

基于多模态用户图的跨领域服务匹配方法

王海艳^{1,2}, 刘万字¹, 骆健^{1,2}, 张少聪¹

(1. 南京邮电大学计算机学院, 江苏南京210023; 2. 大数据安全与智能处理省高校重点实验室, 江苏南京210023)

摘要: 跨领域服务匹配场景中, 用户偏好分布于多个服务领域, 并呈现显著的多模态与异构特征, 准确刻画用户跨领域偏好是提升匹配性能的关键。针对现有方法普遍依赖统一嵌入空间, 易造成多源偏好压缩并忽略模态差异的问题, 本文提出了一种基于多模态用户图的跨领域服务匹配方法(multimodal user graph-based cross-domain service matching, MUG-CDSM)。该方法以偏好标签为锚点构建多模态用户图, 通过显式建模用户在文本、图像、音频和视频等模态下的偏好结构, 并引入开放知识图谱与虚拟辅助节点增强图连通性。在此基础上, 设计模态感知图池化模块 MUG POOL, 实现不同模态偏好的自适应聚合。实验结果表明, 所提出的方法在 Amazon 多领域多模态数据集上优于多种跨领域服务匹配基线模型。

关键词: 跨领域服务匹配; 多模态用户图; 模态感知图池化; 模态聚合

中图分类号: TP399

文献标志码: A

A cross-domain service matching method based on multimodal user graph

Wang Haiyan^{1,2}, Liu Wanyu¹, Luo Jian^{1,2}, Zhang Shaocong¹

1. School of Computer Science, Nanjing University of Post and Telecommunications, Nanjing 210023, China

2. Key Laboratory of Big Data Security & Intelligent Processing, Nanjing 210023, China

Abstract: In cross-domain service matching scenarios, user preferences are distributed across multiple service domains and exhibit pronounced multimodal and heterogeneous characteristics. Accurately modeling users' cross-domain preferences is therefore crucial for improving matching performance. Existing methods predominantly rely on unified embedding spaces, which tend to compress multi-source preferences and overlook modality-specific discrepancies. To address these limitations, we propose a Multimodal User Graph-based Cross-Domain Service Matching method. The proposed approach constructs a multimodal user graph using preference labels as anchors, explicitly modeling user preferences across text, image, audio, and video modalities, while incorporating an open knowledge graph and virtual auxiliary nodes to enhance graph connectivity. Furthermore, a modality-aware graph pooling module, termed MUG POOL, is designed to adaptively aggregate preferences across different modalities. Experimental results on the Amazon multimodal multi-domain dataset demonstrate that the proposed method outperforms state-of-the-art cross-domain service matching baselines.

Keywords: cross-domain service matching, multimodal user graph, modality-aware graph pooling, modality aggregation

0 引言

跨领域服务匹配(cross-domain service matching, CDSM)是利用用户在一个或多个源领域中积累

的行为数据, 推断其在目标领域中的偏好, 从而在目标域实现个性化的服务匹配。与传统的单领域服务匹配不同, CDSM 的目标是打破服务平台之间

收稿日期: 2026-04-03; 修回日期: 2026-04-04

基金项目: 国家自然科学基金资助项目(No. 62272243)

Foundation Items: The National Natural Science Foundation of China (No.62272243)

的数据壁垒, 实现用户偏好的跨平台迁移与共享, 以提升服务匹配在冷启动用户、稀疏数据环境下的服务质量。面对用户行为在多个领域中的分布不均和用户偏好的模态异构问题, 构建能够泛化到多个领域的用户表示是 CDSM 的核心挑战^[1]。

现有研究表明^[2-4], 跨领域服务匹配是缓解平台间信息孤岛、提升用户服务体验的重要技术手段, 其核心在于挖掘和迁移用户在不同服务领域中的潜在偏好关联。针对用户在多源服务平台中行为分布异构、数据稀疏等问题, 已有工作^[8-18]提出了多种跨域建模方法, 以支持用户偏好在不同领域间的共享与泛化。目前, 大多数跨领域服务匹配方法依赖共享或统一嵌入空间对用户在不同服务领域中的行为进行建模, 通过压缩多域信息实现跨域知识迁移。然而, 不同服务平台在内容模态构成、偏好表达方式上存在显著差异, 基于统一嵌入的建模方式难以刻画这种异构性, 往往导致多源偏好细粒度特征的丢失, 从而限制匹配精度。尤其在多模态服务场景中, 忽略模态差异的跨域表示容易引入语义混淆, 造成跨域偏好迁移偏离, 进而影响匹配效果。为缓解上述问题, 部分研究引入多模态特征辅助跨领域服务匹配, 但现有方法多将多模态信息作为附加特征进行简单融合^[5-7], 缺乏对多模态偏好之间结构化关系的系统建模, 难以有效刻画不同模态间的交互与互补特性, 同样很难考虑到跨域偏好的模态差异性。此外, 多模态建模在单一领域内仍易受到用户交互稀疏性的制约, 其性能提升空间有限。综上, 如何在跨领域场景下为目标领域补充稀缺的多模态行为信息, 同时充分考虑不同模态特征的差异性, 是实现更有效的跨平台偏好建模与服务匹配的关键挑战。

为应对上述跨领域服务匹配的挑战, 本文提出了一种基于多模态用户图(multimodal user graph, MUG)的跨领域服务匹配方法。该方法以用户在多个服务领域中的偏好标签为锚点, 系统性融合多模态信息, 构建模态一致性子图, 从结构层面刻画不同模态下的偏好特征, 避免多源信息混叠。进一步地, 引入开放知识图谱(open knowledge graph, OKG)补充标签之间的语义关系, 并通过虚拟辅助节点增强多模态用户图的整体连通性, 从而缓解跨领域场景下偏好分布稀疏问题。在此基础上, 设计了一种模态感知图池化模块 MUG POOL, 通过

“全局模态引导—局部子图注意力聚合”的双阶段策略, 实现模态级特征压缩与跨模态信息融合, 生成统一且充分考虑各个模态差异性的用户表示。最终, 结合所生成的统一用户表示, 构建多模态服务匹配模型, 通过模态感知的相似度计算与注意力加权机制, 对候选服务项进行评分与排序, 从而完成跨领域服务匹配决策。综上, 本文的主要工作包括:

1) 提出一种面向服务匹配的多模态用户建模方法, 通过构建多模态用户偏好图, 在以不同领域中的偏好标签为锚点, 融合文本、图像、音频与视频四类模态信息构建模态一致性子图的基础上, 引入开放知识图谱补充标签语义边, 同时使用虚拟辅助节点增强图结构的全局连通性。该建模方式不仅提升了用户在多领域、异构模态下的行为表达能力, 也为后续服务匹配提供了结构可解释、模态覆盖全面的偏好表示基础。

2) 设计模态感知图池化模块 MUG POOL, 将多模态用户图的结构建模与服务匹配任务有效连接。该机制在多模态异构图中实现了“全局模态引导+局部子图注意力聚合”的双阶段池化策略。通过 Group GCN 独立建模每种模态下的全局偏好向量, 再以其引导对多模态子图的模态级聚合, 实现模态间特征压缩与语义对齐, 最终生成统一用户表示, 用于服务匹配中与多模态服务项进行匹配计算。

3) 构建统一的多模态服务表示与匹配机制, 完成服务级别的跨模态对齐与偏好匹配任务。服务项被建模为结构化多模态实体, 通过模态特征提取与投影实现服务嵌入, 结合用户表示采用模态注意力计算匹配评分。该机制能够自适应处理模态缺失与偏好异质性, 在多模态服务匹配环境下展现出良好的匹配精度与泛化能力。

4) 在 Amazon 多领域多模态数据集上开展大规模实证研究, 结果表明本文方法在 Recall@K 和 NDCG@K 等服务匹配指标上优于跨领域服务匹配模型, 进一步验证了所提出的多模态图结构建模、模态感知图池化模块与匹配策略在提升跨领域服务匹配性能方面的有效性。

1 相关工作

随着多源用户数据的广泛可得, 领域对齐与共

享表示学习逐渐成为跨领域服务匹配的核心研究方向,其主要目标是在保持用户个性差异的前提下,实现跨域偏好的有效泛化与用户建模。典型的跨领域匹配方法多基于共享表示学习框架展开。Liu 等^[8]提出了一种 LLM 增强的跨领域序列匹配模型 (LLMs enhanced cross-domain sequential re-recommendation model, LLM4CDSR),通过引入大语言模型对不同领域物品进行统一语义表示,从语义层面建立跨域关联,并结合带有对比正则化的可训练适配器,以提升模型对跨领域序列匹配任务的适配能力。Cao 等^[9]通过构建共享-私有子空间实现跨领域特征的分离,使模型能够同时捕捉用户的领域无关偏好与领域特定特征。Zhu 等^[10]引入元网络生成个性化迁移函数,以自适应地建模不同用户的跨域行为转移模式,从而显著提升了迁移过程的灵活性与个体化程度。Li 等^[11]进一步结合多视图建模思想,从时间序列、协同关系与个体特征三方面联合捕捉跨领域的动态偏好演化路径,在多粒度层面增强领域间信息共享的有效性。Li 等^[12]提出一种对比自编码器方法,用于编码全局用户嵌入表示,并设计了一种从各个源领域向目标领域迁移用户嵌入的机制。Hwang 等^[13]不仅区分每个域的域特定序列偏好及域共享序列偏好,还引入“域空间”和域间适应机制,以便在目标域的行为稀疏或不频繁时能够借助其它域的交互填补序列上下文。

与此同时,部分研究聚焦于迁移鲁棒性与泛化性能的提升。Tang 等^[14]提出分层特征抽取机制,通过多任务结构解耦不同层次的语义信息,以减少跨领域迁移过程中的特征干扰与共线性问题。Li 等^[15]以互信息约束为核心,最大化源域与目标域之间的共享信息量,从理论上强化领域泛化能力与特征迁移的稳定性。Ning 等^[16]在跨域场景中同时在嵌入与模型层面进行知识解耦,并引入基于图的随机游走对齐策略来辨识跨域中“相似用户/物品”,以改善域间重叠稀疏情况下的知识传递。Zhao 等^[17]提出一种基于自适应锚点链接学习的框架,在源域与目标域的用户表示空间中构建锚点链接,以实现跨域语义的自适应对齐。其通过锚点生成模块筛选高置信度的跨域匹配样本,并结合迭代优化机制动态调整锚点权重,从而提升跨域迁移的稳定性与精度。Du 等^[18]提出了一种分层子空+可逆变换+因果图机制,用以显式拆分域间共享因素与

域特定因素,并保证联合分布的可识别性,从而在弱相关/非重叠用户场景中依然保持良好泛化性。

此外,近期也有研究从 open-world 跨领域服务匹配的视角重新审视用户重叠稀疏与数据偏置问题。Xu 等^[19]提出的自适应多兴趣解耦方法通过构建多兴趣信息模块,在不依赖完全重叠用户的情况下,实现了对重叠与非重叠用户跨领域偏好的统一建模,并结合去偏估计机制缓解实际系统中由选择偏置引起的性能退化。该类方法从系统部署与泛化稳定性的角度增强了跨领域服务匹配在真实场景中的可行性,但其核心仍主要依赖于序列行为建模与隐式兴趣对齐,尚未显式刻画用户偏好的结构组成及多模态差异。

上述方法在领域级特征对齐、跨领域知识迁移以及泛化鲁棒性方面取得了显著进展,但是它们普遍依赖于统一嵌入空间假设,即认为来自不同领域的用户行为数据可以映射到同一潜在语义空间中进行建模。在实际应用中,该假设往往难以成立,因为不同领域在服务内容在模态构成上通常存在显著差异。仅依赖统一嵌入或浅层特征融合的建模方式,难以显式刻画跨领域偏好在结构与模态层面的差异性,从而限制了模型在复杂多源场景下的判别能力与泛化性能。

2 基于用户多模态图的跨领域服务匹配方法

为解决当前跨领域服务匹配中存在的用户表示压缩所导致的“模态无差”问题,本文提出一种基于多模态用户图的服务匹配方法。总体框架如图 1 所示,方法主要包括以下部分:1) MUG 构建,以用户在各领域中的偏好标签为锚点,构建其各个模态下的多模态节点,显式保留模态行为的领域特征与偏好分布。不同模态下的偏好节点被组织为模态一致性子图,从结构上避免了行为信息的混叠与丢失。同时为增强多模态结构建模能力,MUG 引入图结构表达方式,并基于开放知识图谱构建标签之间的语义连接边,引入虚拟辅助节点增强图的全局连通性,形成统一的多模态用户图。2) 多模态用户图双阶段池化,通过 Group GCN 对各个模态子图进行独立建模,提取全局模态偏好表示,再以全局向量引导局部子图特征聚合,并通过注意力机制实现模态间的表示融合。3) 多模态服务匹配,将

目标服务项视为多模态实体，分别提取其多模态特征。并基于用户在 MUG POOL 第一阶段中获得的全局模态偏好向量，与服务项模态特征共同计算跨模态注意力权重，以刻画用户对不同模态的偏好强度。最终，模型以模态注意力加权融合各模态点积分得到用户对目标服务项的匹配评分。

2.1 多模态用户图构建

在跨领域服务匹配场景中，用户的偏好往往散落在不同源领域，其行为数据在内容形式、模态分布以及标签体系上均存在显著差异。直接将多个领域的交互记录简单合并会导致语义混叠与结构冲突，因此在构建 MUG 之前，需要针对多领域行为进行统一化与结构化处理。为此，本文在 MUG 构建阶段引入了跨领域偏好抽取与标签对齐机制，使多模态用户图能够同时表达多个领域的偏好信息并支持跨域迁移。首先从用户历史行为中提取偏好标签，并确保所选标签可在开放知识图谱中链接，获取对应的语义描述。每个标签若具备图像、音频或视频等模态信息，则为每种模态生成独立节点，并连接至标签节点，形成模态一致子图。为建模标签间的语义关系，若两个标签在 OKG 中存在限定长度的可达路径，则在其节点间添加边。同时，引入

虚拟辅助节点连接所有标签子图，增强图结构连通性，避免信息孤岛。最终形成的 MUG 实现了用户多领域、多模态偏好的结构化建模。

2.1.1 节点特征

在多模态用户图中，节点特征由其对应模态数据生成。本文针对不同模态（文本，图片，音频，视频）分别采用成熟的预训练模型进行编码。这样能够显式保留各模态的独立表征能力，并形成围绕标签节点的“模态一致子图”。具体而言，对于文本模态，本文采用预训练语言模型 BERT-base 对服务描述进行编码，并取输出序列中 [CLS] token 对应的隐藏向量作为文本的全局语义表示。对于图像模态，采用预训练的 ResNet-50 模型提取视觉特征，并使用其全局平均池化层的输出作为图像嵌入表示。对于音频模态，采用 VGGish 模型提取音频特征表示。对于视频模态，采用 Clip4Clip 模型对视频帧进行编码，并结合时序建模机制生成视频的时空语义表示，从而增强视频模态的语义表达能力。在实验过程中，所有预训练模型参数均保持冻结，仅训练后续投影层与图模型参数。同时，为实现跨模态特征对齐，本文通过线性投影将不同模态映射至统一嵌入空间

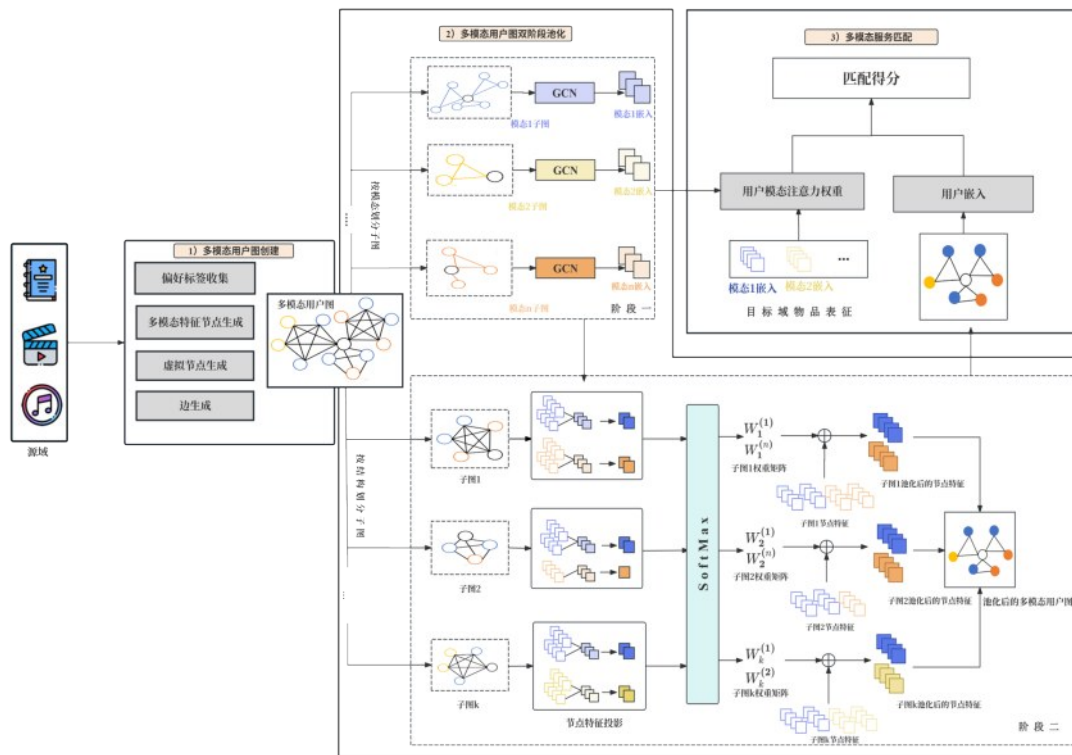


图1 总体框架

$$\mathbf{Mf}_i^{(m)} = W_m F_i^{(m)} + b_m \quad (1)$$

其中, $F_i^{(m)}$ 表示第 i 个标签在模态 m 下的特征表示, $\mathbf{Mf}_i^{(m)}$ 为 $F_i^{(m)}$ 经过线性投影之后的特征表示, $W_m \in \mathbb{R}^{d \times d(m)}$ 为模态 m 对应的线性投影矩阵, $b_m \in \mathbb{R}^d$ 为偏置项, $d(m)$ 表示原始模态特征维度, d 表示统一嵌入空间维度。采用线性映射而非深层 MLP 的原因在于: 1) 线性投影能保持原特征的几何结构, 避免深层非线性造成跨领域语义扭曲; 2) 对齐作用强于拟合作用, 更适合多模态、跨领域表征空间统一; 3) 结构简单且稳定, 在图结构学习与池化中能提供更稳健的节点表示。该投影机制确保不同模态在 MUG 中保持一致空间, 使多模态特征在跨领域场景中具备良好的可对齐性与结构一致性。

此外, 为避免不同领域的模态子图之间存在结构断裂, 本文引入虚拟辅助节点, 该节点在进行图的构建的过程中所有的特征维度初始化为 0。作为领域间与模态间的全局结构桥, 使多模态信息在 MUG 中能够进行跨模态、跨领域统一传播。

通过上述设计, 本文构建的多模态用户图不仅能够对标签的多模态语义进行细粒度建模, 更能够在跨领域场景中保持结构统一、语义一致与可迁移性, 为后续的跨领域服务匹配提供强结构支持。

2.1.2 边的添加

为构建多模态用户图中的结构边, 本文设计三步连接机制, 分别用于表示模态内聚合、语义标签关系与结构连通性。

1) 模态内连接

将每个标签的不同模态嵌入相互连接。该步骤构建了以标签为中心的模态一致子图 G_i , 用于后续的结构建模与服务匹配特征学习。

2) 标签语义连接

由于不同领域的偏好标签可能指向语义有关联但名称不同的实体, 仅依赖领域内标签无法实现跨领域偏好迁移。为此, 本文提出让多模态特征直接参与边的构建。文本语义用于知识图谱链接, 而视觉与音频模态进一步提升跨领域标签相似度的稳健性, 使得跨领域迁移不再仅依赖文本或 ID 级信息。这一机制显著增强了 MUG 的跨域表达能力。具体而言, 对于来自不同领域的标签, 若其在 OKG 中可在限定跳数内通过概念、属性或共现路径关联, 则将其视为“相关标签”, 在 MUG 中加入语义连接边。该机制显式桥接了不同领域的偏好, 使

MUG 能够捕捉跨领域迁移中真实存在的语义依赖关系, 而非单纯依赖共享嵌入空间。

3) 结构统一连接

考虑到部分标签之间可能在知识图谱中无直接关联, 为避免 MUG 被划分为多个孤立子图, 本文通过引入虚拟辅助节点, 将所有标签的模态特征节点统一连接至虚拟辅助节点。该操作保证了 MUG 的结构完整性与可传播性, 并为后续的用户偏好表示学习提供全局连接基础。

最终, 整合上述三类边连接策略后, 得到完整的多模态用户偏好图。

2.2 MUG POOL: 多模态用户图池化

在完成基于用户偏好构建的多模态用户图之后, 如何有效地从该多模态用户图中学习用户的统一表示向量, 成为接下来的关键问题。图池化技术为从复杂图结构中提取紧凑表示提供了一种有效路径, 但现有方法对于如 MUG 这样融合多模态图结构支持较为有限。考虑到 MUG 中同时包含多模态特征节点, 且结构上引入了虚拟辅助节点以增强连通性, 该图天然具备异构图的特征。因此, 传统的池化方法难以直接适用, 需设计一种面向多模态异构图的池化模块。

本小节提出 MUG POOL, 一种适用于 MUG 的双阶段图池化模块。第一阶段是 Group GCN (GGCN), 针对不同模态信息, MUG 被分解为对应模态子图, 每类子图包含相同模态的所有节点及虚拟辅助节点。分别对每个模态子图进行图卷积操作, 提取其对应模态的偏好嵌入。第二阶段, 在获得各个模态的偏好向量后, MUG POOL 引入一种基于“模态感知注意力”的池化策略, 对每个模态下的全连接子图进行权重计算与特征聚合。该操作使得池化结果不仅保留模态间的差异性, 也提升了结构表达的可解释性与鲁棒性。最终, 所有模态子图的池化输出将被整合为压缩后的图表示, 作为用户多模态空间下的聚合偏好表示, 并输入下游模块进行进一步处理。

2.2.1 Group GCN

由于多模态用户图中各类模态信息存在显著的异构性, 直接在整图上进行统一的图神经网络处理将造成模态间干扰和表示冲突。为此, 本文引入 GGCN 模块, 旨在对每种模态特征进行独立建模, 并保留其在全图中的联通性。首先, 以构建完成

的 MUG 为输入, 执行模态划分操作, 将原图划分为多个模态子图。每个模态子图仅保留其对应模态的特征节点与虚拟辅助节点, 忽略与该模态无关的其他节点。需要指出的是, 虽然子图内部不再包含多种模态节点, 但所有节点仍通过虚拟辅助节点相连, 从而在保留模态内局部结构信息的同时, 实现跨子图的信息传递路径。这种结构使得每个模态子图不仅具备清晰的模态语义边界, 还保持了原始图的全局可连通性。根据收集的模态种类, 多模态用户图被划分为以下多个模态子图, 令 $G^{(m)} = \{A^{(m)}, X^{(m)}\}$ 表示模态 m 的子图。每个子图由该模态下的所有节点及虚拟辅助节点共同组成。具体来说:

- $A^{(m)} \in \mathbb{R}^{(N(m)+1) \times (N(m)+1)}$ 为模态 m 对应子图的邻接矩阵, 其中 $N(m)$ 表示模态 m 下节点的数量, 额外的 1 表示虚拟辅助节点;

- $X^{(m)} \in \mathbb{R}^{(N(m)+1) \times d}$ 为模态 m 子图的特征矩阵, 由模态节点与虚拟辅助节点的特征拼接而成。虚拟辅助节点的特征初始化为全零向量, 用于增强图结构的连通性。

在构建出每个模态的子图后, 分别对每个子图应用图卷积网络进行特征提取, 以获得用户在该模态下的全局偏好表示。整个过程可表示为

$$U(m) = \text{GCN}(G^{(m)}) = \text{GCN}(A^{(m)}, X^{(m)}) \quad (2)$$

其中, $U(m)$ 表示用户在模态 m 下的全局偏好嵌入。

2.2.2 MUG 池化

MUG Pool 的核心创新在于引入模态感知的引导机制, 利用用户在模态 m 下的全局偏好表示 $U(m)$, 指导多模态用户图中的局部结构进行池化操作, 从而实现结构压缩与个性偏好对齐的统一建模。首先通过 GGCN 获取模态 m 下的全局向量 $U(m)$, 作为用户对每种模态的整体偏好概括。然后, 对于 MUG 中每块全连接子图 G_i , 根据 G_i 其所含模态构建模态特征矩阵, 并按行与 $U(m)$ 做内积, 计算响应权重向量, 衡量局部节点对全局偏好的契合程度。不同于传统图池化依赖邻接矩阵和固定结构规则, MUG Pool 在无需依赖图结构的前提下, 通过模态感知注意力机制自适应地保留高响应特征, 提升了信息选择性与计算效率。最终, 各模态聚合结果拼接形成结构紧凑、偏好一致的池化图表示。具体而言, 整个池化操作可以划分为以下三个步骤:

步骤 1 子图划分

首先, 将多模态用户图以虚拟辅助节点为界分解为若干局部全连接子图

$$\text{MUG} = \{G_1, G_2, \dots, G_N\} \quad (3)$$

其中每个子图 $G_i (i=1, 2, \dots, N)$ 均由虚拟辅助节点和若干偏好标签的多模态特征节点组成。由于每个 G_i 本身为全连接结构, 池化阶段可直接对特征集合进行操作, 无需重复加载或计算邻接矩阵, 从而有效降低计算开销。随后, 对每个子图 G_i 中的节点按所携带的模态类型进行分类, 形成模态特征子集。

通过上述子图划分与模态分类, MUG POOL 能够在后续步骤中针对每个子图的每种模态独立执行特征聚合, 提高池化效率与表达精度。

步骤 2 特征拼接

在完成子图划分与模态分类后, 本文对每个子图 G_i 内各模态节点的特征向量进行统一整理, 以便后续进行模态引导的特征聚合操作。具体而言, 本文为每种模态构建对应的特征矩阵 $X_i^{(m)}$ 。每个模态特征矩阵 $X_i^{(m)}$ 由子图 G_i 中所有模态为 m 的节点特征按行拼接而成, 定义如下

$$X_i^{(m)} = (\mathbf{x}_{i,1}^{(m)}, \mathbf{x}_{i,2}^{(m)}, \dots, \mathbf{x}_{i,N_i(m)}^{(m)}) \quad (4)$$

其中, $\mathbf{x}_{i,j}^{(m)} \in \mathbb{R}^{1 \times d}$ 表示第 i 个子图中模态 m 的第 j 个节点的特征向量; $N_i(m)$ 为子图 G_i 中属于模态 m 的节点数量。这种模态特征矩阵的构建方式为后续模态引导的注意力聚合操作提供了统一且高效的输入表示。

步骤 3 特征聚合

在完成模态特征矩阵构建后, 本文通过将模态局部特征与全局偏好表示进行线性投影, 衡量节点在全局语义方向上的响应程度。具体而言, 对于第 i 个子图中模态 m 的特征矩阵 $X_i^{(m)}$, 其按行与对应模态的全局偏好向量 $U(m)$ 相乘, 得到投影向量 $P_i^{(m)} = (p_{i,1}^{(m)}, p_{i,2}^{(m)}, \dots, p_{i,N_i(m)}^{(m)})^T$, 表示子图中每个模态 m 节点在全局偏好方向上的投影分值。第 j 个节点的投影值 $p_{i,j}^{(m)}$ 具体计算如下

$$p_{i,j}^{(m)} = \mathbf{x}_{i,j}^{(m)} \cdot U(m) \quad (5)$$

其中 $\mathbf{x}_{i,j}^{(m)}$ 是模态特征矩阵 $X_i^{(m)}$ 中第 j 行, 对应第 j 个模态节点的局部特征表示。该步骤实现了局部特征向量在全局偏好方向上的显性对齐, 为后续加权聚合提供了注意力基础。

完成特征投影后, 需根据节点在全局偏好方向

上的响应程度,对局部特征进行加权聚合。为此,本文对投影向量 $\mathbf{P}_i^{(m)}$ 应用 SoftMax 操作,生成归一化的注意力权重,用于衡量每个节点对最终模态表征的贡献程度,具体地,第 j 个节点对应的归一化权重 $w_{ij}^{(m)}$ 计算如下

$$w_{ij}^{(m)} = \frac{\exp(p_{ij}^{(m)})}{\sum_{j=1}^{Ni(m)} \exp(p_{ij}^{(m)})} \quad (6)$$

其中, $p_{ij}^{(m)} \in \mathbb{R}$ 为投影向量 $\mathbf{P}_i^{(m)}$ 第 j 个节点的值。该权重反映了节点 $\mathbf{x}_{ij}^{(m)}$ 相对于用户全局偏好 $\mathbf{U}(m)$ 的相似度。最终,利用归一化权重对特征矩阵进行加权求和,得到模态 m 在第 i 个子图下的池化输出表示

$$\bar{\mathbf{X}}_i^{(m)} = \sum_{j=1}^{Ni(m)} w_{ij}^{(m)} \cdot \mathbf{x}_{ij}^{(m)} \quad (7)$$

该操作在保证保留关键节点信息的同时,也充分考虑了非显著节点的模态上下文贡献,为最终多模态用户偏好融合打下基础。

2.3 多模态服务匹配

在 section2.2 中,本文通过 MUG POOL 模块实现了对原始多模态用户图中各个子图的模态级特征聚合,得到池化后的图结构。在该图中,除虚拟辅助节点外的每个节点均表示一个模态内子图的压缩嵌入。由于池化后的 MUG 为一个模态融合后的异构图,节点类型仍存在差异,因此不可直接将所有节点特征进行求平均或拼接。为进一步获取用户在整体图结构下的统一表示,本文引入图池化模块 DiffPool^[20],以实现从结构图到单点嵌入的自适应聚合。DiffPool 层通过结构学习与特征学习的联合优化策略,将其映射为仅包含一个虚拟辅助节点的紧凑图结构,进而获得最终的用户嵌入表示向量 \mathbf{u} 。

本文提出的模型将在一个对目标域的评分预测任务中进行训练与评估。对于每一个目标项目

$$Loss = -\frac{1}{|\text{TS}|} \sum_{n=1}^{|\text{TS}|} \log \left(\frac{\exp(\text{score}_{u, \text{positive}})}{\exp(\text{score}_{u, \text{positive}}) + \sum_{i \in \text{item}_{\text{negative}}} \exp(\text{score}_{u, i})} \right) \quad (11)$$

其中, $\text{score}_{u, \text{positive}}$ 为用户 u 对正样本 $\text{item}_{\text{positive}}$ 的预测分数; $\text{score}_{u, i}$ 为用户 u 对第 i 个负样本的预测分数; TS 为训练集中所有正样本的集合。该损失函数旨在最大化用户对正向项目的预测评分相对优势,从而提升匹配结果的排序能力。

本文将每个服务项建模为一个结构化的多模态

item _{n} , 根据其所具备的多模态信息,进行相应的模态特征提取,并通过线性映射投影至相同的维度空间。

考虑到不同用户在多模态空间下的偏好差异,以及跨模态语义间潜在的互补关系,本文借鉴 CAMR 模型^[21]提出的跨模态注意力增强策略,设计了一种自适应的模态感知注意力机制。具体而言,用户 u 对于项目 n 的模态 m , 首先利用用户在 MUG 中的模态偏好向量 $\mathbf{U}(m)$ (由 GGCN 得到) 与项目模态表示 $\mathbf{E}_n(m)$ (由预训练模型提取之后,经过线性映射层得到) 共同计算跨模态注意力得分 $s_{u,m}$

$$s_{u,m} = \text{LeakyReLU}(a^T [W_u \mathbf{U}(m) \| W_m \mathbf{E}_n(m)]) \quad (8)$$

其中, $W_u \in \mathbb{R}^{d \times d}$ 与 $W_m \in \mathbb{R}^{d \times d}$ 为可学习权重矩阵, $a \in \mathbb{R}^{2d}$ 为注意力投影向量。// 表示拼接操作。随后通过 softmax 归一化得到用户对项目模态的注意力权重 $\alpha_{u,m}$

$$\alpha_{u,m} = \frac{\exp(s_{u,m})}{\sum_{m' \in M_n} \exp(s_{u,m'})} \quad (9)$$

其中 M_n 表示项目 n 所具备的模态集合。最终,用户偏好评分 $\text{score}_{u,m}$ 通过各模态特征的加权点积计算

$$\text{score}_{u,m} = \sum_{m \in M_n} \alpha_{u,m} (\mathbf{u}^T \mathbf{E}_n(m)) \quad (10)$$

该设计使注意力权重由用户在全局偏好空间的语义表示与具体项目模态特征共同决定,能够自适应捕捉跨模态的协同与补充关系,从而在服务匹配中实现更精准的模态重要性建模。

为训练上述偏好预测模型,本文采用负采样的 pairwise ranking 策略。对于每个正样本 $\text{item}_{\text{positive}}$, 随机采样 k 个用户未交互的负样本 $\text{item}_{\text{negative}}$, 并使用如下基于 softmax 的 ranking loss 进行优化

实体,包含文本、图像、音频和视频四种基本模态。利用 section2.1.1 中对应的预训练模型对每个模态进行特征提取,并将不同模态嵌入投影至统一空间。

3 实验与结果分析

本文模型嵌入维度设为 $d = 128$ ，图神经网络层数为 2 层，各层输出维度保持一致，并采用 ReLU 作为激活函数。为缓解过拟合，在每层后引入 dropout (rate=0.5)。对于多模态特征，通过线性投影层映射至统一空间。模型训练采用 Adam 优化器，学习率设为 0.001，batch size 为 256，训练轮数为 100，并引入 L2 正则化 ($\lambda = 1e^{-5}$)。实验基于 PyTorch 框架，在 NVIDIA RTX 4060 GPU 上完成。

3.1 数据集和领域设置

本文采用 Amazon 多领域数据集，该数据集包含丰富的用户行为、商品属性与模态内容信息，广泛应用于多领域匹配与跨模态建模研究。为明确不同领域中模态信息的可用性，表 1 总结了各领域所包含的模态类型（文本、图像、视频、音频）：

	Text	Image	Video	Audio
Book	√			
Movie	√	√	√	√
CD	√			√
Game	√	√	√	

同时，为验证本文方法在跨模态偏好建模任务中的有效性，选取四个典型子领域（Book、Movie、CD 和 Game）开展实验。为确保用户在多个领域中具有充分的行为数据，本文筛选在至少两个领域中存在评分记录的用户作为有效样本。不同跨领域组合规模下的数据统计如表 2 所示。表中给出了不同领域组合规模对应的用户数量、商品数量，以及正负反馈交互分布，其中评分 ≥ 4.0 的交互被视为正样本，评分 ≤ 2.0 的交互被视为负样本。

Combination Size	#Users	#Items	#Rating ≥ 4.0	#Rating ≤ 2.0
4	2807	127673	300252	31686
3	12613	188860	597061	60006
2	29086	165948	601500	69304

3.2 评估指标

在数据预处理阶段，本文首先移除交互次数少

于 5 的用户与物品，以减少极端稀疏样本带来的干扰。对于模态数据，若某一模态缺失，则不构建对应的模态节点，模型仅基于已有该模态的节点进行图表示学习与注意力计算，从而避免无效信息的引入。

在模型训练与测试阶段，本文采用基于评分阈值构建隐式反馈信号：

- 对于评分 ≥ 4.0 的项目，视为该用户的正向偏好物品；
- 若存在评分 ≤ 2.0 的项目，则作为其负面反馈示例，其余评分不参与训练；
- 从正反馈中随机采样 60% 的项目用于构建该用户的多模态用户图，剩余 40% 的正反馈项目作为测试集。

在训练过程中，采用负采样策略，训练集中的每一个正样本配合随机采样的 5 个未交互负样本构成训练批次，并通过 MUG POOL 学习全局模态表示。对实验结果的评估，根据文献[9]，考虑使用：

- **Recall@K (R@K)**: 衡量正样本是否出现在预测结果的前 K 项中；
- **Normalized Discounted Cumulative Gain@K (NDCG@K)**: 在考虑排名位次的基础上评估匹配结果的相关性。

本文主要报告 $K \in \{10, 20\}$ 下的 R@K 与 NDCG@K，评价模型在多模态偏好表示下的排序性能。

3.3 基线方法

为全面评估所提出的跨领域服务匹配方法 MUG-CDSM 在多模态跨领域匹配任务中的有效性，本文选取了多种经典与最新的跨领域匹配模型作为对比基线，涵盖传统矩阵分解、跨域表征、多模态建模以及时序匹配等多个类别。在相同实验设置下，通过设定不同随机种子进行多次独立运行，所有实验结果均为 5 次独立实验结果的平均值。具体如下：

BPRMF^[22]: 基于贝叶斯个性化排序的矩阵分解方法，是匹配系统中经典的隐语义建模基线。

PLE^[14]: 一种参数可共享的多任务学习架构，常用于建模多领域用户行为的服务匹配。

UniCDR^[9]: 一个统一的跨领域匹配模型，旨在构建领域无关的用户表示以促进迁移学习。

EDDA^[11]: 基于事件驱动和双重注意力机制的

跨领域匹配模型, 建模用户在多领域中的偏好演化过程。

PTUPCDR^[10]: 基于时间序列建模的跨领域用户偏好建模方法, 融合个性化时间依赖与偏好表示。

CAT-ART^[12]: 将上下文感知、时序注意力与强化学习机制相结合的匹配模型, 能够动态调整用户偏好状态。

LLM4CDSR^[8]: 通过引入大语言模型在语义层面统一跨领域物品表示, 并结合适配与用户画像机制以缓解用户重叠稀疏与迁移复杂性的跨领域序列服务匹配方法。

表 3 实验结果

方法	指标			
	R@10	N@10	R@20	N@20
BPRMF	0.0125	0.0132	0.0228	0.0236
PLE	0.0109	0.0118	0.0187	0.0189
UniCDR	0.0261	0.0275	0.0501	0.0511
EDDA	0.0207	0.0211	0.0399	0.0401
PTUPCDR	0.0221	0.0237	0.0445	0.0447
CAT-ART	0.0275	0.0277	0.0511	0.0526
LLM4CDSR	0.0281	0.0291	0.0533	0.0542
MUG-CDSM(Ours)	0.0291	0.0304	0.0549	0.0561

从表 3 可以观察到, 不同类别模型在跨模态跨领域匹配任务中的性能差异较为显著, 整体呈现出“建模粒度越细、结构感知能力越强, 性能越优”的趋势。首先, 传统矩阵分解与参数共享类方法表现较弱。BPRMF 和 PLE 在各项指标上均处于最低水平, 尤其在 R@10 与 N@10 上差距明显。这主要是由于此类方法依赖隐语义向量或任务级参数共享, 难以刻画跨领域场景下用户偏好的结构异质性与模态差异, 对跨模态信息的建模能力有限。其次, 统一用户表示与序列建模方法带来了显著提升, 但仍存在瓶颈。UniCDR、EDDA、PTUPCDR 以及 CAT-ART 在各项指标上均明显优于传统方法, 说明通过构建共享用户表示、引入事件驱动或时间序列机制, 能够有效缓解跨领域数据稀疏与偏好迁移问题。其中, CAT-ART 和 LLM4CDSR 的性能较为突出, 表明上下文感知、注意力机制在跨域匹配中具有重要作用。然而, 这类方法大多将多模

态信息视为特征层补充或统一嵌入空间中的输入信号, 缺乏对用户偏好内部模态结构的显式建模, 容易在跨模态场景下造成语义压缩或模态贡献不均的问题。相比上述方法, MUG POOL 在所有评价指标上均取得最优结果, 且在高阶排序指标上的提升尤为稳定, 表明其不仅提高了召回能力, 也显著增强了排序质量。这一性能优势主要来源于以下几个方面: 1) 模态粒度的偏好建模方式, MUG POOL 并非直接对用户行为进行统一编码, 而是显式构建用户的多模态偏好子图, 将不同模态视为结构化节点进行建模, 充分考虑模态差异性, 从而保留了用户偏好内部的细粒度模态信息。2) 模态感知池化模块缓解表示压缩问题, 通过模态感知池化, 模型能够根据不同模态在当前领域和任务中的贡献程度, 自适应地聚合关键信息, 避免传统平均或注意力聚合中对弱模态的忽略。3) DiffPool 支持跨领域结构级迁移, 在模态级建模基础上, 引入 DiffPool 将多模态偏好结构进一步聚合为领域级表示, 使用户表示具备更强的结构一致性与迁移鲁棒性, 有效支持跨领域匹配。

3.4 消融实验

为进一步验证本文提出模型中关键设计模块对匹配性能的实际贡献, 本文围绕 MUG 的知识图构建深度与模态全局特征提取策略两方面进行了系统的消融实验。通过对比不同设置下的性能指标, 本文旨在明确各模块的有效性, 并为后续系统部署提供实用参数指导。

3.4.1 OKG 检索层数对 MUG 构建与性能的影响

在多模态用户图构建阶段, 标签实体间的联系依赖于公共知识图谱的多跳检索机制。尽管检索层数的增加可引入更丰富的语义边以增强图连通性, 但过深的检索也可能带来噪声连接与计算负担。因此, 检索层数是 MUG 构建中一个关键的超参数。为验证其影响, 本文将 OKG 检索深度设置为 $Level \in \{0, 1, 2, 3, 5\}$, 并控制其余实验变量一致, 评估在不同 Level 值下的匹配效果。实验结果如图 2 所示, 图中未填充柱表示 Recall@10 (R@10), 斜线填充柱表示 NDCG@10 (N@10):

实验表明, 当检索层数为 2 时模型性能最优。进一步加深检索深度反而导致 Recall 与 NDCG 指标下降, 说明冗余实体或噪声关系的引入降低了特征表达的精度。此外, 图构建复杂度也随着层数呈

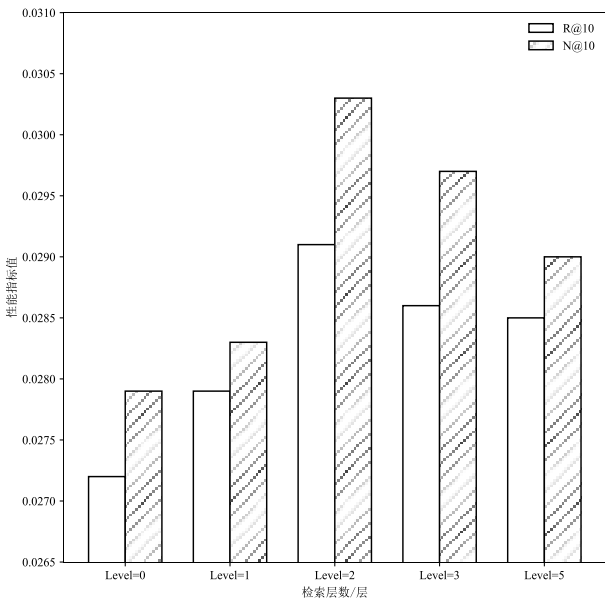


图2 OKG 检索的层数对于实验结果的影响

指数增长，构成时间效率的瓶颈。因此，检索两层 OKG 在保证实体关联丰富性的同时，也能维持图结构的表达可控性与计算效率。该现象说明用户在匹配中的偏好往往集中于具备相似语义或属性的实体，而浅层语义连接已足以捕捉这些相关性。

3.4.2 不同模态全局特征提取策略对比分析

为了构建每个模态的全局偏好嵌入，本文采用 GGCN 模块对 MUG 中的模态子图进行结构化建

模。为验证 GGCN 的有效性，本文对比了五种不同的全局图特征提取方法在匹配任务中的表现，具体包括：

- DeepWalk^[23]: 基于随机游走的节点嵌入方法；
- Graph Attention Network^[24]: 引入注意力机制的图注意力网络；
- GraphSAGE^[25]: 基于邻居采样与聚合的 GNN 模型；
- GCN + GAP: GCN 提取结构特征后进行全局平均池化；
- GCN + GAP + 负例优化: 本文最终采用方案，加入判别式训练策略。

实验结果如图 3 所示，图中未填充柱表示 Recall@10 (R@10)，斜线填充柱表示 NDCG@10 (N@10)：

根据图 3 中的实验结果，可以发现 DeepWalk 的性能显著低于其他方法。这主要归因于 DeepWalk 专注于捕捉图的局部结构特征，忽视了全局语义关联。而在 GGCN 模块中，每一个模态子图本身结构已被划分并简化，节点之间的全局语义连接成为建模效果的关键。因此，DeepWalk 的局部路径建模策略与该任务目标存在根本性偏离，难以有效建构全图级的模态偏好表征。Graph Attention Network 试图通过学习不同邻居节点的注意力权重

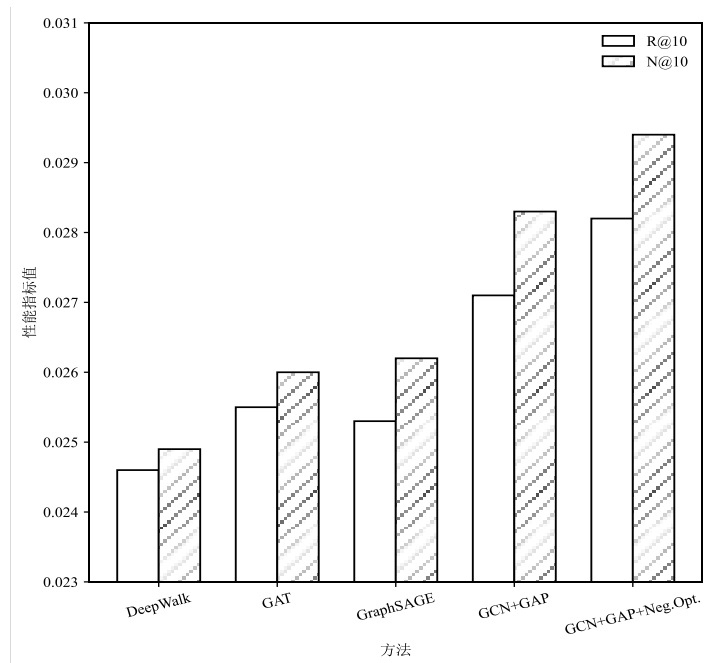


图3 不同全局特征提取策略结果

来增强特征表达能力。然而,在 MUG 的图结构中,为保证图的连通性,本文引入了一个虚构的虚拟辅助节点将原本语义无关联的模态偏好子图连接成整体。虽然结构上虚拟节点是图的连接枢纽,但其并不携带任何实际语义信息,仅作为结构稳定性的辅助元素存在。这一特性使得 GAT 难以合理学习其注意力权重,反而导致模型在更新邻接节点特征时引入噪声,从而影响整体性能。类似地,GraphSAGE 强调基于邻居节点的特征采样与聚合策略。但由于 MUG 中的邻居连接主要依赖虚拟节点的人工连接,邻居采样无法真实反映语义邻接性,使得 GraphSAGE 的设计假设在该结构下失效,难以捕捉用户模态偏好间的高质量聚合特征。相比之下,GCN + GAP 在处理包含结构性异常节点(如虚拟节点)时表现更为稳健。GCN 能在结构层面有效传播全图信息,而 GAP 则通过对所有节点进行均值聚合,避免了对个别节点的过度依赖,从而实现了更具鲁棒性的模态全局嵌入生成。进一步地,本文在 GCN + GAP 基础上引入负例优化机制,可显式引导模型区分语义相关与不相关的节点,从而提升模态嵌入表示的判别能力与泛化效果。综上所述,传统依赖局部结构或节点可解释性的图神经网络(如 GAT、GraphSAGE)在 MUG 中由于虚拟节点的结构特殊性难以发挥优势,而 GCN + GAP + 负例优化能有效兼容异构结构,成为适用于本任务的最佳特征提取策略。

3.4.3 模态有效性

为了进一步验证各模态在 MUG 模型中的实际贡献与必要性,本文通过分别移除图构建与特征提取过程中的部分模态,观察服务匹配性能的变化,从而探究了这样一个核心问题:每种模态在最终匹配效果中发挥了多大作用?

在保持图结构、图池化模块以及下游评分匹配模块不变的前提下,本文构建了以下模型变体:

w/o Image: 移除所有图像模态节点及其特征;

w/o Audio: 移除音频模态节点及其特征;

w/o Video: 移除视频模态节点及其特征;

w/o Text: 移除文本特征,仅保留其他模态;

Text Only (w/o I V A): 仅保留文本模态,模拟传统单模态用户建模方式;

Full MUG-CDSM (完整模型): 保留全部模态信息。

其余所有模块(如 OKG 语义边、虚拟辅助节点、GGCN 建模、模态感知池化等)均保持一致,以确保公平对比。

表 4 模态有效性验证的实验结果

实验设置	指标			
	R@10	N@10	R@20	N@20
w/o Image	0.0266	0.0271	0.0511	0.0523
w/o Audio	0.0278	0.0289	0.0524	0.0530
w/o Video	0.0276	0.0285	0.0521	0.0525
w/o Text	0.0194	0.0207	0.0403	0.0417
Text Only	0.0248	0.0252	0.0502	0.0511
MUG-CDSM(Ours)	0.0291	0.0304	0.0549	0.0561

根据表 4 的实验结果,可以发现文本模态是最核心的信息来源:移除文本模态后,模型性能下降最为显著,说明文本在标签语义建模与偏好表达中具有锚定作用。而图像、音频和视频模态提供关键补充信息:当分别去除图像、音频或视频模态时,相关指标均有下降,表明这些模态有助于补充用户偏好的细粒度特征。多模态融合具备显著增益:完整模型在所有指标上表现最佳,表明 MUG-CDSM 能够有效融合并协调多种模态,提升整体匹配性能。

4 结束语

本文提出了一种面向跨领域服务匹配任务的用户建模方法,构建了基于用户偏好多模态信息的多模态用户图。与交互序列的建模方式不同,MUG-CDSM 能够充分挖掘用户在多种模态维度上的细粒度偏好特征,从而构建更为全面且个性化的用户表示。为解决多模态偏好节点之间可能存在的结构断裂问题,本文在图构建过程中引入了一个虚拟辅助节点,用于增强图的连通性并促进各模态子图间的信息传递,使得用户在不同偏好模态下的相关性得以更好建模。此外,本文设计了一种面向多模态用户图的池化模块 MUG POOL,该方法通过结合模态全局语义信息与子图结构特征,实现了对用户多模态偏好的有效聚合,最终生成统一的用户嵌入表示。在多个大规模跨领域多模态服务匹配数据集上的实验结果表明,本文方法在服务匹配的性能上优于现有主流基线模型,验证了细粒度模态建模与图结构融合策略在跨领域服务匹配中的有效性。

参考文献:

- [1] Purificato E, Boratto L, De Luca E W. User Modeling and User Profiling: A Comprehensive Survey[J]. arXiv preprint arXiv: 2402.09660, 2024.
- [2] Ju H, Kang S K, Lee D, et al. Multi-domain recommendation to attract users via domain preference modeling[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2024, 38(8): 8582-8590.
- [3] Sun C, Gu J, Hu B B, et al. REMIT: reinforced multi-interest transfer for cross-domain recommendation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2023, 37(8): 9900-9908.
- [4] Liu J, Huang W, Li T, et al. Cross-domain knowledge graph chiasmal embedding for multi-domain item-item recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 35(5): 4621-4633.
- [5] 秦志金, 赵葵葵, 李凡, 等. 多模态语义通信研究综述[J]. 通信学报, 2023, 44(5): 28-41.
- Qin Z J, Zhao T T, Li F, et al. A survey of multimodal semantic communications[J]. Journal on Communications, 2023, 44(5): 28 - 41.
- [6] Guo X, Li S, Guo N, et al. Disentangled representations learning for multi-target cross-domain recommendation[J]. ACM Transactions on Information Systems, 2023, 41(4): 1-27.
- [7] Yang W, Yang J, Liu Y. Multimodal optimal transport knowledge distillation for cross-domain recommendation[C]//Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. 2023: 2959-2968.
- [8] Liu Q, Zhao X, Wang Y, et al. Bridge the domains: Large language models enhanced cross-domain sequential recommendation[C]//Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2025: 1582-1592.
- [9] Cao J, Li S, Yu B, et al. Towards universal cross-domain recommendation[C]//Proceedings of the sixteenth ACM international conference on web search and data mining. 2023: 78-86.
- [10] Zhu Y, Tang Z, Liu Y, et al. Personalized transfer of user preferences for cross-domain recommendation[C]//Proceedings of the fifteenth ACM international conference on web search and data mining. 2022: 1507-1515.
- [11] Li H, Yu L, Niu X, et al. Sequential and graphical cross-domain recommendations with a multi-view hierarchical transfer gate[J]. ACM Transactions on Knowledge Discovery from Data, 2023, 18(1): 1-28.
- [12] Li C, Xie Y, Yu C, et al. One for all, all for one: Learning and transferring user embeddings for cross-domain recommendation[C]//Proceedings of the sixteenth ACM international conference on web search and data mining. 2023: 366-374.
- [13] Hwang J, Ju H, Kang S K, et al. Multi-Domain Sequential Recommendation via Domain Space Learning[C]//Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2024: 2134-2144.
- [14] Tang H, Liu J, Zhao M, et al. Progressive layered extraction (ple): A novel multi-task learning (mtl) model for personalized recommendations[C]//Proceedings of the 14th ACM conference on recommender systems. 2020: 269-278.
- [15] Li Z, Amagata D, Zhang Y, et al. Mutual information-based preference disentangling and transferring for non-overlapped multi-target cross-domain recommendations[C]//Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2024: 2124-2133.
- [16] Ning W, Yan X, Liu W, et al. Multi-domain recommendation with embedding disentangling and domain alignment[C]//Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. 2023: 1917-1927.
- [17] Zhao Y, Li C, Peng J, et al. Beyond the overlapping users: Cross-domain recommendation via adaptive anchor link learning[C]//Proceedings of the 46th international ACM SIGIR conference on research and development in information retrieval. 2023: 1488-1497.
- [18] Du J, Ye Z, Guo B, et al. Identifiability of cross-domain recommendation via causal subspace disentanglement[C]//Proceedings of the 47th international ACM SIGIR conference on research and development in information retrieval. 2024: 2091-2101.
- [19] Xu W, Wu Q, Wang R, et al. Rethinking cross-domain sequential recommendation under open-world assumptions[C]//Proceedings of the ACM Web Conference 2024. 2024: 3173-3184.
- [20] Ying Z, You J, Morris C, et al. Hierarchical graph representation learning with differentiable pooling[J]. Advances in neural information processing systems, 2018, 31.
- [21] Wang R, Li C, Zhao Z. Towards user-specific multimodal recommendation via cross-modal attention-enhanced graph convolution network [J]. Applied Intelligence, 2025, 55(1): 2.
- [22] Rendle S, Freudenthaler C, Gantner, Z, and SchmidtThieme, L. BPR: Bayesian Personalized Ranking from Implicit Feedback[C]//Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence. 2009:454-461.
- [23] Jeyaraj R, Balasubramaniam T, Balasubramaniam A, et al. DeepWalk with Reinforcement Learning (DWRL) for node embedding[J]. Expert Systems with Applications, 2024, 243: 122819.
- [24] Chang H, Rong Y, Xu T, et al. Spectral graph attention network with fast eigen-approximation[C]//Proceedings of the 30th ACM international conference on information & knowledge management. 2021: 2905-2909.
- [25] Jiawei E, Zhang Y, Yang S, et al. GraphSAGE++: Weighted multi-scale GNN for graph representation learning[J]. Neural Processing Letters, 2024, 56(1): 24.

[作者简介]



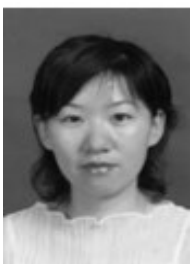
王海艳 (1974-), 女, 江苏东台人, 博士, 南京邮电大学教授、博士生导师, 主要研究方向为服务计算、可信计算、大数据应用与云计算技术、隐私保护技术等。



刘万宇 (2001-), 男, 江苏南京人, 南京邮电大学硕士生, 主要研究方向为多模态用户偏好的提取与表示。



张少聪 (2002-), 女, 河北衡水人, 南京邮电大学硕士生, 主要研究方向为边缘计算任务卸载。



骆健 (1976-), 女, 江西赣州人, 南京邮电大学副教授, 主要研究方向为服务计算、可信计算、服务推荐等。