIM	41.98	48.53	30.35	31.31	48.40	52.39	35.78	36.39	63.54	78.44	45.93	52.19
GBERT	49.43	52.82	35.31	36.43	48.67	53.14	37.46	38.11	65.20	79.90	47.22	54.58

行训练,然而这些数据往往呈现出高度稀疏的特点,即用户与项目之间的交互数据十分有限.在基于深度学习的群组推荐系统中,这种稀疏性导致训练数据不足,难以捕捉群组和项目之间的潜在关系,进而导致系统难以准确评估新群组的兴趣偏好和新项目的特征表示,无法基于历史数据做出精准的推荐.因此,在深度学习群组推荐系统中,有效的特征表示学习成为解决稀疏性问题的关键.

改善深度学习模型的性能需要设计合适的嵌入层或编码方法,以从用户和项目的特征中提取有用的信息并嵌入到模型中.例如,利用预训练的深度学习模型,如 GPT^[58] 或 BERT^[57] 等,将知识迁移到群组推荐系统中,有助于系统更好地理解用户和项目之间的关系,进而改善数据稀疏性问题^[74].这一策略为系统提供了更丰富的特征表示,有助于提高模型的性能.除了特征表示学习外,还可以通过整合不同来源的数据,如社交网络信息、评论、评分等,采用多源数据融合的方法来克服稀疏性问题.这种综合性策略为系统提供更全面、多维度的信息,弥补了单一数据源的不足,提高了系统对用户和项目关系的理解能力.此外,在群组推荐系统中,通过迁移学习实现共享知识、特征或模型参数,有助于解决冷启动问题.例如, Jafri 等^[75] 提出利用多模态数据嵌入和深度迁移学习的多阶段模型,来实现个性化的产品推荐. Gan 等^[76] 提出一种基于深度迁移学习的推荐模型 DeepTransferR, 它从多个用户活动中进行知识迁移.这些方法利用迁移学习使不同领域的推荐系统共享知识和模型参数,以提高模型的泛化能力.

综上所述,通过合理设计模型结构、特征表示学习、多源数据融合和迁移学习等方式,可以使系统更具适应性和准确性.这些方法为克服稀疏性问题提供了一系列可行的方案,为深度学习在群组推荐领域的应用提供了有益的启示.

4.2 推荐的可解释性

可解释性推荐系统的发展旨在增强群组推荐的透明度,使用户能够理解推荐背后的原因,从而提高其认可度和接受度.这种方法不仅有助于群组成员更迅速地做出决策,还提升了系统的整体易用性.然而,基于深度学习的群组推荐系统往往被视为黑盒模型,其内部工作原理难以理解和解释,引发了一系列问题,包括用户对推荐结果的不信任、推荐理由的难以解释,以及难以满足监管和法规的要求.

为了解决深度学习模型可解释性的问题,一种常见的方法是设计更可解释的深度学习模型结构.

例如,采用浅层神经网络和引入可解释性的特征工程,或将深度学习模型与传统推荐算法相结合,以兼顾二者的优势.深度学习模型可以捕捉更复杂的关系,而传统算法通常更易于解释.通过结合这两种方法,可以在提供准确推荐的同时提高系统的可解释性.除了模型层面的改进外,创建用户友好的界面也是提高可解释性的重要途径.通过可视化工具或自然语言解释,向用户清晰呈现推荐结果的理由,有助于用户更好地理解系统的决策过程.

4.3 推荐的公平性

推荐系统的公平性是指系统在决策和产生输出时能够确保结果公正、无偏、非歧视.在群组推荐中,公平性主要包括个体公平性、群体公平性和多样公平性.个体公平性强调系统应该尊重每一位用户的兴趣和需求,而不只是满足少数活跃用户的期望;群体公平性关注系统对待不同用户群体的方式,避免在推荐中偏袒某一特定群体;多样公平性要求系统提供多样化的推荐,避免用户陷入信息茧房,只看到与其观点一致的内容.在推荐系统中存在一些不公平现象,其中包括训练数据的偏见^[77-78],导致模型学习到的结果存在非公平性.为了解决这些问题,需要深入分析不公平推荐的原因,并提出相应的解决方案.

为确保推荐结果的公平性,有一些有效的方法可供采用.首先,采集均衡的数据至关重要,确保训练数据集涵盖各种不同类型的用户和项目,以减少偏见的影响.这需要整合来自多个来源的数据,以确保各类用户和项目都得到足够的代表性.其次,过采样和欠采样是解决不平衡数据的有效手段.通过过采样(增加少数类样本)或欠采样(减少多数类样本),可以平衡数据分布,减少对某些群体的过度侧重.这可以通过随机采样或使用某些算法,如 SMOTE^[79] 来实现.未来的研究应更深入地分析导致不公平推荐的根本原因,并提出更具体、有效的解决方案.除了数据层面的改进,还可以考虑模型设计、算法优化等方面的创新,以进一步提高推荐系统的公平性.

5 群组推荐的前景展望

5.1 融合群决策机制的推荐

群决策理论在群组决策中引入了共识和影响力两个核心概念,这与群组推荐系统的目标密切相关^[80].共识被视为群决策中的关键因素,直接影响着决策结果的满意度和公平性,与群组推荐系统的核心理

念一脉相承;影响力则指在群体中对其他成员观点的影响程度.群决策机制旨在实现群体成员在决策中相互影响,促使观点趋于一致.群决策理论认为,群体成员在决策过程中相互交流观点,不断调整自身观点,最终达成共识.

基于这一理念,Castro等^[81]提出了一个基于共识达成的群组推荐系统,通过成员对候选项目的评分和共识达成过程生成群组的推荐列表.类似地,Garcia等^[82]设计了一个基于反馈机制的协商框架,用于达成最终推荐列表的共识.另外,Mandal等^[80]通过共识函数建立每个专家或成员的共识关系,并通过反馈机制调整专家的偏好.这些研究为群组推荐中的共识达成提供了重要基础,通过群决策机制的信息交互与融合,解决了群组推荐领域的冲突和用户满意度提升等问题.

群决策机制在群组推荐中的应用取得了一系列积极效果.首先,它解决了冲突的问题,通过协调不同用户之间的冲突,找到适合多数人的共同兴趣或建议.其次,它提升了用户满意度,系统更好地满足多数用户的需求,从而提高了整体用户的满意度,增强了群组推荐系统的可用性和用户接受度.尽管群决策机制在群组推荐中取得了显著成果,但仍有一些潜在的研究方向值得深入探讨.首先,需要进一步研究共识达成的机制,以更好地理解 and 优化这一过程.其次,可以考虑结合深度学习等技术,提高群决策机制的智能化水平.此外,针对不同应用场景的群组推荐系统,定制化的群决策机制也是未来研究的一个重要方向.

5.2 群组推荐中知识与数据的迭代进化

在群组推荐系统中,用户的需求和偏好随时间和情境不断变化,因此系统需要持续更新和优化数据与推荐模型,进行知识与数据的迭代进化.Xu等^[83]提出一种基于图神经网络的增量更新推荐模型方法,通过图神经网络的结构,实现了对用户兴趣的动态捕捉和模型的高效增量更新.这一方法显著减少了计算时间,使系统能够更频繁地更新模型,以应对用户兴趣的动态变化.另外,Song等^[84]采用在线学习方法,根据用户的实时点击行为,在线学习用户的物品偏好,通过构建在线物品集群树实现了实时推荐.通过采用在线学习等方法,模型能够根据用户的实时反馈不断调整推荐策略,以更好地适应兴趣的漂移.这些方法为推荐系统提供了更灵活、实时的用户建模方法,提高了推荐的准确性,并有效地解决了用户兴趣漂移的问题^[85],同时也提升了模型更新的效率.

尽管知识与数据的迭代进化方法在群组推荐系统中取得了显著进展,但仍然存在一些挑战.首先,如何更好地平衡推荐模型的精度与计算效率仍然是一个开放性问题.其次,对于不同领域的推荐系统,需要探索更灵活、可扩展的迭代进化方法.未来的研究可以进一步深化对知识与数据迭代进化的理论认识,并结合深度学习等技术,提高迭代进化的智能化水平.

5.3 大型语言模型对群组推荐的影响

大型语言模型(Large language model, LLM)通过广泛应用大规模预训练语料和自监督训练,显著提升了通用语义理解和生成能力.这种增强的能力使得LLM在泛化和对知识的深入理解方面具备了更为卓越的表现.LLM不仅能够更好地理解跨领域的语义关系,而且通过引入外部的开放知识,其语义信号更加丰富.在群组推荐的应用中,LLM的引入将带来深远的影响^[86].首先,LLM能够实现对用户和项目更深层次的理解,通过学习大规模语料库,模型获得了对丰富语境和多样性信息的高度感知^[87].其次,LLM具备卓越的泛化能力,能够在不同领域之间进行知识迁移,从而为跨领域的群组推荐提供新的可能性^[88].在群组推荐系统中,LLM的数据扩充能力也是一项重要的优势.通过利用LLM在大量无监督数据上的学习,系统能够获得更为全面和多样的数据表示,从而提高模型对用户和项目的理解能力^[89],这为群组推荐系统的性能提供了新的提升途径.

将LLM与群组推荐模型相结合,可以实现对用户和内容的更全面理解.这种整合包括用户理解、内容理解、样本扩充以及打分排序等多个方面,为群组推荐系统注入了更多的智能和个性化元素.通过充分发挥LLM在信息抽象和表示学习方面的优势,群组推荐系统能够更准确地捕捉群组偏好和项目特征,从而提供更为个性化和精准的推荐服务.总体而言,LLM的引入为群组推荐系统带来了全新的发展机遇.通过充分挖掘LLM的泛化能力、跨域知识能力以及数据扩充能力,群组推荐系统不仅能够提高用户体验,还能够在推荐效果上取得更为显著的提升.

6 结束语

本文首先介绍群组推荐系统的背景知识,并按照数据输入层、表示学习层和模型层对目前基于深度学习的群组推荐算法进行分类介绍.在表示学习层方法中,根据偏好获取方法的不同,将其分为基于启发式聚合策略、基于概率模型、基于注意力机

制、基于图神经网络和面向临时性群组的偏好表示方法, 并对各类方法进行深入的探讨. 此外, 本文系统性地整理群组推荐的数据集, 以便于对群组推荐系统领域进行实验评估. 通过对基于深度学习的群组推荐算法进行全面调研与综述, 进一步分析该领域的研究难点, 包括数据的稀疏性、深度学习的可解释性和推荐的公平性等问题, 这些问题使得群组推荐系统依然面临着诸多技术挑战. 针对这些难点, 本文结合群决策理论和大语言模型等技术, 提出了未来有价值的研究方向.

References

- Tang Jia-Rui, Tian Ran, Liang Dan-Hui. In-depth report on the E-commerce industry: Deepening online consumption mindset and actively exploring new consumption scenarios [Online], available: <https://baijiahao.baidu.com/s?id=1760070652649669963&wfr=spider&for=pc>, March 11, 2023 (唐佳睿, 田然, 梁丹辉. 电商行业深度报告: 线上消费心智加深, 积极探索新消费场景 [Online], available: <https://baijiahao.baidu.com/s?id=1760070652649669963&wfr=spider&for=pc>, 2023-03-11)
- Zhang Hong-Lei, Li Yi-Dong, Wu Jun, Chen Nai-Yue, Dong Hai-Rong. A survey on privacy-preserving federated recommender systems. *Acta Automatica Sinica*, 2022, **48**(9): 2142–2163 (张洪磊, 李滢东, 邬俊, 陈乃月, 董海荣. 基于隐私保护的联邦推荐算法综述. *自动化学报*, 2022, **48**(9): 2142–2163)
- Yang Y, Zhang C B, Song X, Dong Z, Zhu H S, Li W J. Contextualized knowledge graph embedding for explainable talent training course recommendation. *ACM Transactions on Information Systems*, 2024, **42**(2): Article No. 33
- Cheng M Y, Liu Q, Zhang W Y, Liu Z D, Zhao H K, Chen E H. A general tail item representation enhancement framework for sequential recommendation. *Frontiers of Computer Science*, 2024, **18**(6): Article No. 186333
- Hao B W, Yin H Z, Li C P, Chen H. Self-supervised graph learning for occasional group recommendation. *International Journal of Intelligent Systems*, 2022, **37**(12): 10880–10902
- McCarthy J F, Anagnost T D. MusicFX: An arbiter of group preferences for computer supported collaborative workouts. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work. Seattle Washington, USA: ACM, 1998. 363–372
- Xu Xiao-Ming, Mei Hong-Yan, Yu Heng, Li Xiao-Hui. Review of group recommendation methods based on preference fusion. *Journal of Chinese Computer Systems*, 2020, **41**(12): 2500–2508 (许晓明, 梅红岩, 于恒, 李晓会. 基于偏好融合的群组推荐方法研究综述. *小型微型计算机系统*, 2020, **41**(12): 2500–2508)
- Zhang Yu-Jie, Du Yu-Lu, Meng Xiang-Wu. Research on group recommender systems and their applications. *Chinese Journal of Computers*, 2016, **39**(4): 745–764 (张玉洁, 杜雨露, 孟祥武. 组推荐系统及其应用研究. *计算机学报*, 2016, **39**(4): 745–764)
- Zhao Hai-Yan, Cheng Ruo-Yao, Chen Qing-Kui, Cao Jian. Group recommender system: Current status and future trends. *Journal of Chinese Computer Systems*, 2021, **42**(6): 1144–1151 (赵海燕, 成若瑶, 陈庆奎, 曹健. 群组推荐系统: 现状与展望. *小型微型计算机系统*, 2021, **42**(6): 1144–1151)
- Dara S, Chowdary C R, Kumar C. A survey on group recommender systems. *Journal of Intelligent Information Systems*, 2020, **54**(2): 271–295
- Kompan M, Bielikova M. Group recommendations: Survey and perspectives. *Computing and Informatics*, 2014, **33**(2): 446–476
- O'Connor M, Cosley D, Konstan J A, Riedl J. PolyLens: A recommender system for groups of users. *ECSCW 2001*. Dordrecht: Springer, 2001. 199–218
- Wu X X, Xiong Y, Zhang Y, Jiao Y Z, Zhang J W, Zhu Y Y, et al. ConsRec: Learning consensus behind interactions for group recommendation. In: Proceedings of the ACM Web Conference. Austin, USA: ACM, 2023. 240–250
- Zhu R T, Lv D T, Yu Y, Zhu R H, Zheng Z Z, Bu K, et al. LINet: A location and intention-aware neural network for hotel group recommendation. In: Proceedings of the ACM Web Conference. Austin, USA: ACM, 2023. 779–789
- Crossen A, Budzik J, Hammond K J. Flytrap: Intelligent group music recommendation. In: Proceedings of the 7th International Conference on Intelligent User Interfaces. San Francisco, USA: ACM, 2002. 184–185
- Carroll J M, Rosson M B. Paradox of the active user. In: Proceedings of the Interfacing Thought: Cognitive Aspects of Human-computer Interaction. Cambridge, USA: MIT Press, 1987. 80–111
- Zheng N, Li Q D, Liao S C, Zhang L M. Which photo groups should I choose? A comparative study of recommendation algorithms in Flickr. *Journal of Information Science*, 2010, **36**(6): 733–750
- Liu X J, Tian Y, Ye M, Lee W C. Exploring personal impact for group recommendation. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management. Maui, USA: ACM, 2012. 674–683
- Yuan Q, Cong G, Lin C Y. COM: A generative model for group recommendation. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, USA: ACM, 2014. 163–172
- Wang Q, Mao Z D, Wang B, Guo L. Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering*, 2017, **29**(12): 2724–2743
- Chen J W, Wu J K, Wu J C, Cao X Z, Zhou S, He X N. Adaptor: Adaptively modulating embedding magnitude for recommendation. In: Proceedings of the ACM Web Conference. Austin, USA: ACM, 2023. 1085–1096
- Dyer J S, Sarin R K. Group preference aggregation rules based on strength of preference. *Management Science*, 1979, **25**(9): 822–832
- Wu Yun-Chang, Liu Bai-Song, Wang Yang-Yang, Fei Chen-Jie. Review of group recommendation analysis and research. *Telecommunications Science*, 2018, **34**(12): 71–83 (吴云昌, 刘柏嵩, 王洋洋, 费晨杰. 群组推荐分析与研究综述. *电信科学*, 2018, **34**(12): 71–83)
- Garcia I, Pajares S, Sebastia L, Onaindia E. Preference elicitation techniques for group recommender systems. *Information Sciences*, 2012, **189**: 155–175
- Masthoff J. Group recommender systems: Combining individual models. *Recommender Systems Handbook*. New York: Springer, 2010. 677–702
- Agarwal A, Chakraborty M, Chowdary C R. Does order matter? Effect of order in group recommendation. *Expert Systems With Applications*, 2017, **82**: 115–127
- Baltrunas L, Makcinskas T, Ricci F. Group recommendations with rank aggregation and collaborative filtering. In: Proceedings of the 4th ACM Conference on Recommender Systems. Barcelona, Spain: ACM, 2010. 119–126
- Yu Z W, Zhou X S, Hao Y B, Gu J H. TV program recommendation for multiple viewers based on user profile merging. *User Modeling and User-Adapted Interaction*, 2006, **16**(1): 63–82
- Xuan P C, Tang Y, Wang W S. Group recommendation method based on item type proportion factor. In: Proceedings of the IEEE/ACM International Conference on Utility and Cloud Computing Companion. Zurich, Switzerland: IEEE, 2018. 376–381
- Mao Yu-Jia, Liu Xue-Jun, Xu Xin-Yan, Zhang Xin. Method based on sub-group and social behavior for narrowing recommended list for groups. *Computer Science*, 2019, **46**(2): 202–209 (毛宇佳, 刘学军, 徐新艳, 张欣. 基于子组与社会行为的缩小群组

- 推荐列表方法. 计算机科学, 2019, 46(2): 202–209
- 31 Huang Z H, Xu X, Zhu H H, Zhou M C. An efficient group recommendation model with multiattention-based neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(11): 4461–4474
- 32 Yin H Z, Wang Q Y, Zheng K, Li Z X, Yang J L, Zhou X F. Social influence-based group representation learning for group recommendation. In: Proceedings of the 35th International Conference on Data Engineering. Macao, China: IEEE, 2019. 566–577
- 33 Jameson A. More than the sum of its members: Challenges for group recommender systems. In: Proceedings of the Working Conference on Advanced Visual Interfaces. Gallipoli, Italy: ACM, 2004. 48–54
- 34 Du J, Li L, Gu P, Xie Q. A group recommendation approach based on neural network collaborative filtering. In: Proceedings of the 35th International Conference on Data Engineering Workshops. Macao, China: IEEE, 2019. 148–154
- 35 Quintarelli E, Rabosio E, Tanca L. Recommending new items to ephemeral groups using contextual user influence. In: Proceedings of the 10th ACM Conference on Recommender Systems. Boston, USA: USA, 2016. 285–292
- 36 Quijano-Sanchez L, Recio-Garcia J A, Diaz-Agudo B. Happy-movie: A facebook application for recommending movies to groups. In: Proceedings of the 23rd International Conference on Tools with Artificial Intelligence. Boca Raton, USA: IEEE, 2011. 239–244
- 37 Quijano-Sánchez L, Recio-García J A, Díaz-Agudo B. Personality and social trust in group recommendations. In: Proceedings of the 22nd IEEE International Conference on Tools With Artificial Intelligence. Arras, France: IEEE, 2010. 121–126
- 38 Quijano-Sanchez L, Recio-Garcia J A, Diaz-Agudo B, Jimenez-Diaz G. Social factors in group recommender systems. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2013, 4(1): Article No. 8
- 39 Hu L, Cao J, Xu G D, Cao L B, Gu Z P, Cao W. Deep modeling of group preferences for group-based recommendation. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence. Québec City, Canada: AAAI, 2014. 1861–1867
- 40 Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. In: Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA: ICLR, 2015. 1–15
- 41 Pan X R, Ge C J, Lu R, Song S J, Chen G F, Huang Z Y, et al. On the integration of self-attention and convolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE, 2022. 805–815
- 42 Xu H T, Zhang J, Dai L R. Differential time-frequency log-mel spectrogram features for vision transformer based infant cry recognition. In: Proceedings of the Interspeech. Incheon, South Korea: ISCA, 2022. 1963–1967
- 43 Ma Z Y, Li J J, Li G H, Cheng Y J. UniTranSeR: A unified transformer semantic representation framework for multimodal task-oriented dialog system. In: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics. Dublin, Ireland: Association for Computational Linguistics, 2022. 103–114
- 44 Zheng Jian-Xing, Li Qin-Wen, Wang Su-Ge, Li De-Yu. Research on explainable rating prediction by fusing attribute preference and multi-order interaction information. *Acta Automatica Sinica*, 2024, 50(11): 2231–2244
(郑建兴, 李沁文, 王素格, 李德玉. 融合属性偏好和多阶交互信息的可解释评分预测研究. 自动化学报, 2024, 50(11): 2231–2244)
- 45 Tang Wen-Bing, Ren Zheng-Yun, Han Fang. Attention-based collaborative convolutional dynamic network for recommendation. *Acta Automatica Sinica*, 2021, 47(10): 2438–2448
(汤文兵, 任正云, 韩芳. 基于注意力机制的协同卷积动态推荐网络. 自动化学报, 2021, 47(10): 2438–2448)
- 46 Cao D, He X N, Miao L H, An Y H, Yang C, Hong R C. Attentive group recommendation. In: Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval. Ann Arbor, USA: ACM, 2018. 645–654
- 47 Tran L V, Pham T A N, Tay Y, Liu Y D, Cong G, Li X L. Interact and decide: Medley of sub-attention networks for effective group recommendation. In: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. Paris, France: ACM, 2019. 255–264
- 48 He Z X, Chow C Y, Zhang J D, Li N. GRADI: Towards group recommendation using attentive dual top-down and bottom-up influences. In: Proceedings of the IEEE International Conference on Big Data. Los Angeles, USA: IEEE, 2019. 631–636
- 49 Zhang S, Zheng N, Wang D L. A novel attention-based global and local information fusion neural network for group recommendation. *Machine Intelligence Research*, 2022, 19(4): 331–346
- 50 McPherson M, Smith-Lovin L, Cook J M. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 2001, 27: 415–444
- 51 Liang R X, Zhang Q, Wang J Q, Lu J. A hierarchical attention network for cross-domain group recommendation. *IEEE Transactions on Neural Networks and Learning Systems*, 2024, 35(3): 3859–3873
- 52 He Z X, Chow C Y, Zhang J D. GAME: Learning graphical and attentive multi-view embeddings for occasional group recommendation. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2020. 649–658
- 53 Deng Z Y, Li C Y, Liu S J, Ali W, Shao J. Knowledge-aware group representation learning for group recommendation. In: Proceedings of the 37th International Conference on Data Engineering. Chania, Greece: IEEE, 2021. 1571–1582
- 54 Zhang S, Zheng N, Wang D L. HetGRec: Heterogeneous graph attention network for group recommendation. *IEEE Intelligent Systems*, 2023, 38(1): 9–18
- 55 Sankar A, Wu Y H, Wu Y H, Zhang W, Yang H, Sundaram H. Groupim: A mutual information maximization framework for neural group recommendation. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2020. 1279–1288
- 56 Veličković P, Fedus W, Hamilton W L, Liò P, Bengio Y, Hjelm R D. Deep graph infomax. In: Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA: OpenReview.net, 2019. 1–17
- 57 Devlin J, Chang M W, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, Minnesota: Association for Computational Linguistics, 2019. 4171–4186
- 58 Radford A, Narasimhan K, Salimans T, Sutskever H. Improving language understanding by generative pre-training [Online], available: <https://api.semanticscholar.org/CorpusID:49313245>, December 18, 2023
- 59 Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I. Language models are unsupervised multitask learners. *OpenAI Blog*, 2019, 1(8): Article No. 9
- 60 Zhang S, Zheng N, Wang D L. GBERT: Pre-training user representations for ephemeral group recommendation. In: Proceedings of the 31st ACM International Conference on Information and Knowledge Management. Atlanta, USA: ACM, 2022. 2631–2639
- 61 Zhang Y Y, Wu X, Fang Q, Qian S S, Xu C S. Knowledge-enhanced attributed multi-task learning for medicine recommendation. *ACM Transactions on Information Systems*, 2023, 41(1): Article No. 17
- 62 Cao D, He X N, Miao L H, Xiao G Y, Chen H, Xu J. Social-enhanced attentive group recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 33(3): 1195–1209

- 63 Yin H Z, Wang Q Y, Zheng K, Li Z X, Zhou X F. Overcoming data sparsity in group recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 2022, **34**(7): 3447–3460
- 64 Chen H X, Yin H Z, Chen T, Nguyen Q V H, Peng W C, Li X. Exploiting centrality information with graph convolutions for network representation learning. In: Proceedings of the 35th International Conference on Data Engineering. Macao, China: IEEE, 2019. 590–601
- 65 He X N, Liao L Z, Zhang H W, Nie L Q, Hu X, Chua T S. Neural collaborative filtering. In: Proceedings of the 26th International Conference on World Wide Web. Perth, Australia: International World Wide Web Conferences Steering Committee, 2017. 173–182
- 66 Feng S S, Zhang H X, Wang L, Liu L, Xu Y C. Detecting the latent associations hidden in multi-source information for better group recommendation. *Knowledge-Based Systems*, 2019, **171**: 56–68
- 67 Choudhary N, Minz S, Bharadwaj K K. Negotiation framework for group recommendation based on fuzzy computational model of trust and distrust. *Multimedia Tools and Applications*, 2020, **79**(37–38): 27337–27364
- 68 Guo L, Yin H Z, Wang Q Y, Cui B, Huang Z, Cui L Z. Group recommendation with latent voting mechanism. In: Proceedings of the 36th International Conference on Data Engineering. Dallas, USA: IEEE, 2020. 121–132
- 69 Guo L, Yin H Z, Chen T, Zhang X L, Zheng K. Hierarchical hyperedge embedding-based representation learning for group recommendation. *ACM Transactions on Information Systems (TOIS)*, 2021, **40**(1): Article No. 3
- 70 Seo Y D, Kim Y G, Lee E, Seol K S, Baik D K. An enhanced aggregation method considering deviations for a group recommendation. *Expert Systems With Applications*, 2018, **93**: 299–312
- 71 Loveymi S, Hamzeh A. Proposing an evolutionary method based on maximization of precision of group recommender systems. In: Proceedings of the 7th Conference on Information and Knowledge Technology. Urmia, Iran: IEEE, 2015. 1–6
- 72 Guo Z W, Tang C W, Tang H, Fu Y Q, Niu W J. A novel group recommendation mechanism from the perspective of preference distribution. *IEEE Access*, 2018, **6**: 5865–5878
- 73 Sun Y Z, Han J W, Yan X F, Yu P S, Wu T Y. PathSim: Meta path-based top-k similarity search in heterogeneous information networks. *Proceedings of the VLDB Endowment*, 2011, **4**(11): 992–1003
- 74 Zeng Z N, Xiao C J, Yao Y, Xie R B, Liu Z Y, Lin F, et al. Knowledge transfer via pre-training for recommendation: A review and prospect. *Frontiers in Big Data*, 2021, **4**: Article No. 602071
- 75 Jafri S I H, Ghazali R, Javid I, Mahmood Z, Hassan A A A. Deep transfer learning with multimodal embedding to tackle cold-start and sparsity issues in recommendation system. *PLoS One*, 2022, **17**(8): 1–24
- 76 Gan M X, Ma Y X. Knowledge transfer learning from multiple user activities to improve personalized recommendation. *Soft Computing*, 2022, **26**(14): 6547–6566
- 77 Chen J W, Dong H D, Qiu Y, He X N, Xin X, Chen L, et al. AutoDebias: Learning to debias for recommendation. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2021. 21–30
- 78 Chen J W, Dong H D, Wang X, Feng F L, Wang M, He X N. Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems*, 2023, **41**(3): Article No. 67
- 79 Chawla N V, Bowyer K W, Hall L O, Kegelmeyer W P. SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 2002, **16**(1): 321–357
- 80 Mandal P, Samanta S, Pal M, Ranadive A S. Pythagorean linguistic preference relations and their applications to group decision making using group recommendations based on consistency matrices and feedback mechanism. *International Journal of Intelligent Systems*, 2020, **35**(5): 826–849
- 81 Castro J, Quesada F J, Palomares I, Martínez L. A consensus-driven group recommender system. *International Journal of Intelligent Systems*, 2015, **30**(8): 887–906
- 82 Garcia I, Sebastia L. A negotiation framework for heterogeneous group recommendation. *Expert Systems With Applications*, 2014, **41**(4): 1245–1261
- 83 Xu Y S, Zhang Y X, Guo W, Guo H F, Tang R M, Coates M. GraphSAIL: Graph structure aware incremental learning for recommender systems. In: Proceedings of the 29th ACM International Conference on Information and Knowledge Management. ACM, 2020. 2861–2868
- 84 Song L Q, Tekin C, van der Schaar M. Online learning in large-scale contextual recommender systems. *IEEE Transactions on Services Computing*, 2016, **9**(3): 433–445
- 85 Zheng N, Li Q D. A recommender system based on tag and time information for social tagging systems. *Expert Systems With Applications*, 2011, **38**(4): 4575–4587
- 86 Lin J H, Dai X Y, Xi Y J, Liu W W, Chen B, Li X Y, et al. How can recommender systems benefit from large language models: A survey. *ACM Transactions on Information Systems*, DOI: 10.1145/3678004
- 87 Wu L K, Zheng Z, Qiu Z P, Wang H, Gu H C, Shen T J, et al. A survey on large language models for recommendation. *World Wide Web*, 2024, **27**(5): Article No. 60
- 88 Li L, Zhang Y F, Liu D G, Chen L. Large language models for generative recommendation: A survey and visionary discussions. In: Proceedings of the Joint International Conference on Computational Linguistics, Language Resources and Evaluation. Torino, Italia: ELRA and ICCL, 2023. 10146–10159
- 89 Qiu Z P, Wu X, Gao J Y, Fan W. U-BERT: Pre-training user representations for improved recommendation. In: Proceedings of the 35th AAAI Conference on Artificial Intelligence. AAAI, 2021. 4320–4327



郑楠 中国科学院自动化研究所多模态人工智能系统全国重点实验室副研究员。主要研究方向为复杂系统, 系统集成, 数据挖掘, 个性化推荐。

E-mail: nan.zheng@ia.ac.cn

(ZHENG Nan Associate professor at the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences. Her research interest covers complex systems, metasynthesis, data mining, and personalized recommendations.)



章颂 中国科学院自动化研究所多模态人工智能系统全国重点实验室博士研究生。主要研究方向为复杂系统, 系统集成, 自然语言推理, 推荐系统。

E-mail: zhangsong2022@ia.ac.cn

(ZHANG Song Ph.D. candidate at the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers complex systems, metasynthesis, natural language reasoning, and recommender system.)



刘玉桥 中国科学院自动化研究所多模态人工智能系统国家重点实验室硕士研究生. 主要研究方向为数据挖掘, 推荐系统, 自然语言处理.

E-mail: liuyuqiao2022@ia.ac.cn

(LIU Yu-Qiao Master student at the State Key Laboratory of Mul-

timodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers data mining, recommendation system, and natural language processing.)



王雨桐 中国科学院自动化研究所复杂系统管理与控制国家重点实验室助理研究员. 2021 年获得中国科学院大学控制理论与控制工程专业博士学位. 主要研究方向为计算机视觉.

E-mail: yutong.wang@ia.ac.cn

(WANG Yu-Tong Assistant pro-

fessor at the State Key Laboratory for Management

and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. She received her Ph.D. degree in control theory and control engineering from University of Chinese Academy of Sciences in 2021. Her main research interest is computer vision.)



王飞跃 中国科学院自动化研究所复杂系统管理与控制国家重点实验室研究员. 主要研究方向为智能系统和复杂系统的建模、分析与控制. 本文通信作者. E-mail: feiyue.wang@ia.ac.cn

(WANG Fei-Yue Professor at the State Key Laboratory for Manage-

ment and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers modeling, analysis, and control of intelligent systems and complex systems. Corresponding author of this paper.)