

# 异质信息网络高阶层次化嵌入学习与推荐预测\*

荀亚玲, 毕慧敏, 张继福

(太原科技大学 计算机科学与技术学院, 山西 太原 030024)

通信作者: 张继福, E-mail: [zjf@tyust.edu.cn](mailto:zjf@tyust.edu.cn)



**摘要:** 异质信息网络是一种异质数据表示形式, 如何融合异质数据复杂语义信息, 是推荐系统面临的挑战之一. 利用弱关系具有的丰富语义和信息传递能力, 构建一种面向推荐系统的异质信息网络高阶嵌入学习框架, 主要包括: 初始化信息嵌入、高阶信息嵌入聚合与推荐预测 3 个模块. 初始化信息嵌入模块首先采用基于弱关系的异质信息网络最佳信任路径筛选算法, 有效地避免在全关系异质信息网络中, 采样固定数量邻居造成的信息损失, 其次利用新定义的基于多头图注意力的多任务共享特征重要性度量因子, 筛选出节点的语义信息, 并结合交互结构, 有效地表征网络节点; 高阶信息嵌入聚合模块通过融入弱关系及网络嵌入对知识良好的表征能力, 实现高阶信息表达, 并利用异质信息网络的层级传播机制, 将被采样节点的特征聚合到待预测节点; 推荐预测模块利用高阶信息的影响力推荐方法, 实现了推荐任务. 该框架具有嵌入节点类型丰富、融合共享属性和隐式交互信息等特点. 最后, 实验验证 UI-HEHo 学习框架可有效地改善评级预测的准确性, 以及推荐生成的针对性、新颖性和多样性, 尤其是在数据稀疏的应用场景中, 具有良好的推荐效果.

**关键词:** 推荐预测; 异质信息网络; 网络嵌入; 共享特征; 重要性度量因子

**中图法分类号:** TP311

中文引用格式: 荀亚玲, 毕慧敏, 张继福. 异质信息网络高阶层次化嵌入学习与推荐预测. 软件学报, 2023, 34(11): 5230–5248. <http://www.jos.org.cn/1000-9825/6681.htm>

英文引用格式: Xun YL, Bi HM, Zhang JF. Higher-order Hierarchical Embedding Learning and Recommendation Prediction in HIN. Ruan Jian Xue Bao/Journal of Software, 2023, 34(11): 5230–5248 (in Chinese). <http://www.jos.org.cn/1000-9825/6681.htm>

## Higher-order Hierarchical Embedding Learning and Recommendation Prediction in HIN

XUN Ya-Ling, BI Hui-Min, ZHANG Ji-Fu

(College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China)

**Abstract:** Heterogeneous information network is a representation of heterogeneous data. How to integrate complex semantic information of heterogeneous data is one of the challenges faced by recommendation systems. A higher-order embedded learning framework for heterogeneous information networks based on weak ties featured by semantic information and information transmission abilities is constructed. The framework includes three modules of initial information embedding, high-order information embedding aggregation, and recommendation prediction. The initial information embedding module first adopts the best trust path selection algorithm to avoid information loss caused by sampling a fixed number of neighbors in a full-relational heterogeneous information network. Then the newly defined importance measure factors of multi-task shared characteristics based on multi-head attention are adopted to filter out the semantic information of each node. Additionally, combined with the interactive structure, the network nodes are effectively characterized. The high-order information embedding aggregation module realizes high-order information expression by integrating weak ties and good knowledge representation ability of network embedding. The hierarchical propagation mechanism of heterogeneous information networks is utilized to aggregate the characteristics of sampled nodes into the nodes to be predicted. The recommendation prediction module employs the

\* 基金项目: 国家自然科学基金青年科学基金 (61602335); 山西省自然科学基金 (201901D211302); 太原科技大学博士科研启动基金 (20172017)

收稿时间: 2021-08-17; 修改时间: 2021-11-09, 2022-01-09; 采用时间: 2022-03-21; jos 在线出版时间: 2022-05-24

CNKI 网络首发时间: 2023-04-28

influence recommendation method of high-order information to complete the recommendation. The framework is characterized by rich embedded nodes, fusion of shared attributes, and implicit interactive information. Finally, the experiments have verified that UI-HEHo can effectively improve the accuracy of rating prediction, as well as the pertinence, novelty and diversity of recommendation generation. Especially in application scenarios with sparse data, UI-HEHo yields good recommendation effects.

**Key words:** recommendation prediction; heterogeneous information networks (HIN); network embedding; shared characteristics; importance measure factor

大数据时代, 互联网用户既是信息的消费者又是信息的生产者, 信息过载问题日益突出. 推荐系统作为解决信息过载最有效的方式之一, 能够帮助用户从大量信息中筛选有价值的信息, 逐渐成为信息领域的研究热点之一. 传统基于强关系的推荐因其本身圈子小、产生的唠叨型内容较优质内容多, 往往导致单一化推荐<sup>[1-3]</sup>. 随着在线社交网络的流行, 用户的行为不仅是个人兴趣爱好的体现, 而会受到与其有交互用户的影响, 因而产生了大量包含推荐信息的异质数据, 如何有效地利用异质数据, 进一步改善推荐效果和质量, 是推荐系统面临的挑战之一.

异质信息网络是推荐系统中涉及异质数据类型的一种数据表示形式<sup>[4]</sup>, 能够有效地刻画和描述不同类型的节点信息和关系. 但大多数异质信息网络推荐算法依赖于推荐节点路径的相似度, 未能充分利用用户和项目节点的潜在结构特征以及节点的属性等辅助推荐信息, 影响了推荐的准确性、新颖性和多样性<sup>[5,6]</sup>. 网络嵌入作为网络表征学习的一种方法, 有效地促进了信息网络的分析和可视化, 但大多数算法都集中在同构网络上<sup>[7]</sup>, 难以体现不同类型实体和实体间的联系信息. 为了有效利用异质网络中的复杂结构, 融合更多类型节点间的直接或间接交互信息辅助推荐, 本文充分利用弱关系的丰富语义和信息传递, 以及网络嵌入对知识良好的表征能力, 构建了一种异质信息网络高阶分层嵌入学习框架. 该框架通过有效融合更丰富的节点类型及其交互信息, 并通过嵌入学习捕获重要的关系结构, 有效地解决了推荐数据稀疏造成的冷启动等不足, 改善了推荐的准确性、新颖性和多样性. 其主要贡献如下.

- 构建了一种全关系异质信息网络, 并给出最佳信任路径筛选算法.
- 定义了一种基于多头图注意力机制的多任务共享特征重要性度量因子.
- 提出了一种融合多类型用户信息和信息传递的高阶信息表达模型.
- 给出了一种面向高阶信息表达模型的影响力推荐算法.
- 构建了一种异质信息网络高阶嵌入学习框架.

## 1 相关工作

推荐系统可以有效地帮助用户从海量数据中发现自己的信息需求, 但随着在线社交网络的流行, 产生了大量异质数据, 对推荐生成的针对性、新颖性和多样性, 提出了新的挑战, 为此国内外学者从信任关系推荐、异质信息网络、网络嵌入学习等方面开展了大量研究工作.

### 1.1 信任关系推荐

传统推荐算法大多是基于“相似性”或深度信任的强关系推荐, 推荐效果往往趋于同质. 随着 Twitter、Facebook、YouTube 等在线社交网络的流行, 用户的行为不再是个人兴趣和爱好的体现, 一定程度上会受到与其有社会关系的其他用户的影响<sup>[8,9]</sup>. 一般情况下, 消费者更愿意相信朋友或者有影响力的人的推荐, 而信任作为人际关系的核心将会直接影响用户的决策, 因此, 基于信任的弱关系在扩大推荐的选择范围的基础上提供了推荐的依据.

Qin 等人<sup>[10]</sup>和 Lin 等人<sup>[11]</sup>以弱关系为背景, 根据用户间的相似性进行协同过滤, 与强关系推荐相比, 虽产生了大量的候选推荐对象, 但没有解决数据稀疏造成的冷启动问题. 为了改善这一问题, 学者们将“信任”引入推荐, 信任系统的基本理念是允许用户在没有直接邻居的情况下相互评价: Li 等人针对数据稀疏和冷启动问题, 提出了一种结合社会标签和信任关系的社交网络推荐方法, 通过概率矩阵分解收集社会信任关系等信息<sup>[12]</sup>; Cui 等人针对社交网络中项目评价的局限性, 提出了一种将信任、领域倾向和项目声誉综合在一起的信任评价模型 (CTDR)<sup>[13]</sup>; Forsati 等人将显式的不信任关系引入推荐, 提出了一种结合信任和不信任的基于矩阵分解的推荐模型<sup>[14]</sup>; Zhang

等人为了降低恶意实体推荐值对信任评估的影响,提出考虑实体间交互过程、引入奖惩因素和评价可靠性因素的推荐信任动态评估模型<sup>[15]</sup>; Hamdi 等人从传递性的概念出发,研究了社交网络中信任传播的性质,引入了 TISoN 模型生成和评估在线社交网络中的信任推理<sup>[16]</sup>.

## 1.2 信任异质信息网络

将多类型的网络化数据建模成异质信息网络 HIN (heterogeneous information network),能够充分利用网络中丰富的对象和关系信息,改善了数据稀疏造成的冷启动等问题. Hu 等人提出了基于异质图神经网络的半监督短文本分类方法帮助用户找到相关信息,解决了实际应用中缺乏标记训练数据的问题<sup>[17]</sup>; Mnih 等人提出的经典的概率矩阵分解模型 PMF,采取低维度因子模型 (low-dimensional factor) 将评级矩阵分解成两个低维矩阵,在用户评价较少的情况下有一定的预测效果<sup>[18]</sup>; Rendle 等人提出的因子分解机 FM (factorization machine) 模型解决了数据稀疏情况下的特征组合问题<sup>[19]</sup>. 现有的异质信息网络推荐算法大多数是依赖网络路径的相似性来评价对象的相似性: Shi 等人提出基于加权异质信息网络的协同过滤方法,预测目标用户对项目的评级<sup>[4]</sup>; Zheng 等人设计了双重相似正则化方法,利用对偶相似度正则化方法整合不同语义元路径下用户和项目的相似度信息<sup>[5]</sup>; 而元路径不能很好地获取同级别或同类型节点的复杂语义信息, Yuan 等人<sup>[20]</sup>和 Huang 等人<sup>[21]</sup>提出将元图的概念引入 HIN, 计算同类实体间的相似度,从而捕获元路径无法捕获的复杂语义; Zhao 等人将元路径融入 HIN, 采用矩阵分解+因子分解机的方法求解信息融合问题<sup>[6]</sup>.

总的来说,基于路径的相似度推荐是 HIN 推荐的主流研究方法,但不能充分利用 HIN 上用户和项目的潜在交互信息进行推荐.

## 1.3 网络嵌入学习

网络嵌入学习将复杂的信息转换为密集的低维实值向量,最大限度地保留了网络中的结构以及属性等信息<sup>[22]</sup>. 目前的网络嵌入可以分为“基于分解”“基于随机游走”“基于深度学习”这 3 类.

早期的网络嵌入学习大都是基于分解的,通过对描述图数据结构信息的矩阵进行分解,将节点转化到低维向量空间并保留结构的相似性. 知识管理图表征学习模型 GraRep, 通过学习低维向量来表示图的节点,利用矩阵分解技术构造顶点的全局表示<sup>[23]</sup>. Rendle 提出个性化排序矩阵分解模型 BPRMF<sup>[24]</sup>,只依靠用户的显式反馈作为信息交互的目标值; He 等人利用多层感知器和矩阵分解结合捕获用户和项目的非线性特征交互<sup>[25]</sup>. 虽然这些方法都有解析解,但结果过度依赖于相关矩阵的分解计算,造成了较高的时间和空间复杂度.

词向量方法在语言表示上的成功,启发了学者把在图中随机游走产生的序列看作句子,节点看作词,类比词向量方法来学习出节点的表示. 常见的方法有 DeepWalk<sup>[26]</sup>、node2vec<sup>[27]</sup>. Yang 等人提出 text-associated DeepWalk (TADW), 将具有文本特征的节点在 MF 的基础上合并进网络表示学习中<sup>[7]</sup>; Tu 等人提出最大利润深度步行 (MMDW), 学习到的表征不仅包含网络结构,而且具有判别性<sup>[28]</sup>; Ivanov 等人基于匿名行走以显示的和分布式的方式学习图形表示<sup>[29]</sup>; Dong 等人提出 metapath2vec++ 框架保留异构网络中的结构和语义联系<sup>[30]</sup>; Shi 等人设计了基于元路径的随机游走策略进行异构网络嵌入的方法 HERec-spl, 并通过融合函数将嵌入集成到矩阵分解模型<sup>[31]</sup>. 基于随机游走的方式实现了大规模的图表示学习,但是将图转化为序列集合没有利用图本身的结构信息,并且难以融合图中节点的属性信息进行表示学习.

基于深度学习的网络嵌入,大都是基于图神经网络挖掘用户-项目节点间的浅层交互,缺少挖掘高维的能力, DNNGR (deep neural networks for learning graph convolutional) 算法根据图生成共现概率矩阵,在此基础上得到 PPMI 矩阵,使用 SDAE (stacked denoising autoencoder) 编码得到向量<sup>[32]</sup>; 与 DNNGR 相比, SDNE (structural deep network embedding) 利用神经网络挖掘高阶相似性的同时,利用一阶相似性作为已知的监督信息<sup>[19]</sup>; Rianne 等人提出 GC-MC 模型,基于用户-项目二部图从链路预测的角度解决推荐系统中的评级预测问题<sup>[33]</sup>; 与 SDNE 相同,采用自动编码方式学习原始特征到中间表示的映射,虽适用于超大数据集的中间表示,但难以学习语义信息; Gao 等人提出学习二部网络顶点表示的方法 BiNE, 将显式和高阶隐式关系联合建模<sup>[34]</sup>, 但只考虑显示的连接信息; Tang 等人开发了 TranFG 框架,通过跨异质网络学习对社会关系类型进行分类<sup>[35]</sup>; Ebesu 等人提出协同记忆网络

(CMN), 根据特定的用户和项目学习用户邻域的自适应非线性加权, 输出与用户、项目记忆之间的非线性相互作用<sup>[36]</sup>; Tu 等人提出 CANE 模型, 通过相互注意机制对节点的结构和文本信息融合<sup>[37]</sup>; Ying 等人提出 PinSage 将图神经网络应用于通过对项目之间的相似性建模而形成的项目图<sup>[38]</sup>; Shi 等人发现把不同类型的节点映射到同一低维空间造成了信息丢失, 提出了 AspEm 对 HIN 进行层次划分<sup>[39]</sup>; Wang 等人提出的 NGCF 框架将用户-项目二部图集成到嵌入过程, 通过对图中的高阶连通性表达建模把协作信号注入嵌入过程<sup>[40]</sup>.

综上所述, 将弱关系产生的大量异质数据建模成异质信息网络, 尽管刻画了复杂的类型关系, 但大多数推荐算法仅依赖网络路径的相似度实现推荐预测, 不能很好地获取同级别或同类型节点的复杂语义信息. 此外, 当推荐数据稀疏时, 复杂数据类型的网络嵌入方法预测节点, 不能同时表征局部和全局结构特征信息.

## 2 异质信息网络高阶嵌入学习框架

### 2.1 问题提出

在网络嵌入中划分层次, 改善了因关系类型不同导致的相似信息被抵消或削弱的问题, 但现有算法忽略了除用户-项目交互外的其他类型交互或节点信息, 也没有考虑不同层次间的关联. 图 1 是以 User1 为目标节点的异质信息交互图, 描述了用户的偏好信息和项目的属性信息. 节点 Movie1、Movie2、Actor1 和 User2 作为 User1 的直接交互对象, 类型不同且彼此间的相关性较低, 因此 User1 只能嵌入到它们的中间位置, 导致 User1 看起来并没有偏好这几个节点, 出现了信息抵消和丢失. 研究发现, 没有考虑层次间存在关联的划分, 如图 2 中的 (a) 和 (b), 划分层次后只能看出 User1 对 Movie1 和 Movie2 的喜爱; 或是仅局限于用户-项目交互的层次关联的划分, 如图 2(c) 所示通过挖掘 User1-Movie1-User4-Movie5 的交互, 预测 User1 可能对 Movie5 产生偏好. 因此, 拟解决的主要问题如下.

- 1) 节点间的复杂交互和属性信息能否作为辅助信息提高推荐预测效果, 以及该信息如何获取?
- 2) 图 1 包含了复杂的节点信息, 但预测仅利用了用户和项目的交互, 被忽略的其他类型节点交互和属性信息也会对推荐产生影响, 如何利用被忽略的信息?
- 3) 不同节点所处的层次之间也会存在关联, 这些潜在关联如何显式表达?
- 4) 在图 1 中, 推荐结果仅预测到了 User1 对 Movie1、Movie2 和 Movie5 的偏好, 如何衡量其对其他电影有无产生偏好的可能及偏好程度?

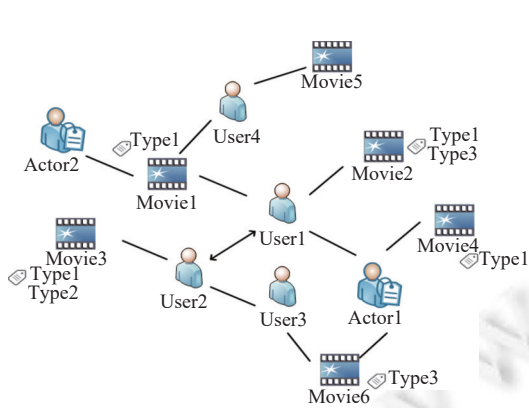


图 1 User1 的异构信息交互

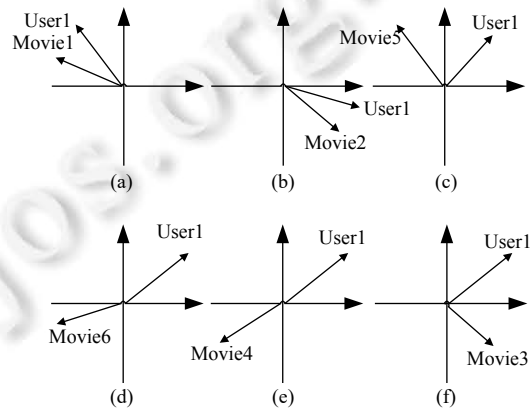


图 2 异构信息网络层次划分

### 2.2 高阶嵌入学习框架结构

将多类型的网络化数据建模为异质信息网络并对其进行信息嵌入, 能够充分利用弱关系丰富的语义信息和强



大的信息传递能力. 构建了以社交网络为基础的高阶嵌入学习框架 UI-HEHo (high-order embedded learning framework), 考虑到层次间存在的关联以及不同类型节点间的交互和属性对推荐的影响, 该框架融入了用户-用户、用户-项目以及项目-项目的高阶交互和节点属性, 并进一步将层次嵌入细分到阶间和阶内, 挖掘高阶连接状态下的信息交互给节点带来的推荐影响力, 图 3 给出了 UI-HEHo 整体示意图. 如图 2(d)-(f) 所示, 图 1 通过 UI-HEHo 挖掘 User1-User2-Movie3、User1-User2-User3-Movie6、User1-Actor1-Movie6、User1-Actor1-Movie4, 预测得到 User1 可能对 Movie3、Movie6 以及 Movie4 产生偏好; 结合节点交互和属性发现 Movie4 和 Movie6 相比 Movie3 属于 User1 的阶内嵌入, 因此对 Movie3 的偏好最小. Movie6 相比 Movie4 多了一条与 User1 的间接连接, 因此对 Movie6 的偏好程度大于 Movie4, 但偏好程度没有较大差距. 所以, User1 对这 3 个项目的预测偏好程度为 Movie6>Movie4>Movie3. 由此看出, UI-HEHo 框架能够充分利用节点信息提高推荐预测的准确性、多样性和新颖性.

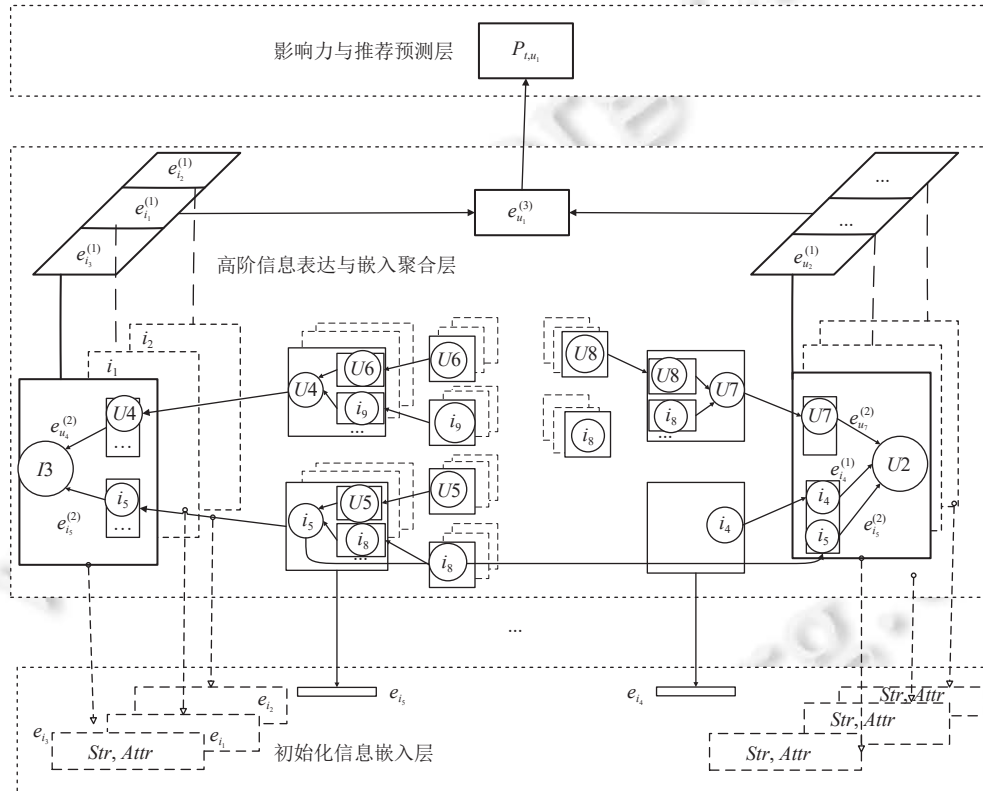


图 3 异质信息网络高阶嵌入学习框架

### 3 初始化信息嵌入层

该嵌入层利用社交网络弱关系具有的信息传递与丰富语义信息, 构建一种基于弱关系的全关系异质信息网络 (FR-HIN), 并给出一种最佳信任路径选择算法对 FR-HIN 进行路径筛选, 在降低计算量的同时避免了采样固定数量邻居带来的信息损失; 提出将多头图注意力与多任务结合, 使得注意力机制的掩码能自动确定每个任务的共享特征的重要性, 不仅学习了不同任务的权重而且筛选出节点属性中更具代表性的语义特征.

#### 3.1 全关系异质信息网络

集成弱关系的社交网络将包含大量不同类型的节点和节点间复杂的交互. 而将这些相互作用的节点建模为异质网络 HIN 是很自然的方式, 正迎合了大数据中数据多样性的显著特征. 现有异质信息网络建模由于仅将注意力

放在异质关系上, 导致其无法充分利用包含丰富信息的网络图结构. 因此, 对传统的异质网络进行拓展, 构建全关系异质信息网络.

**定义 1.** 全关系异质信息网络 (FR-HIN): 至少包含用户-项目节点对, 可增加用户对、项目对以及节点间不同类型关系的带权有向图  $G=(U, I, E(Un, Un'), E(Un, Im), E(Im, Im'), V(Un, Un'), V(Un, Im), S(Im, Im'))$ , 其中  $U$  代表用户集合,  $I$  代表  $U$  的历史项目集合,  $E(Un, Un')$ 、 $E(Im, Im')$  和  $E(Un, Im)$  分别代表用户对、项目对以及用户与项目间的信任关系的有向边,  $V(Un, Un')$ 、 $S(Im, Im')$  和  $V(Un, Im)$  为边上的信任值.

图 4 展示了一种全关系异质信息网络, 该网络中不存在直接信任关系的节点间无信任关系边, 且节点间的信任是非对称的. 其中,  $V(Un, Un')$  的取值一方面以“法定”为支柱的强制性影响形成的用户自身信任值, 如用户的地位、职业、资历等, 另一方面是因用户间的直接信息交互产生;  $V(Un, Im)$  为用户对项目的直接评级;  $S(Im, Im')$  为频繁  $N$  项集中每个项目存在条件下其余项目存在的概率;  $E(Im, Im')$  由关联项目中概率较大值指向较小值.

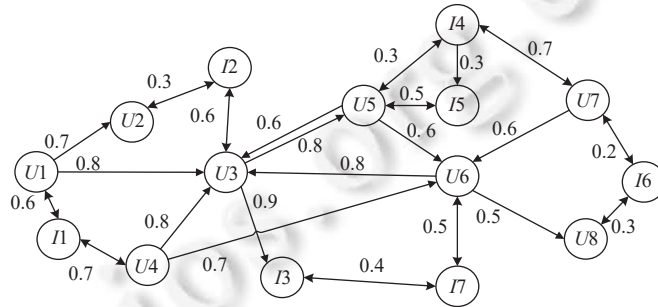


图 4 全关系异质信息网络

### 3.2 异质信息网络最佳信任路径筛选

与基于分解和随机游走的方法相比, 图神经网络能同时处理图中节点的交互和属性信息, 缓解了异质数据稀疏造成的冷启动问题. 根据图卷积模型的局部性质, 节点在第  $L$  层的特征只与其  $L$  阶子图有关, 但对于大规模的图数据 (例如: 社交网络), 子图中的节点数成指数级增长, 存在很多超级节点 (节点的度呈现幂律分布), 因此, 对目标节点的邻居节点进行采样可以降低遍历子图的时间、模型的计算以及数据的存储代价.

在异质信息网络中, 由网络中信息的传递性发现, 具有间接关系的节点间往往会包含多条连接路径. 如图 4 中, 节点  $U1$  到  $U6$  存在:  $U1 \rightarrow U2 \rightarrow I2 \rightarrow U3 \rightarrow U5 \rightarrow U6$ ,  $U1 \rightarrow U2 \rightarrow I2 \rightarrow U3 \rightarrow U5 \rightarrow I4 \rightarrow U7 \rightarrow U6, \dots, U1 \rightarrow I1 \rightarrow U4 \rightarrow U6$  共 11 条路径. 显然, 全部路径可以涵盖网络中的所有信息, 然而对待预测节点和目标节点之间的所有路径, 分别设置最大请求长度 (MTL) 和最小信任阈值 (MTT), 会大大增加处理大量节点以及节点间不同信任关系的工作量; 而使用固定长度路径又会损失有价值的节点信息. 因此, 提出 BTP 算法对 FR-HIN 的路径进行筛选, 以得到定义 2 表述的最佳信任路径, 详见算法 1.

**定义 2.** 最佳信任路径  $L$ .  $L$  满足  $R_L \ll TTL, R_{VT} \gg T_{VT}$ .  $R_L$  和  $R_{VT}$  分别表示路径  $L$  的长度和信任总值;  $TTL$  和  $T_{VT}$  分别为网络路径的最大请求长度和最小信任阈值.

---

#### 算法 1. 最佳信任路径算法 (BTP).

---

输入: 目标节点  $S_n$  的  $N$  个邻居节点集合  $T_n$ ;

输出: 最佳信任路径集合  $PATH$ .

---

算法过程:

1.  $EnQueue(Q, T_n)$ ;

2. for  $i$  in  $range(N)$ :

---

```

3.   if  $Q.front == Q.rear$ 
4.       return error;
5.   DeQueue( $Q, e$ );
6.    $ptp(Sn, e)$  /* $ptp()$  用来查找得到节点  $e$  和  $Sn$  间的所有路径*/
7.    $path \leftarrow e$ ; /* $path$  存放已构路径,  $e$  为路径的起点*/
8.    $R \leftarrow [path, Sn]$ ; /* $R$  存放节点  $e$  和  $Sn$  之间的所有路径*/
9.    $P = 0$ ; /*算法开始定义节点间路径数量为 0*/
10.   $ADJ \leftarrow getadj(adj, getLast(path))$ ; /*将已构建路径最后节点的邻居  $adj$  放入  $ADJ$  中*/
11.  for each  $adj$  in  $ADJ$ 
12.      if  $adj \in getadj(Sn)$  then /*若  $adj$  不属于终止节点邻居节点则将  $adj$  加入当前路径*/
13.           $add(adj, path)$ ;
14.           $ptp(Sn, adj)$ ;
15.      else /*否则当前路径将作为节点间的路径*/
16.           $add(adj, path)$ ;
17.           $P++$ ;
18.           $R \leftarrow [Sn, path]$ ;
19.      end for
20.  end
21.  while  $P$  /*筛选  $R$  中路径得到最佳信任路径*/
22.      if  $(R_L > AR_L \&\& R_L - AR_L \leq \beta \&\& R_{VT} \geq A_{VT})$  /* $R_L, R_{VT}$  为当前路径请求长度和总信任值,  $AR_L, A_{VT}$  为与当前路径有相同终止节点路径的平均请求长度和总信任值*/
23.           $PATH \leftarrow [Sn, R[p]]$ ;
24.      else if  $(R_{VT} < A_{VT} \&\& A_{VT} - R_{VT} \leq \gamma)$ 
25.           $PATH \leftarrow [Sn, R[p]]$ ;
26.      else
27.           $PATH \leftarrow [Sn, \Phi]$ ;
28.  end for
29.  end

```

在 FR-HIN 中利用 Personalized PageRank 算法获得算法输入的邻居节点  $Tn$ , 算法第 6–20 行获取  $Tn$  中的每个邻居节点与  $Sn$  间的所有可能路径, 算法第 21–27 行是最佳信任路径的筛选操作,  $\beta = |getLength(SP) - AR_L| \geq 0$ , 为允许的最大请求长度差值;  $\gamma = |getValue(SP) - A_{VT}| \geq 0$ , 为允许的最大信任差值;  $SP$  是所有可能路径中信任值最高的路径。

### 3.3 多头图注意力机制与多任务共享特征重要性度量

由于节点类型以及属性特征维度的不同, 对节点共享特征提取带来了较大困难. 因此, 将节点间不同类型交互看作不同的任务, 提出将多头图注意力和多任务机制结合, 确定任务间的共享特征, 实现对节点的属性表征. 在共享网络中对每个任务赋予注意力掩码, 刻画每个任务共享特征的重要性, 从而允许以一种自我监督、端到端的方式学习任务共享特征, 与具有明确任务分离的多任务体系结构相比, 学习参数更少.

将用户  $u_n$  和项目  $i_m$  通过  $d$  维信息嵌入特征向量表征为  $e_{u_n}, e_{i_m}$ , 则初始嵌入特征向量  $e_{u_n}^{(0)}, e_{i_m}^{(0)}$  分别为:

$$e_{u_n}^{(0)} = \underbrace{[V_{u_n u_n}, N_{u_n I}, V_{u_n I}, N_{u_n U}, V_{u_n U}]_{Str}}_{Attr} \underbrace{[F_{u_n}^I]_{Attr}}_{Attr} \quad (1)$$

$$e_{i_m}^{(0)} = \underbrace{[V_{i_m i_m}, N_{i_m U}, V_{i_m U}, N_{i_m I}, V_{i_m I}]_{Str}}_{Attr} \underbrace{[F_{i_m}^I]_{Attr}}_{Attr} \quad (2)$$

其中, 节点的结构特征向量  $Str$  中,  $N_{u_n I}, V_{u_n I}$  分别为  $u_n$  一阶信任项目集合  $I$  和  $u_n$  对  $I$  每个项目  $i_m$  的评级矩阵,  $N_{u_n U}, V_{u_n U}$  分别为  $u_n$  的一阶信任用户集合  $U$  和  $u_n$  对  $U$  中每个用户  $u'_n$  信任矩阵,  $N_{i_m I'}, V_{i_m I'}$  分别为  $i_m$  的一阶信任项目集合  $I$  以及  $i_m$  与  $I$  中每个项目  $i'_m$  的信任矩阵;  $V_{u_n u_n}, V_{i_m i_m}$  为节点自身信任值. 定义  $F_{u_n}^{l_u}$  和  $F_{i_m}^{l_i}$  分别表示用户和项目节点属性特征矩阵  $Attr$ .  $E = [e_{u_1}, \dots, e_{u_N}, e_{i_1}, \dots, e_{i_M}]$ ,  $e_{u_1}, \dots, e_{u_N}$  表示网络中的用户嵌入,  $e_{i_1}, \dots, e_{i_M}$  表示项目嵌入.

如图 5, 使用 *look up* 层将每个用户的  $l_u$  维和项目的  $l_i$  维属性  $attr$  分别转换为矩阵  $F_{u_n}^{l_u}$  和  $F_{i_m}^{l_i}$ , 由此得到每个用户属性嵌入向量  $attr_u = (attr_{u,1}, \dots, attr_{u,l_u})$  和项目属性嵌入向量  $attr_i = (attr_{i,1}, \dots, attr_{i,l_i})$ , 用户对、项目对以及用户-项目 3 种关系  $R$  的属性注意力权重矩阵中, 第  $p, q$  项的权重为  $\varphi_{p,q}^{(R)}$ :

$$\varphi_{p,q}^{(R)} = \begin{cases} F(attr_{u,p})^T W F(attr_{i,q}), & \text{if } R = u-i \\ F(attr_{i,p})^T W F(attr_{i,q}), & \text{if } R = i-i \\ F(attr_{u,p})^T W F(attr_{u,q}), & \text{if } R = u-u \end{cases} \quad (3)$$

其中,  $W \in A^{l_u \times l_i}$  为属性间的权重矩阵,  $F$  为具有  $l_F$  层的前馈神经网络. 池化是为了选择与所有属性相关性最大的属性. 用户对、项目对, 用户-项目节点对的属性选择分别为:

$$\begin{cases} a_p = \max_{q=1, \dots, l_u} \varphi_{p,q}^{(u-i)} \\ b_p = \max_{q=1, \dots, l_i} \varphi_{p,q}^{(i-i)} \\ c_p = \max_{q=1, \dots, l_i} \varphi_{p,q}^{(u-i)} \\ d_q = \max_{p=1, \dots, l_u} \varphi_{p,q}^{(u-i)} \end{cases} \quad (4)$$

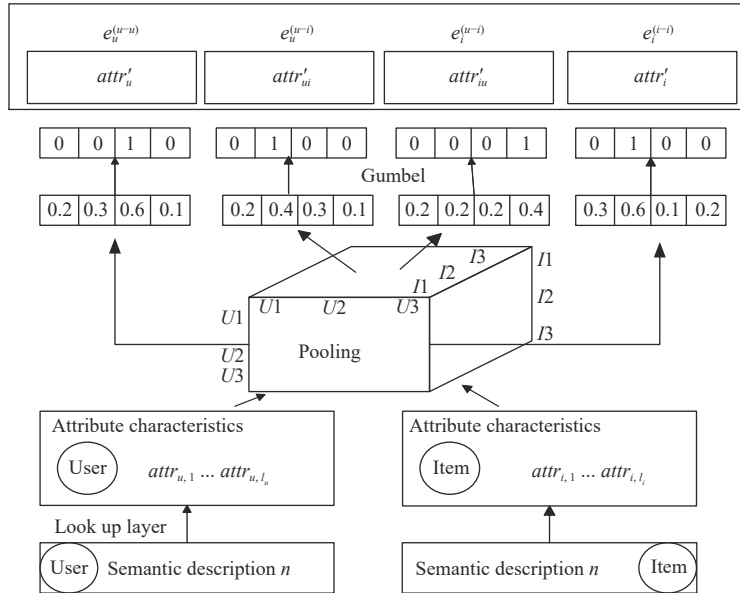


图 5 多头图注意力

将非标准化向量  $a = (a_1, a_2, \dots, a_{l_u})$ 、 $b = (b_1, b_2, \dots, b_{l_i})$ 、 $c = (c_1, c_2, \dots, c_{l_u})$  和  $d = (d_1, d_2, \dots, d_{l_i})$  通过 Gumbel-Softmax<sup>[41]</sup> 函数转换为概率分布, 则被选中的用户、项目以及用户与项目间属性分别为:

$$\begin{cases} attr'_u = (\text{Gumbel}(a))^T attr_u \\ attr'_i = (\text{Gumbel}(b))^T attr_i \\ attr'_{ui} = (\text{Gumbel}(c))^T attr_{u,i} \\ attr'_{iu} = (\text{Gumbel}(d))^T attr_{i,u} \end{cases} \quad (5)$$

图数据中多类型的节点形成了节点间的复杂关系, 不同关系间也存在着信息传递. 若要学习所有关系间的权



重  $W_R$ , 将导致网络模型所要学习的参数量将十分庞大的. 而对于不常见的关系,  $W_R$  所对应的学习数据很少, 容易造成过拟合现象. 为了解决这个问题, 本文把不同类型关系看作不同的任务, 将关系分为用户-用户、项目-项目、用户-项目 3 类任务, 并实现任务间  $W_R$  的共享, 进而得到任务间的共享特征 (见图 6). 将多任务学习机制中的 MMoE 模型转化为基本构建块 MMoE 层嵌入到图神经网络中, 为不同的任务分配不同的  $W_R$ , 并为每个任务都配备一个 Gate<sup>[42]</sup>. 针对不同的任务, 特定 Gate 的输出表示不同的 Expert 被选择的概率, 将多个 Expert 加权求和得到  $f_{(x)}^k$ , 并输出给特定的 Tower 模型, 用于最终每个任务  $k$  的输出  $y_k: y_k = h^k(f_{(x)}^k)$ . 其中,  $f_{(x)}^k = \sum_{i=1}^n g_i^k(x) f_i(x)$ ,  $g^k(x) = \text{Softmax}(W_{gk}x)$ ,  $w_{gk} \in R^{n \times d}$  是可训练矩阵,  $x$  为输入特征,  $n=3$  为 expert 的个数,  $d$  是输入的特征维度.  $g^k(x)$  为 Gate<sup>k</sup> 的输出, 即每个 expert 的权重.

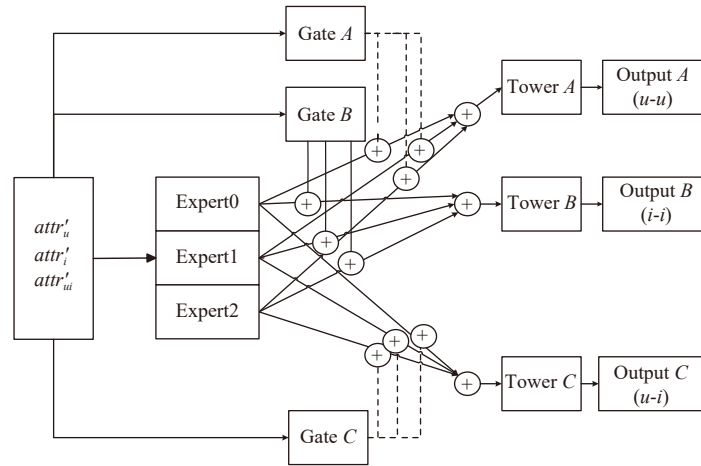


图 6 多任务共享特征

#### 4 高阶信息表达与嵌入聚合层

FR-HIN 是一类具有高阶连接的信息交互网络, 与给定节点直接或间接连接的节点均会对该节点产生推荐影响, 图 7(a) 展示了以节点  $U1$  为目标节点的 FR-HIN. 同阶层节点即便类型以及与其交互的节点的信息和个数不同, 也可能存在交互. 每个节点对目标节点的推荐影响力是不同的, 其本身也形成了带有高阶交互信息的小型 FR-HIN. 因此, 区分节点类型以及计算不同阶层上节点对目标节点的推荐影响力会产生很大代价.

该层使用多头图注意力机制, 实现对不同邻居权重的自适应分配, 将同一阶层的节点放入一个信息片中, 信息片间依据阶间节点的交互设置连接, 聚合邻居特征, 更新当前节点的特征, 通过网络的层级传播机制将节点的特征信息

由远及近地聚合到待预测节点上. 多头图注意力机制更新节点特征向量的公式一般为  $e'_i = \frac{1}{\|k=1\|} \sigma \left( \sum_{v_j \in \tilde{N}(v_i)} \alpha_{ij}^{(k)} W^{(k)} e_j \right)$ <sup>[43]</sup>.

为了对不同节点间不同类型的关系区分, 本文将公式中的  $W$  进一步细化, 设置邻居节点  $S'$  到目标节点为  $S$  的权重系数为  $e_{SS'} = \sigma(\partial W(S, S'))$ ,  $W'$  为节点特征变换的权重参数, 定义  $\partial(\cdot)$  (节点相关度函数) 为 2 个节点向量的哈达玛积, 使嵌入信息依赖于节点之间的交互密切程度. 激活函数  $\sigma(\cdot)$  为  $LeakyReLU$  则  $e_{SS'} = LeakyReLU(W'(e_S \circ e_{S'}))$ , 为了更好地分配权重, 对  $e_{SS'}$  通过  $\frac{1}{\sqrt{N_S}}$  进行了归一化处理,  $N_S$  代表节点  $S$  的一阶邻居数量, 从表示学习的角度看, 该系数分别为项目对用户偏好程度的贡献和用户对用户信任程度的贡献. 则  $e_{SS'}$  以及对应更新后的节点特征向量  $e_{SS'}$  分别为:

$$e_{SS'} = \frac{LeakyReLU(W'(e_S \circ e_{S'}))}{\sqrt{N_S}}, e'_i = \frac{1}{\|k=1\|} \sigma \left( \sum_{v_j \in \tilde{N}(v_i)} \alpha_{ij}^{(k)} e_{SS'}^{(k)} e_j \right) \quad (6)$$

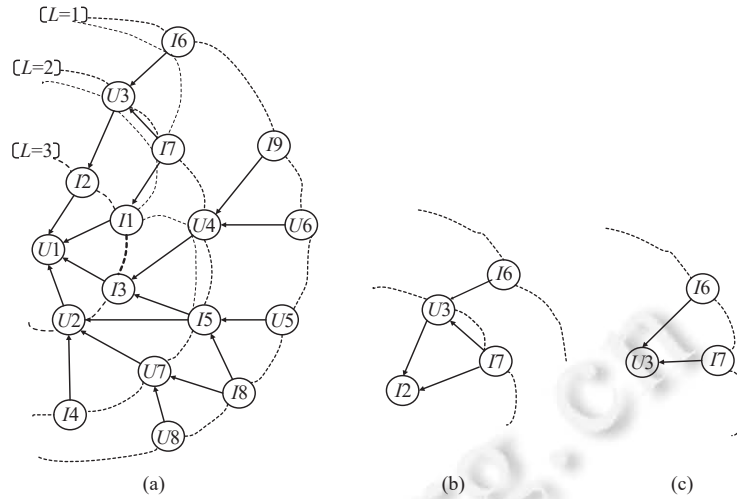


图 7 以  $U1$  为目标节点的 FR-HIN

如图 7(a) 中节点  $I2$  既是  $U1$  的一阶邻居节点, 也形成了以其为目标节点的子 FR-HIN (如图 7(b)). 同样的, 节点  $U3$  和  $I7$  既是  $I2$  的一阶邻居, 也形成了以其为目标节点的子 FR-HIN (如图 7(c)).  $U3$  的一阶邻居为  $I7$  和  $I6$ ,  $I7$  的一阶邻居为  $I1$  和  $U3$ , 值得注意的是,  $I7$  和  $U3$  属于  $I2$  的同一阶层,  $I7$  既作是  $U3$  的一阶邻居又是  $I2$  的一阶邻居, 其产生的推荐影响力肯定强于  $I6$ . 考虑到共享特征对推荐的影响是不同的, 而且同类型节点和不同类型节点间推荐影响也不同, 因此, 本文将高阶信息嵌入聚合分为阶间与阶内节点特征向量聚合两个过程.

(1) 阶间节点特征向量嵌入聚合

如图 4 信息聚合路径:  $U1 \leftarrow U2 \leftarrow I2 \leftarrow U3 \leftarrow I3 \leftarrow I7$  和  $U1 \leftarrow I1 \leftarrow U4 \leftarrow U6 \leftarrow I7$ , 其中  $I2$  对  $U2$  同  $U4$  对  $I1$  进行的特征向量嵌入聚合操作是不同的, 以用户  $u_n$  作为路径目标嵌入节点进行信息嵌入聚合, 因为图中存在连接的用户-项目节点对  $(u_n, i_m)$ 、用户-用户节点对  $(u_n, u'_n)$ , 所以对  $e_{SS'}$  作了划分,  $e_{SS'_U}$  是目标节点的邻居用户对应的权重参数,  $e_{SS'_I}$  是目标节点的邻居项目对应的权重参数. 定义信息嵌入聚合函数  $u_n$  为:

$$e_{u_n}^{(L)} = \begin{cases} \|\|_{k=1}^K \sigma \left( \sum_{u'_n \in U} \alpha_{u_n, u'_n}^{(k)} e_{SS'_U}^{(k)} e_{u'_n}^{(L-1)} \right), & \text{if } (u-u) \\ \|\|_{k=1}^K \sigma \left( \sum_{i_m \in I} \alpha_{u_n, i_m}^{(k)} e_{SS'_I}^{(k)} e_{i_m}^{(L-1)} \right), & \text{if } (u-i) \end{cases} \quad (7)$$

$\alpha_{(i)}^{(k)}$  和  $e_{(i)}^{(k)}$  分别是第  $k$  组注意力机制得到的权重系数和对应的学习参数. 同样地, 对于图中存在连接的用户-项目节点对  $(i_m, u_n)$ 、项目-项目节点对集合  $(i_m, i'_m)$  定义  $i_m$  的一阶信息嵌入聚合函数为:

$$e_{i_m}^{(L)} = \begin{cases} \|\|_{k=1}^K \sigma \left( \sum_{u'_n \in U} \alpha_{i_m, u'_n}^{(k)} e_{SS'_U}^{(k)} e_{u'_n}^{(L-1)} \right), & \text{if } (u-i) \\ \|\|_{k=1}^K \sigma \left( \sum_{i'_m \in I} \alpha_{i_m, i'_m}^{(k)} e_{SS'_I}^{(k)} e_{i'_m}^{(L-1)} \right), & \text{if } (i-i) \end{cases} \quad (8)$$

(2) 阶内节点信息聚合

同样,  $U1 \leftarrow I2 \leftarrow U3 \leftarrow I6$  和  $U1 \leftarrow U2 \leftarrow U7 \leftarrow I8$  两条路径,  $I2$  和  $U2$  均作为  $U1$  的一阶邻居属于同一阶都对  $U1$  产生了推荐影响, 因此, 提出对于同一阶层内的节点进行一个特征向量的聚合操作, 考虑到减少输出特征向量的维度, 我们将平均操作代替拼接操作, 而且注意力机制的分组  $K$  由当前节点间的关系决定. 用户节点、项目节点的层内聚合分别为:

$$e_{u_n}^{(L)} = \sigma \left( \frac{1}{K_{u \leftarrow u'}} \sum_{k=1}^{K_{u \leftarrow u'}} \sum_{u'_n \in U} \alpha_{u_n, u'_n}^{(k)} e_{SS_U}^{(k)} e_{u'_n}^{(L)} + \frac{1}{K_{i \leftarrow i'}} \sum_{k=1}^{K_{i \leftarrow i'}} \sum_{i_m \in I} \alpha_{u_n, i_m}^{(k)} e_{SS_I}^{(k)} e_{i_m}^{(L)} \right) \quad (9)$$

$$e_{i_m}^{(L)} = \sigma \left( \frac{1}{K_{i \leftarrow u}} \sum_{k=1}^{K_{i \leftarrow u}} \sum_{u_n \in U} \alpha_{i_m, u_n}^{(k)} e_{SS_U}^{(k)} e_{u_n}^{(L)} + \frac{1}{K_{i \leftarrow i'}} \sum_{k=1}^{K_{i \leftarrow i'}} \sum_{i_m \in I} \alpha_{i_m, i'_m}^{(k)} e_{SS_I}^{(k)} e_{i'_m}^{(L)} \right) \quad (10)$$

算法 2 描述了 FR-HIN 中带预测节点的高阶信息嵌入聚合过程。

**算法 2.** 高阶信息嵌入聚合算法 (high order information embedding aggregation).

1.  $List \leftarrow DFS(\text{FR-HIN})$  /\*遍历 FR-HIN 中的节点存入列表  $List$ \*/
2. for  $B$  in  $List$ :
3.  $L \leftarrow BFS(B)$  /\*广度遍历节点  $B$ , 得到  $B$  的深度阶数  $L$ \*/
4. for  $l=L...1$  do
5.  $B^{(l)} \leftarrow B$  /\*得到  $B$  的  $l$  阶子 FR-HIN
6. end for
7. end for
8. for  $l=L...1$  do
9. for  $v \in B^{(l)}$  do
10. if  $v^{(l)} \leftarrow v$  /\*如果节点  $v$  有  $l$  阶子 FR-HIN, 进行阶间嵌入聚合\*/
11. for  $l=l...1$  do
12. if  $v \in user$
13.  $v = u_n$
14. 执行公式 (7)
15. if  $v \in item$
16. 执行公式 (8)
17. else /\*阶内嵌入聚合\*/
18. 执行公式 (9)、公式 (10)
19. end for
20. end for
21. end for
22. end

## 5 影响力与推荐预测层

该预测层给出了一种推荐影响力算法, 计算待预测节点对目标节点的推荐影响力来完成推荐。

设  $P_{t,f}$  为待预测节点  $f$  对  $t$  目标节点的推荐影响值, 其取值为  $f$  与  $t$  高阶信息嵌入聚合值的内积, 意为带预测节点特征向量在目标节点特征向量上的投影, 值越大表明两节点的关联性越强,  $f$  对  $t$  的推荐影响力也就越大。值得注意的是, 若  $f$  是  $t$  的一阶节点, 则推荐影响值为节点间高阶信息嵌入聚合的内积; 若  $f$  是  $t$  的高阶节点, 则推荐影响值等于  $f, t$  的高阶信息嵌入聚合内积与由  $f$  到  $t$  路径上的节点间的信任值的乘积。  $P_{t,f}$  的形式化描述如下:

$$P_{t,f} = \begin{cases} e_f^{(L)\top} e_t^{(L)}, & \text{if } l = 1 \\ e_f^{(L)\top} e_t^{(L)} \cdot \prod_{b=1}^n V_b, & \text{if } l > 1 \end{cases} \quad (11)$$

筛选  $P_{t,f}$  取值高的  $N$  个待预测节点进行 top- $N$  推荐, 若带预测节点为项目节点且不存在于用户的历史项目列

表中, 则直接推荐给目标节点, 若待预测节点是用户节点, 则将以该用户为目标节点, 得到其推荐影响值最高, 且不属于  $N$  个待预测节点的项目推荐给目标用户. 算法 3 描述了详细的推荐影响力算法生成推荐的过程.

**算法 3.** 推荐影响力算法 (influence recommendation).

输入: 待预测节点的高阶信息嵌入传播值  $e_f^{(L)}$ ,  $N_t$  为目标节点  $t$  的一阶节点集合,  $n, V_b$  分别为  $f$  到  $t$  的传播路径长度以及路径上节点间的信任值, 目标节点的历史项目列表  $t_{hist}$ , 待预测用户节点的历史项目列表  $f_{hist}^u$ ;

输出:  $P_{t,f}$ , 推荐列表  $List$ .

1.  $IRec(t, f)$
2. if  $f \in N_t$
3.  $P_{t,f} = e_t^{(L)\top} e_f^{(L)}$
4. else
5.  $P_{t,f} = e_t^{(L)\top} e_f^{(L)} \cdot \prod_{b=1}^n V_b$
6. for  $f$  in  $list \leftarrow \text{top-}N(P_{t,f})$ :
7. if  $f$  is item and  $f \notin t_{hist}$
8.  $List \leftarrow f$
9. else
10.  $IRec(f, f')$
11. end

**6 实验结果及分析**

为了实验验证 UI-HEHO 框架的有效性, 采用真实数据集, 并选取目前具有代表性的主流推荐方法作为实验对比算法. 同时, 为了进一步验证 UI-HEHO 框架能更充分利用异质网络中的有价值信息, 而且能够帮助预测用户对项目的评级, 实验在具有评级预测目的的文献 [31] 中, 将用户和项目嵌入替换为 UI-HEHO 框架生成的用户嵌入  $e_u^{(L)}$  和项目嵌入  $e_i^{(L)}$ , 则有评级预测公式:  $\hat{R}_{u,i} = x_u^T \cdot y_i + \alpha \cdot e_u^{(L)\top} \cdot \gamma_i^{(L)} + \beta \cdot \gamma_u^{(L)} \cdot e_i^{(L)}$ , 其中,  $\hat{R}_{u,i}$  为用户对项目的预测评级,  $\gamma_i^{(L)}, \gamma_u^{(L)}$  分别为用户和项目的潜在因子,  $\alpha, \beta$  为调谐参数,  $x_u^T, y_i$  为矩阵分解模型中用户和项目的潜在因素.

**6.1 数据集**

为了验证 UI-HEHO 框架能有效弥补数据稀疏导致评级预测准确度较低的问题, 采用 Douban Movie、Douban Book 和 Yelp 这 3 种不同领域和稀疏度 (Yelp 数据集最稀疏, Douban Movie 数据集最密集) 数据集验证评级预测, 且均包含了用户对、项目对以及用户和项目的交互记录, 其详细描述见后文表 1. 采用含有属性及评论等语义信息的数据集 Gowalla、Amazon-Books、Yelp2018、serendipity-sac2018<sup>[44]</sup> 和 Douban Movie-2019 (过滤掉其少于 10 次交互的用户和项目, 具体数据描述见后文表 2), 验证了 UI-HEHO 框架的有效性.

**6.2 评级预测**

利用第 6.1 节中表 1 数据集, 对评级预测效果进行评价, 采用平均绝对误差 MAE 和均方根误差 RMSE 作为评价指标. 对 Double Movie 和 Book 数据集, 设置 {80%、60%、40%、20%} 这 4 个训练比率; 对于稀疏数据集 Yelp, 设置更高的训练比为 {90%、80%、70%、60%}. 对于每个比率, 随机生成 10 个评估集并取结果的平均值作为最终表现. 实验对比算法为:

- (1) PMF<sup>[14]</sup>: 核心思想为用户对项目的偏好可以由较少的几个因素的线性组合决定.
- (2) SemRec<sup>[16]</sup>: 基于加权异质信息网络的协同过滤方法, 使用带权或不带权元路径计算用户间的相似度进行评级预测.
- (3) HERec\_mp<sup>[30]</sup>: 结合了 metapath2vec++ 的异构网络嵌入.

(4) HERec\_spl<sup>[31]</sup>: 为了嵌入 HIN, 基于元路径的随机游走策略生成节点序列, 用一组融合函数对学习到的节点嵌入进行变换并将其集成到扩展的矩阵分解模型中, 对扩展的 MF 模型和融合函数进行了联合优化用于预测任务.

表 1 评级预测任务数据集

数据集 (density)	关系	Entity1	Entity2	Meta-paths
Douban Movie (0.63%)	User-Movie (1 068 278)	User (13 367)	Movie (12 677)	UMU, MUM
	User1-User2 (4 085)	User1 (2 440)	User2 (2 294)	UMDMU
	User-Group (570 047)	User (13 337)	Group (570 047)	MDM, MAM
	Movie-Director (11 276)	Movie (10 179)	Director (2 449)	UMAMU
	Movie-Actor (33 587)	Movie (11 718)	Actor (6 311)	UMTMU
	Movie-Type (27 668)	Movie (12 678)	Type (38)	MTM
Douban Book (0.27%)	User1-User2 (169 150)	User1 (12 748)	User2 (12 748)	UBU, BUB
	User-Book (792 062)	User (13 024)	Book (22 347)	UBPB, BPB
	Book-Author (21 907)	Book (21 907)	Author (10 805)	UBYBU, BYB
	Book-Publisher (21 907)	Book (21 773)	Publisher (1 815)	UBABU
Yelp (0.08%)	Book-Year (21 192)	User (21 192)	Group (64)	
	User-Business (198 397)	User (16 239)	Business (14 284)	UBU, BUB
	User1-User2 (158 590)	User1 (10 580)	User2 (10 580)	UBCiBU
	User-Compliment (76 875)	Use (14 411)	Compliment (11)	BCiB, BCaB
	Business-City (14 267)	Business (14 267)	City (47)	UBCaBU,
	Business-Category (40 009)	Business (14 180)	Category (511)	

表 2 推荐生成任务数据集

数据集	关系	User	Item	Text Semantic
Gowalla	1 027 370	29 858	40 981	签到时间和地点经纬度
Yelp2018	1 666 869	31 831	40 841	21种酒店属性信息以及用户对酒店的评论
Amazon-Book	2 984 108	52 643	91 599	图书的语义描述等5种属性
serendipity-sac2018	2 150	481	1 678	有关项目的9个问题以及7种不同的答案
Douban Movie2019	1 048 576	278 297	21 359	21 359部电影的278 297条评论和140 502部电影内容概述

PMF 是经典的基于矩阵分解的方法; SemRec、HERec\_mp 和 HERec\_spl 是基于元路径相似性的 HIN 方法, 表 2 给出了对应数据集的元路径信息. 元路径过长会带来噪声的影响, 因此元路径的长度设置为 $\leq 4$ . 为了公平起见, 实验均采用提出的 BTP 算法实现对输入数据的节点筛选, 通过 UI-HEHo 框架得出的评级预测公式的参数学习遵循文献 [31] 的实验设置, 其他参数采用原始最佳设置. 实验结果见表 3.

表 3 展示了预测评级实验效果, 结果表明引入异质信息的基于 HIN 的方法 (SemRec、HERec\_mp 和 HERec\_spl) 比基于传统 MF 的方法 (PMF) 表现更好. SemRec 根据相似用户进行评级预测, 过于依靠历史数据, 忽略了 HIN 环境下信息嵌入对节点的影响, 相比其他两种基于 HIN 的方法效果欠佳; 与 HERec\_mp 将所有类型的实体映射到相同的表示空间中不同, HERec\_spl 方法将异构网络信息转换为异构邻域学习用户和项目的嵌入, 其评级预测效果更优, 表明了执行特定类型的信息嵌入在 HIN 推荐的重要性. UI-HEHo 框架生成的用户和项目嵌入在评级预测方面总体表现是最好的, 但在较为密集的双电影和双书数据集中, 尤其是在 20%–40% 数据量的情况下, UI-HEHo 表现次优, 其主要原因是 UI-HEHo 框架增加了辅助评级信息, 包括: 在初始信息嵌入模块中, 增加了节点及其交互信息, 可提取多类型节点间的共享语义特征; 在高阶信息嵌入聚合模块中, 捕获了节点在整个异质信息网络中的隐式交互. 但在密集数据集中, 由于已包含了丰富的评级信息, 辅助评级信息的增加会致使原真实评级信息损失, 从而偏离了真实评级. 而对于稀疏数据集 Yelp, UI-HEHo 则通过辅助评级信息可有效缓解数据不足造成的冷启动问题.



表 3 评级预测实验结果 (粗体为最佳性能取值)

数据集	Training (%)	Metrics	PMF	SemRec	HERec_mp	HERec_spl	UI-HEHo
Douban Movie	80	MAE	0.5741	0.5695	0.5515	0.5519	<b>0.5492</b>
		RMSE	0.7641	0.7399	0.7121	0.7053	<b>0.7024</b>
	60	MAE	0.5867	0.5738	0.5611	0.5587	<b>0.5543</b>
		RMSE	0.7891	0.7551	0.7264	0.7148	<b>0.7086</b>
	40	MAE	0.6078	0.5945	0.5747	0.5699	<b>0.5654</b>
		RMSE	0.8321	0.7836	0.7429	0.7315	<b>0.7296</b>
	20	MAE	0.7247	0.6392	0.6063	<b>0.5900</b>	0.6276
		RMSE	0.9440	0.8599	0.7877	<b>0.7660</b>	0.7790
Douban Book	80	MAE	0.5774	0.5675	0.5519	0.5502	<b>0.5379</b>
		RMSE	0.7414	0.7283	0.7081	0.6811	<b>0.6673</b>
	60	MAE	0.6065	0.5833	0.5666	<b>0.5600</b>	0.6024
		RMSE	0.7908	0.7505	0.7318	<b>0.7123</b>	0.7263
	40	MAE	0.6800	0.6025	0.5954	<b>0.5774</b>	0.5818
		RMSE	0.9203	0.7751	0.7703	0.7400	<b>0.7258</b>
	20	MAE	1.0344	0.6481	0.6785	0.6450	<b>0.6386</b>
		RMSE	1.4414	<b>0.8350</b>	0.8869	0.8581	0.8561
Yelp	90	MAE	1.0412	0.9043	0.8822	0.8395	<b>0.8319</b>
		RMSE	1.4268	1.1637	1.1309	1.0907	<b>1.0802</b>
	80	MAE	1.0791	0.9176	0.8953	<b>0.8475</b>	0.8553
		RMSE	1.4816	1.1771	1.1516	1.1117	<b>1.1105</b>
	70	MAE	1.1170	0.9407	0.9043	0.8580	<b>0.8497</b>
		RMSE	1.5387	1.2108	1.1639	1.1256	<b>1.1232</b>
	60	MAE	1.1778	0.9637	0.9257	<b>0.8759</b>	0.8785
		RMSE	1.6167	1.2380	1.1887	1.1488	<b>1.1389</b>

### 6.3 生成推荐

采用表 2 所示的数据集验证和评估了该学习框架推荐结果的有效性、多样性和新颖性, 并随机选择每个数据集 80% 的历史交互数据构成训练集, 其余作为测试集; 将观察到的用户-项目交互视为积极实例, 通过负采样将无历史交互的用户-项目作为消极实例. 对比算法描述如下.

(1) GC-MC<sup>[33]</sup>: 基于 user-item 二部图的图自编码器框架, 只考虑一阶邻居.

(2) CMN<sup>[36]</sup>: 以非线性方式利用潜在因子模型的全局结构和邻域模型局部结构的优势将记忆组件和神经注意机制融合为邻域组件. 关联寻址用户和物品存储器在记忆模块中编码复杂的用户项目关系, 与神经注意机制相结合, 以学习用户物品特定的邻域, 值得注意的是一阶连接用于查找与相同项目交互的相似用户.

(3) PinSage<sup>[38]</sup>: 被设计为在项目-项目图中使用 GraphSAGE. 本文将其应用于用户-项目交互图.

(4) NGCF<sup>[40]</sup>: 通过传播嵌入利用用户项目图结构对图中的高阶连通性表达建模, 有效地以显式方式将协作信号注入嵌入过程.

#### • 参数学习

所有方法的嵌入大小固定为 64, 批大小固定为 1024, 使用调参函数 grid\_search 在 {0.0001, 0.001, 0.01, 0.1} 和 {10<sup>-5</sup>, 10<sup>-4</sup>, ..., 10<sup>-1</sup>} 中分别选择学习率和 L2 正则化. 从小批量中中选择唯一的用户  $u$  和项目节点  $i = \{i_m, i_n\}$ , 然后在信息融合后获得低维初始信息嵌入  $\{e_u^{(0)}, e_{i_m}^{(0)}, e_{i_n}^{(0)}\}$ , 对广泛使用的贝叶斯个性化推荐 (BPR) 损失进行随机梯度下降以优化推荐模型. 目标函数如下:

$$loss = \sum_{(u, i_m, i_n) \in o} -\log \sigma(e_u^{(L)} \cdot e_{i_m}^{(L)} - e_u^{(L)} \cdot e_{i_n}^{(L)}) + \lambda \|\Theta\|_2^2 + \beta (\|e_u^{(L)}\|_2^2 + \|e_{i_m}^{(L)}\|_2^2 + \|e_{i_n}^{(L)}\|_2^2) \quad (12)$$

其中,  $o = \{(u, i_m, i_n) | (u, i_m) \in R^+, (u, i_n) \in R^-\}$  表示一个训练样本,  $R^+, R^-$  分别表示观察到的和未观察到的用户项目交

互,  $\Theta = \{W, \alpha_{u_n, u_n}^{(k)}, \alpha_{u_n, i_m}^{(k)}, e_{SS'_U}^{(k)}, e_{SS'_I}^{(k)}\}$  为模型参数集,  $e_u^{(L)}, e_{i_m}^{(L)}, e_{i_n}^{(L)}$  为学习到的用户项目嵌入. 为了防止过度拟合, 设置正则化系数  $\lambda = 0.01, \beta = 0.02$  对  $\Theta$  和生成的嵌入进行正则化.

- 推荐有效性评估

该组实验采用召回率 Recall 和归一化折损累计增益 NDCG 指标评价推荐结果的有效性, 实验结果见表 4.

表 4 推荐有效性 (粗体为最佳性能取值)

模型	Gowalla		Yelp2018		Amazon-Book		serendipity-sac2018		Douban Movie2019	
	Recall	NDCG	Recall	NDCG	Recall	NDCG	Recall	NDCG	Recall	NDCG
GC-MaC	0.1392	0.1977	0.0365	0.0812	0.0310	0.0561	0.1003	0.0553	<b>0.0797</b>	0.0776
CMN	0.1404	0.2129	0.0364	0.0745	0.0299	0.0556	0.1007	<b>0.0574</b>	0.0664	0.0794
PinSage	0.1380	0.1947	0.0372	0.0803	0.0283	0.0545	0.0983	0.0489	0.0728	0.0818
NGCF	0.1531	0.2067	<b>0.0538</b>	0.0825	0.0339	0.0637	0.0849	0.0453	0.0678	0.0759
UI-HEHo	<b>0.1601</b>	<b>0.2211</b>	0.0472	<b>0.0953</b>	<b>0.0359</b>	<b>0.0651</b>	<b>0.1032</b>	0.0541	0.0789	<b>0.1092</b>

在表 4 中, GC-MC 和 CMN 均引入一阶邻居对网络中的节点表示学习, 但由于 GC-MC 使用相同或启发式权重定义相邻用户权重, 而 CMN 采用注意机制来指定权重, 从而使得 CMN 较 GC-MC 表现更优. 在 Yelp2018 数据集上, PinSage 和 NGCF 的性能提升相对明显, 这是因为 PinSage 在嵌入函数中引入了高阶连通性, NGCF 也通过高阶邻居信息丰富了用户项目交互数据, 而 CMN 只考虑一阶相似的用户, 这体现了高阶信息传播对信息嵌入及对邻居建模的积极作用. UI-HEHo 虽然因为引入大量隐式关系导致推荐效果次优, 但在数据集上整体表现是最佳的, 特别是在 serendipity-sac2018 和 Douban Movie2019 数据集上的召回率较 NGCF 分别提高了 21.6% 和 16.37%, 因其利用多头图注意力机制的多任务共享特征重要性度量因子, 充分利用了原数据集语义等辅助信息, 丰富了异质网络中节点间的交互信息, 并通过信息的层次传播对各阶层网络中的交互信息聚合, 以显示的方式表达了网络中的高阶连通性, 提高了推荐性能.

- 推荐多样性和新颖性评估

推荐的多样性要求推荐列表能覆盖用户不同的兴趣领域, 同时能够满足用户的主要兴趣, 现有的推荐多样性指标  $ILS$  公式仅用推荐列表中物品间的相似度来衡量, 没有考虑用户的主要兴趣, 因此, 对传统推荐多样性指标  $ILS$  做了如下改进:

$$ILS'(L_{Rec}) = \frac{2N_1}{N(N-1)} \sum_{b_i \in L_{Rec}} \sum_{b_j \in L_{Rec}, (b_i \neq b_j) \in L_{His}} S(b_i, b_j) + \frac{2N}{N(N-1)} \sum_{b_i \in L_{Rec}} \sum_{b_j \in L_{His}} S(b_i, b_j) \quad (13)$$

其中,  $L_{Rec}$  为推荐列表种类,  $N$  为推荐列表长度,  $L_{His}$  为用户历史列表种类,  $S(b_i, b_j)$  表示项目间的相似度,  $N_1$  为  $L_{Rec}$  与  $L_{His}$  中不同类型的项目个数.  $ILS'$  反映了项目间的相似度, 因此  $ILS'$  值越小表示项目间差异越大, 推荐的多样性也就越好.

图 8 展示了在表 2 Yelp2018、serendipity-sac2018 和 Douban Movie2019 数据集上的实验结果. UI-HEHo 对应的  $ILS'$  值最小, 即多样性表现最好, 其主要原因是 UI-HEHo 不仅引入语义信息作为不同类型节点交互的共享特征, 还增加了项目对和用户对的高阶交互, 使无直接交互节点间的隐式关联更易被发现, 推荐涉及的项目类型更丰富.

$ILS'$  虽然能说明推荐列表的多样性, 但不能明确表示用户的主次兴趣以及新推荐兴趣在列表中的占比. 因此, 采用具有明确类型信息的 3 个数据集 Yelp2018、serendipity-sac2018 和 Douban Movie2019 对推荐多样性类型进行评估, 随机统计 20 个用户在每种方法下生成的 top-20 推荐项目列表及类型 (项目类型取各个类型数量的平均值), 评估结果见图 8-图 11. 其中, 横坐标表示项目类型, 纵坐标代表每种方法生成的对应类型的项目数量. 参照数据集 Douban Movie2019 原本内容, 图 11 项目类型用数字表示.

从实验结果可以看出, UI-HEHo 生成的列表涵盖的项目类型最丰富, 对用户的主次要兴趣以及可能兴趣的推荐占比较为平均. 而 GC-MC、CMN 和 PinSage 只考虑一阶邻居节点的影响, 使推荐类型过于突出用户主要兴趣, 推荐的类型相似, 导致了高  $ILS'$  值. NGCF 的  $ILS'$  评估和类型的统计评估仅次于 UI-HEHo, 因为其虽考虑了用户项目的高阶连通性, 但忽略了源数据中的语义等可作为共享特征辅助推荐, 也未考虑用户间以及项目存在的高阶信息交互.

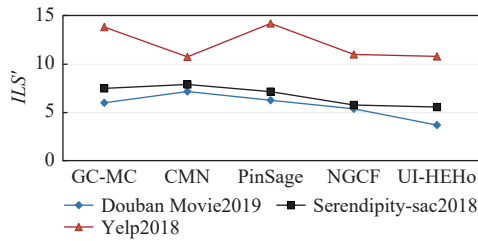


图 8 推荐多样性 ILS' 指标评估

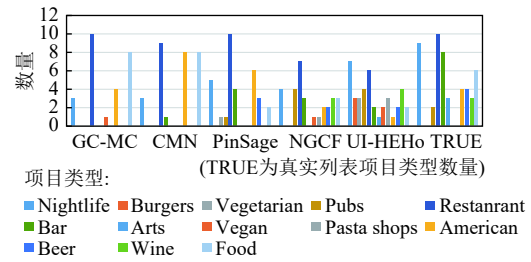


图 9 推荐多样性类型评估 (Yelp2018)

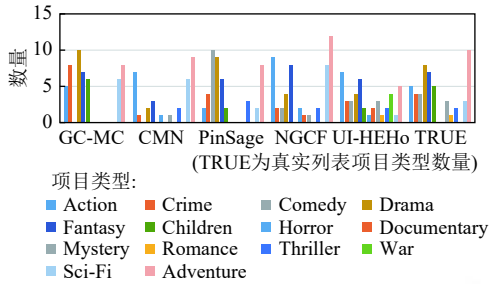


图 10 推荐多样性类型评估 (serendipity-sac2018)

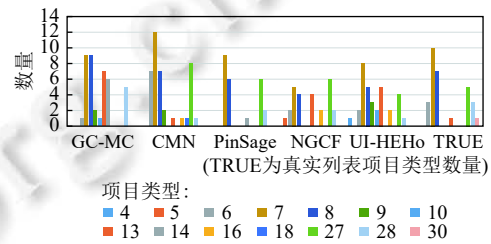


图 11 推荐多样性类型评估 (Douban Movie2019)

一般而言, 推荐的新颖性要求尽量推荐给用户没有看过的内容, 现有的新颖性评价多采用相似度和平均流行度进行度量<sup>[44]</sup>, 但均会导致推荐列表忽视用户主要兴趣, 降低了用户体验的满意度. 本文亦引入了相似度和流行度概念, 区别在于: 在计算流行度时, 从推荐列表中过滤掉目标用户的历史项目和历史相似项目, 考虑了用户的主要兴趣; 同时, 为了避免列表中“冷门”项目过多导致流行度评价差异化弱, 进一步利用用户历史项目流行度进行约束. 因此, 定义了公式 (14) 所示的一种新颖性评价指标.

$$Nov = \frac{Pop\left(\sum_{i=1}^N His_{(i)}\right)}{Pop\left(\sum_{i=1}^{N'} [Rec - His_{(i)} - SRH_{(i)}]\right)} \quad (14)$$

其中,  $Pop(\cdot)$  表示项目在当前数据中的流行度 (即被使用过的记录条数),  $Rec$  表示推荐列表,  $His_{(i)}$  表示目标用户的历史项目,  $SRH_{(i)}$  表示与目标用户历史项目相似的项目集合,  $N$  和  $N'$  分别表示推荐列表的项目数量及过滤掉  $His_{(i)}$  和  $SRH_{(i)}$  的项目数量.

由公式 (14) 可知, 在推荐列表中, 过滤掉目标用户的历史和相似项目后, 其剩余推荐列表项目的流行度越小  $Nov$  取值越高, 表明推荐就越新颖. 实验过程随机选取每个数据集中的 20 个用户的 top-20 推荐列表, 记录列表项目的平均流行度.

从后文图 12 推荐新颖性评估实验结果可以看出, UI-HEHo 的  $Nov$  值整体是最高的, 其主要原因是通过节点间的共享特征, 使原本无交集的节点也会存在交互的可能, 在用户-项目高阶交互的基础上增加用户对和项目对交互, 丰富了网络中的交互信息, 使推荐列表更容易生成用户未曾听说或使用的项目, 从而有利于长尾项目被发掘.

## 7 结 论

本文充分利用弱关系所具有的丰富语义和信息传递能力, 提出了一种面向推荐系统的异质信息网络高阶嵌入学习框架. 该学习框架利用了目标用户社交网络的丰富信息, 即: 异质信息网络中节点的交互结构和属性特征可以作为辅助信息实现其推荐. 尽管该学习框架具有准确的评级预测能力, 以及能生成有效、多样和新颖的推荐列表

等优点, 但如何考虑结合用户的实时位置信息实现超前推荐、探索用户间不信任关系对推荐产生的作用、同时优化评级预测和生成推荐将是下一步研究的重点.

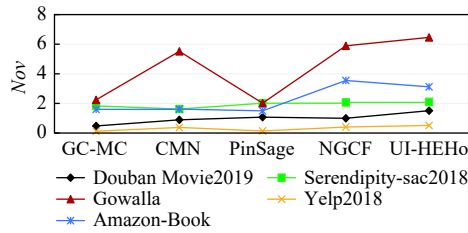


图 12 推荐的新颖性评估

## References:

- [1] Liu HF, Jing LP, Yu J. Survey of matrix factorization based recommendation methods by integrating social information. *Ruan Jian Xue Bao/Journal of Software*, 2018, 29(2): 340–362 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5391.htm> [doi: 10.13328/j.cnki.jos.005391]
- [2] Shi YC, Yin Y, Zhao YH, Zhang B, Wang GR. User recommendation algorithm based on multi-developer community. *Ruan Jian Xue Bao/Journal of Software*, 2019, 30(5): 1561–1574 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5733.htm> [doi: 10.13328/j.cnki.jos.005733]
- [3] Yang C, Fan Q, Wang T, Yin G, Wang HM. Multi-feature based personal recommendation approach for open source project. *Ruan Jian Xue Bao/Journal of Software*, 2017, 28(6): 1357–1372 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5230.htm> [doi: 10.13328/j.cnki.jos.005230]
- [4] Shi C, Hu BB, Zhao WX, Yu PS. Heterogeneous information network embedding for recommendation. *IEEE Trans. on Knowledge and Data Engineering*, 2018, 31(2): 357–370. [doi: 10.1109/TKDE.2018.2833443]
- [5] Zheng J, Liu J, Shi C, Zhuang FZ, Li JZ, Wu B. Recommendation in heterogeneous information network via dual similarity regularization. *Int'l Journal of Data Science and Analytics*, 2017, 3(1): 35–48. [doi: 10.1007/s41060-016-0031-0]
- [6] Zhao H, Yao QM, Li JD, Song YQ, Lee DL. Meta-graph based recommendation fusion over heterogeneous information networks. In: *Proc. of the 23rd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. Halifax: ACM, 2017. 635–644. [doi: 10.1145/3097983.3098063]
- [7] Yang C, Liu ZY, Zhao DL, Sun MS, Chang EY. Network representation learning with rich text information. In: *Proc. of the 24th Int'l Conf. on Artificial Intelligence*. Buenos: AAAI Press, 2015. 2111–2117.
- [8] Chen T, Zhu Q, Zhou MX, Wang S. Trust-based recommendation algorithm in social network. *Ruan Jian Xue Bao/Journal of Software*, 2017, 28(3): 721–731 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5159.htm> [doi: 10.13328/j.cnki.jos.005159]
- [9] Khan MM, Ibrahim R, Younas M, Ghani I, Jeong SR. Facebook interactions utilization for addressing recommender systems cold start problem across system domain. *Journal of Internet Technology*, 2018, 19(3): 861–870. [doi: 10.3966/160792642018051903021]
- [10] Qin D, Zhou XM, Chen L, Huang GY, Zhang YC. Dynamic connection-based social group recommendation. *IEEE Trans. on Knowledge and Data Engineering*, 2020, 32(3): 453–467. [doi: 10.1109/TKDE.2018.2879658]
- [11] Lin X, Zhang M, Liu YQ, Ma SP. Enhancing personalized recommendation by implicit preference communities modeling. *ACM Trans. on Information Systems*, 2019, 37(4): 1–32. [doi: 10.1145/3352592]
- [12] Li H, Zhang S, Hu Y, Shi J, Zhong ZM. Research of social recommendation based on social tag and trust relation. *Cluster Computing*, 2018, 21(1): 933–943. [doi: 10.1007/s10586-017-0962-9]
- [13] Cui LZ, Ou P, Lu N, Zhang GJ. A comprehensive trust-based item evaluation model for recommendation in social network. In: *Proc. of the 2016 IEEE Symp. on Computers and Communication (ISCC)*. Messina: IEEE, 2016. 1090–1096. [doi: 10.1109/ISCC.2016.7543881]
- [14] Forsati R, Mahdavi M, Shamsfard M, Sarwat M. Matrix factorization with explicit trust and distrust side information for improved social recommendation. *ACM Trans. on Information Systems*, 2014, 32(4): 17. [doi: 10.1145/2641564]
- [15] Zhang Y, Sun GM, He P, Zhai P, Sun YG. Dynamic evaluation of recommendation trust in open networks. *Journal of Web Engineering*, 2020, 19(7–8): 218–229. [doi: 10.13052/jwe1540-9589.197811]
- [16] Hamdi S, Gancarski AL, Bouzeghoub A, Yahia SB. Tison: Trust inference in trust-oriented social networks. *ACM Trans. on Information*



- Systems, 2016, 34(3): 1–32. [doi: [10.1145/2858791](https://doi.org/10.1145/2858791)]
- [17] Hu LM, Yang TC, Shi C, Ji HY, Li XL. HGAT: Heterogeneous graph attention networks for semi-supervised short text classification. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing (EMNLP-IJCNLP). Hong Kong: ACL, 2019. 4821–4830. [doi: [10.18653/v1/D19-1488](https://doi.org/10.18653/v1/D19-1488)]
- [18] Salakhutdinov RR, Mnih A. Probabilistic matrix factorization. In: Proc. of the 20th Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2008. 1257–1264.
- [19] Rendle S. Factorization machines with libFM. ACM Trans. on Intelligent Systems and Technology, 2012, 3(3): 57. [doi: [10.1145/2168752.2168771](https://doi.org/10.1145/2168752.2168771)]
- [20] Yuan F, Lin WQ, Zheng VW, Wu M, Chang KCC, Li XL. Semantic proximity search on graphs with metagraph-based learning. In: Proc. of the 32nd IEEE Int'l Conf. on Data Engineering (ICDE). Helsinki: IEEE, 2016. 277–288. [doi: [10.1109/ICDE.2016.7498247](https://doi.org/10.1109/ICDE.2016.7498247)]
- [21] Huang ZP, Zheng YD, Cheng R, Sun YZ, Mamouli N, Li X. Meta structure: Computing relevance in large heterogeneous information networks. In: Proc. of the 22nd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016. 1595–1604. [doi: [10.1145/2939672.2939815](https://doi.org/10.1145/2939672.2939815)]
- [22] Sheikh N, Kefato Z, Montresor A. gat2vec: Representation learning for attributed graphs. Computing, 2019, 101(3): 187–209. [doi: [10.1007/s00607-018-0622-9](https://doi.org/10.1007/s00607-018-0622-9)]
- [23] Cao SS, Lu W, Xu QK. GraRep: Learning graph representations with global structural information. In: Proc. of the 24th ACM Int'l on Conf. on Information and Knowledge Management. Melbourne: ACM, 2015. 891–900. [doi: [10.1145/2806416.2806512](https://doi.org/10.1145/2806416.2806512)]
- [24] Rendle S, Freudenthaler C, Gantner Z, Schmidt-Thieme L. BPR: Bayesian personalized ranking from implicit feedback. In: Proc. of the 25th Conf. on Uncertainty in Artificial Intelligence. Montreal: AUAI Press, 2009. 452–461.
- [25] He XN, Liao LZ, Zhang HW, Nie LQ, Hu X, Chua TS. Neural collaborative filtering. In: Proc. of the 26th Int'l Conf. on World Wide Web. Perth: Int'l World Wide Web Conf. Steering Committee, 2017. 173–182. [doi: [10.1145/3038912.3052569](https://doi.org/10.1145/3038912.3052569)]
- [26] Perozzi B, Al-Rfou R, Skiena S. DeepWalk: Online learning of social representations. In: Proc. of the 20th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. New York: ACM, 2014. 701–710. [doi: [10.1145/2623330.2623732](https://doi.org/10.1145/2623330.2623732)]
- [27] Grover A, Leskovec J. node2vec: Scalable feature learning for networks. In: Proc. of the 22nd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016. 855–864. [doi: [10.1145/2939672.2939754](https://doi.org/10.1145/2939672.2939754)]
- [28] Tu CC, Zhang WC, Liu ZY, Sun MS. Max-margin DeepWalk: Discriminative learning of network representation. In: Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence. New York: ACM, 2016. 3889–3895.
- [29] Ivanov S, Burnaev E. Anonymous walk embeddings. In: Proc. of the 35th Int'l Conf. on Machine Learning. Stockholm: ICML, 2018. 2186–2195.
- [30] Dong YX, Chawla NV, Swami A. metapath2vec: Scalable representation learning for heterogeneous networks. In: Proc. of the 23rd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Halifax: ACM, 2017. 135–144. [doi: [10.1145/3097983.3098036](https://doi.org/10.1145/3097983.3098036)]
- [31] Shi C, Zhang ZQ, Luo P, Yu PS, Yue YD, Wu B. Semantic path based personalized recommendation on weighted heterogeneous information networks. In: Proc. of the 24th ACM Int'l on Conf. on Information and Knowledge Management. Melbourne: ACM, 2015. 453–462. [doi: [10.1145/2806416.2806528](https://doi.org/10.1145/2806416.2806528)]
- [32] Cao SS, Lu W, Xu QK. Deep neural networks for learning graph representations. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Phoenix: AAAI Press, 2016. 1145–1152.
- [33] Van Den Berg R, Kipf TN, Welling M. Graph convolutional matrix completion. arXiv:1706.02263, 2017.
- [34] Gao M, Chen LH, He XN, Zhou AY. BiNE: Bipartite network embedding. In: Proc. of the 41st Int'l ACM SIGIR Conf. on Research & Development in Information Retrieval. Ann Arbor: ACM, 2018. 715–724. [doi: [10.1145/3209978.3209987](https://doi.org/10.1145/3209978.3209987)]
- [35] Tang J, Lou TC, Kleinberg J, Wu S. Transfer learning to infer social ties across heterogeneous networks. ACM Trans. on Information Systems, 2016, 34(2): 7. [doi: [10.1145/2746230](https://doi.org/10.1145/2746230)]
- [36] Ebesu T, Shen B, Fang Y. Collaborative memory network for recommendation systems. In: Proc. of the 41st Int'l ACM SIGIR Conf. on Research & Development in Information Retrieval. Ann Arbor: ACM, 2018. 515–524. [doi: [10.1145/3209978.3209991](https://doi.org/10.1145/3209978.3209991)]
- [37] Tu CC, Liu H, Liu ZY, Sun MS. CANE: Context-aware network embedding for relation modeling. In: Proc. of the 55th Annual Meeting of the Association for Computational Linguistics. Vancouver: ACL, 2017. 1722–1731. [doi: [10.18653/v1/P17-1158](https://doi.org/10.18653/v1/P17-1158)]
- [38] Ying R, He RN, Chen KF, Eksombatchai P, Hamilton WL. Graph convolutional neural networks for Web-scale recommender systems. In: Proc. of the 24th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. London: ACM, 2018. 974–983. [doi: [10.1145/3219819.3219890](https://doi.org/10.1145/3219819.3219890)]
- [39] Shi Y, Gui H, Zhu Q, Kaplan L, Han JW. AspEm: Embedding learning by aspects in heterogeneous information networks. In: Proc. of the 2018 SIAM Int'l Conf. on Data Mining (SDM). San Diego: SIAM, 2018. 144–152. [doi: [10.1137/1.9781611975321.16](https://doi.org/10.1137/1.9781611975321.16)]



- [40] Wang X, He XN, Wang M, Feng FL, Chua TS. Neural graph collaborative filtering. In: Proc. of the 42nd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Paris: ACM, 2019. 165–174. [doi: 10.1145/3331184.3331267]
- [41] Chen ZX, Wang XT, Xie X, Wu T, Bu GQ, Wang YN, Chen EH. Co-attentive multi-task learning for explainable recommendation. In: Proc. of the 28th Int'l Joint Conf. on Artificial Intelligence. Macao: AAAI Press, 2019. 2137–2143.
- [42] Ma JQ, Zhao Z, Yi XY, Chen JL, Hong LC, Chi EH. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In: Proc. of the 24th ACM SIGKDD Int'l Conf. on Knowledge Discovery & Data Mining. London: ACM, 2018. 1930–1939. [doi: 10.1145/3219819.3220007]
- [43] Vaswani A, Shazeer N, Parmar N, Parmar J, Jones L, Gomez AN, Kaiser Ł, Polosukhin I. Attention is all you need. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- [44] Kotkov D, Konstan JA, Zhao Q, Veijalainen J. Investigating serendipity in recommender systems based on real user feedback. In: Proc. of the 33rd Annual ACM Symp. on Applied Computing. Pau: ACM, 2018. 1341–1350. [doi: 10.1145/3167132.3167276]

#### 附中文参考文献:

- [1] 刘华锋, 景丽萍, 于剑. 融合社交信息的矩阵分解推荐方法研究综述. 软件学报, 2018, 29(2): 340–362. <http://www.jos.org.cn/1000-9825/5391.htm> [doi: 10.13328/j.cnki.jos.005391]
- [2] 时宇岑, 印莹, 赵宇海, 张斌, 王国仁. 基于多开发者社区的用户推荐算法. 软件学报, 2019, 30(5): 1561–1574. <http://www.jos.org.cn/1000-9825/5733.htm> [doi: 10.13328/j.cnki.jos.005733]
- [3] 杨程, 范强, 王涛, 尹刚, 王怀民. 基于多维特征的开源项目个性化推荐方法. 软件学报, 2017, 28(6): 1357–1372. <http://www.jos.org.cn/1000-9825/5230.htm> [doi: 10.13328/j.cnki.jos.005230]
- [8] 陈婷, 朱青, 周梦溪, 王珊. 社交网络环境下基于信任的推荐算法. 软件学报, 2017, 28(3): 721–731. <http://www.jos.org.cn/1000-9825/5159.htm> [doi: 10.13328/j.cnki.jos.005159]



荀亚玲(1980—), 女, 博士, 教授, CCF 专业会员, 主要研究领域为大数据分析及应用, 并行与分布式计算.



张继福(1963—), 男, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为大数据挖掘及应用, 人工智能及应用.



毕慧敏(1996—), 女, 硕士生, 主要研究领域为大数据分析及应用, 并行与分布式计算.