计算机研究与发展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202221045
61(7): 1825−1835, 2024

完全图高阶关系驱动的链接预测

张惠鹃[1] 黄钦阳[3] 胡诗彦[2] 杨 青[2] 张敬伟[1]
[1](广西可信软件重点实验室(桂林电子科技大学) 广西桂林 541004)
[2](广西自动检测技术与仪器重点实验室(桂林电子科技大学) 广西桂林 541004)
[3](桂林电子科技大学计算机与信息安全学院 广西桂林 541004)
(965727190@qq.com)

Link Prediction Driven by High-Order Relations in Complete Graph

Zhang Huijuan[1], Huang Qinyang[3], Hu Shiyan[2], Yang Qing[2], and Zhang Jingwei[1]
[1] (*Guangxi Key Laboratory of Trusted Software* (*Guilin University of Electronic Technology*), *Guilin*, *Guangxi* 541004)
[2] (*Guangxi Key Laboratory of Automatic Detecting Technology and Instruments* (*Guilin University of Electronic Technology*), *Guilin*, *Guangxi* 541004)
[3] (*School of Computer Science and Information Security*, *Guilin University of Electronic Technology*, *Guilin*, *Guangxi* 541004)

Abstract Graph convolutional network (GCN) has been widely used in recommendation systems due to its unique advantages in processing graph data. GCN propagates node attribute information by exploiting the dependencies between nodes in the graph, which greatly improves the accuracy of node representation and thus improves recommendation performance. However, existing GCN-based recommendation methods still have difficulty in modeling deeper layers due to the over-smoothing problem, which limits the representation of higher-order relationships between users and items. To this end, we propose a novel link prediction method driven by high-order relations among items in complete graph (short for LinkCG). The LinkCG method utilizes a heterogeneous graph consisting of a user-item interaction graph and a global graph representing implicit item associations to predict user-item links. By directly modeling item associations based on user interactions, LinkCG captures higher-order relationships and mitigates the issue of data sparsity. In addition, unlike node embedding-based methods, LinkCG improves the training process and enhances recommendation accuracy by assigning link weights that represent the degree of item association. Experimental results on three publicly available datasets show that LinkCG method, as a non-deep learning model containing only two hyperparameters, provides better performance compared with the deep learning-based baseline methods. The application on social relationship data further demonstrates LinkCG can capture rich enough information about user preferences from their historical interaction items.

Key words recommendation system; link prediction; complete graph; high-order relations; associated relationship

摘 要 图卷积网络(graph convolutional network, GCN)因其在处理图数据方面的独特优势而被广泛应用于推荐系统中,它通过利用图中节点之间的依赖关系传播节点属性信息,极大地提高了节点表示的准确度从而提升推荐性能.然而现有基于GCN的推荐方法仍因过平滑问题而难以进行更深层的建模,从而限

收稿日期: 2022-12-30; **修回日期**: 2023-11-24
基金项目: 国家自然科学基金项目(62267002, 62167002, 61862013); 广西重点研发计划项目(桂科 AB22080047); 广西可信软件重点实验室主任基金重点课题(KX202317); 广西自动检测技术与仪器重点实验室主任基金课题(YQ21102)
This work was supported by the National Natural Science Foundation of China (62267002, 62167002, 61862013), Guangxi Key Research and Development Program (Guike AB22080047), Guangxi Key Laboratory of Trusted Software (KX202317), and Guangxi Key Laboratory of Automatic Detecting Technology and Instruments (YQ21102).
通信作者: 张敬伟(gtzjw@hotmail.com)

制了用户与项目间高阶关系的表达.为此,提出了一种基于项目间关系的完全图高阶关系驱动的链接预测（link prediction driven by high-order relations in complete graph, LinkCG）方法. LinkCG通过用户-项目交互图与项目间隐式关联关系全局图组成的异构图预测用户到项目的链接,跳过了中间的用户节点直接利用完全图建模每个用户历史交互的项目间的局部隐式关联关系,获得项目间的高阶关系从而缓解数据稀疏性问题;此外,不同于基于节点嵌入的推荐方法,LinkCG通过赋予项目间的链接权重来表示项目间关系的紧密程度,并根据紧密程度进行链接预测,优化了模型的训练过程.在3个公开数据集上的实验结果表明,LinkCG作为只包含2个超参数的非深度学习模型,与一些先进的基于深度学习的基线方法相比提供了更好的性能.在社交关系数据上的应用进一步表明LinkCG能够从用户历史交互项目中获取足够丰富的用户偏好信息.

关键词 推荐系统;链接预测;完全图;高阶关系;关联关系

中图法分类号 TP274；TP301

信息爆炸时代,为用户提供高效且个性化的服务成为在线平台保持用户粘性的关键.推荐系统作为缓解信息过载问题的有效工具,在提升用户体验方面取得了诸多成果[1-3].协同过滤是推荐系统中影响最广泛的技术之一,它利用群体的行为获得隐式反馈或显式反馈来做决策,为用户推荐最可能偏好的项目.早期使用矩阵分解的推荐模型[4]的核心假设是用户和项目可以通过隐向量表示,越相似的用户和项目的表示在嵌入空间中越相近.

近年来的研究发现,将用户历史交互数据建模成图结构数据时,可以显式地沿着用户-项目交互边传播信息,通过递归地聚合邻居节点信息细化用户和项目嵌入.文献[5−7]通过捕获基于图结构数据的协同信号进行推荐,大量研究工作表明了基于图神经网络的协同过滤模型能够提升推荐的准确性.这些方法依赖于邻节点聚合递归地传递节点信息,使得邻节点之间共享特征,高阶信息的获取需要堆叠多层神经网络,然而神经网络层数越深会使得中心节点和邻居节点越相似而导致过平滑问题.为此,本文提出一种直接对图中节点之间的链接进行建模的链接预测方法,提取了更丰富的高阶关系并有效地避免了过平滑问题.

早期的链接预测方法,如将用户-项目交互数据建模为图结构数据[8],证明了使用链接预测方法来提出协同过滤推荐建议比标准的协同过滤算法有更好的性能.文献[9]进一步证明,对于像Flickr这样庞大的用户生成内容系统库,使用链接预测算法对经典协同过滤算法进行扩展能获得更好的性能;还证明了只考虑用户-项目图中用户的近邻在性能上要优于使用整个图结构来进行推荐的算法.经典的链接预测方法中,CN(common neighbors)[8]从图的拓扑结构角度出发,通过考虑2个节点的共同邻居数定义节点之间的相似度,从而得到2个节点之间的链接分数.CN方法因其简单且有效的特点被广泛应用,如文献[10−11]就是在CN方法的基础上考虑节点度的影响提出的新的链接预测方法.这些工作[8−11]都指出了在项目推荐中使用链接预测方法处理图结构数据的巨大潜力.

上述链接预测方法虽然有效,但依然是致力于对图中的节点进行建模而不是直接建模节点间的链接,此外,直接在用户-项目交互图中进行链接预测难以获得高阶节点间的链接信息.因此,本文提出了一种基于项目间隐式关联关系的完全图高阶关系驱动的链接预测(link prediction driven by high-order relations in complete graph, LinkCG)方法,直接对图中节点之间的链接进行建模,并通过建模项目间关系获得节点间高阶关系信息.

LinkCG首先对用户历史交互项目建模,构建一个与用户对应的局部有向完全图来表示项目间隐式关联关系,通过2个超参数控制有向边的权重,所有用户对应的局部图叠加得到项目间隐式关联关系的全局图F,最后通过用户-项目交互图G与项目间隐式关联关系的全局图F组成的异构图E预测从用户到项目的链接.以往的链接预测方法直接在用户-项目交互图G中进行用户到未交互项目的链接预测,这种链接预测方式属于低阶链接,如图1低阶链接方式中,用户u_1通过已交互项目v_1,经过3跳只能链接到1个未交互项目v_2,而LinkCG通过对项目间关系的建模,能够直接链接2个存在隐式关联关系的项目,用户到未交互项目的链接预测方式属于高阶链接,如图1高阶链接方式中,用户u_1通过已交互项目v_1,经过3跳能够链接到2个未交互项目v_2和v_3.由

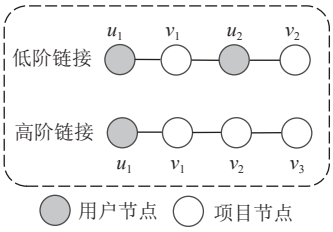


Fig. 1 Comparison of linking methods

图 1 链接方式对比

此可见，LinkCG 通过建模项目间关系的链接，用户节点可以更快地链接到未交互的高阶项目. 为缓解数据稀疏性问题，LinkCG 还将全局图 F 中项目间的 2 跳关联关系作为 1 跳关系的补充和增强，探索更多用户可能偏好的项目. 最后，在包含社交关系数据的实验结果表明，本文方法通过用户历史交互项目所获取的用户偏好信息已足够丰富，额外的社交关系无法带来显著的性能提升.

综上所述，本文的主要贡献有 3 个方面：

1）提出了利用项目间隐式关联关系解决链接预测问题的新视角，构建局部有向完全图表示每个用户历史交互的项目之间的隐式关联关系，并将局部完全图叠加获得项目间的隐式关联关系全局图.

2）提出了一种新的在异构图上进行用户到项目的链接预测方法 LinkCG，并在此异构图上提出利用项目间的 2 跳关系缓解数据稀疏性问题，丰富用户偏好信息，提高推荐准确性.

3）在 3 个公开数据集上进行了充分的对比实验和消融实验，实验结果表明 LinkCG 相比于基线提供了更好的性能. 此外，通过添加额外的社交关系的实验表明，LinkCG 在用户历史交互项目中能够获得足够丰富的用户偏好信息，性能上超过一些先进的社交推荐方法.

1 相关工作

在本节中，我们主要介绍基于图的推荐方法以及推荐系统中用户到项目的链接预测方法.

1.1 基于图的推荐方法

图数据可以很好地表达推荐系统中的数据关系，如用户-项目型数据、用户-用户-项目型社交关系数据，这使得基于图的模型成为当今推荐系统的研究热点. 早期的工作[12-14]通过在图上进行随机游走来学习节点的局部潜在表征，将截断的随机游走访问到的节点作为可能的推荐结果，这些工作具有较大的随机性，无法取得最优的推荐性能. 为了改进这些方法，一些工作[15-18]考虑图嵌入的方式，首先为图中用户节点和项目节点都分配统一长度的嵌入表示，再将节点表示输入模型进行训练. 近年来，随着图神经网络在各领域的成功[19-21]，吸引了一大批学者对基于图神经网络的推荐模型的研究，如社交推荐[22]、序列推荐[23]以及会话推荐[24]等. 图神经网络主要是利用图中节点之间的依赖关系通过迭代聚合邻节点嵌入的方式传播嵌入，节点可以访问到高阶邻居的信息，学习到更准确的节点嵌入，如 NGCF（neural graph collaborative filtering）[5]通过在用户-项目交互 2 部图上执行嵌入传播获得高阶协同信号；LightGCN[6]将 NGCF 中的特征变换和非线性激活部分去掉，得到一个更简单但性能更好的模型；LR-GCCF[25]使用线性变换替换图卷积网络（graph convolutional network, GCN）的非线性变换，并设计一种基于用户-项目交互图建模的残差网络缓解图卷积邻域聚合操作中的过平滑问题. 与这些工作类似，LinkCG 也是利用图来表示一个用户的历史交互项目之间的隐式关联关系，不同的是，LinkCG 没有采用节点嵌入的方式，而是直接对图中的链接建模，能够更有效地避免过平滑问题.

1.2 链接预测驱动的推荐方法

本文的工作也与链接预测紧密相关，LinkCG 主要解决的是根据现有观察到的用户-项目交互图 G 中的链接预测未观察到的用户与项目之间的链接问题. 文献 [26] 首先将社会网络进化背后的计算问题定义为链接预测问题，并提出了通过分析网络中节点的“接近度”来进行链接预测的方法. Chen 等人[8]将用户-项目交互作为图来分析，采用网络建模中的链接预测方法进行协同过滤推荐，利用 2 种基于路径的度量同时兼顾全局网络结构和单个链接度量，缓解了数据稀疏性问题，表明了链接预测方法的性能优于标准协同过滤算法. 文献 [9] 证明流行的链接预测算法比经典的协同过滤更适合在大规模在线社交网络中应用. Zhang 等人[27]成功利用图神经网络进行链接预测，表明了对链接周围标记的子图进行学习可以获得高度表达和可概括的链接表示. 将用户-项目交互 2 部图中用户和项目的度与链接分数联系起来已经取得了不错的成果[10-11,28]，与这些方法类似，LinkCG 也将用户和项目的度与链接分数关联；不同的是，LinkCG 不是直接在用户-项目交互图 G 中进行链接预测，而是在构建的包含项目间隐式关联关系的异构图上进行链接预测，这一做法相比于以往的链接预测方法能够获得更高阶的项目信息.

2 LinkCG 方法

2.1 问题定义

传统的链接预测方法被定义为在用户-项目交互 2 部图中通过已观察到的链接来预测用户和项目之间未观察到的链接任务，如 CN[8]、SC[10] 和 LHN[11]. 然而，项目之间的隐式关联关系才是促成潜在的链接产生的根本原因. 例如，项目 v_1 和 v_2 被多个用户同时交互，可认为这 2 个项目间存在较强的隐式关联关系，与项目 v_1 有交互的用户很可能对项目 v_2 感兴趣. 于是本文从用户历史交互的项目出发，通过建模项目之间隐式关联关系来直接链接 2 个项目，利用项目间关系图和用户-项目交互图组成的异构图获得用户到未交互项目的链接. 用户集合和项目集合分别用 U 和 V 表示，n 是用户的数量，m 是项目的数量. 去掉用户和项目之间重复的交互，用集合 C 表示观察到的用户 $u \in U$ 和项目 $v \in V$ 之间的交互. 图 $G=(U, V, C)$ 表示用户-项目交互 2 部图，并将其表示成式(1)所示的矩阵 $\boldsymbol{M} \in \mathbb{R}^{n \times m}$，其中 $\boldsymbol{M}$ 的元素 M_{ij} 为：

$$M_{ij}=\begin{cases}1, & \text{if } (u_i, v_j) \in C,\\ 0, & \text{其他.}\end{cases} \tag{1}$$

本文的任务是通过构建项目间的链接表示项目间的隐式关联关系，再利用用户历史交互项目链接到用户未交互项目，最终获得表示用户对未交互项目偏好的链接评分矩阵 $\boldsymbol{S} \in \mathbb{R}^{n \times m}$，其中 $\boldsymbol{S}$ 的元素 S_{ij} 为：

$$S_{ij}=\text{用户}u_i\text{到项目}v_j\text{的链接评分预测.} \tag{2}$$

2.2 LinkCG 方法总体架构

本文提出的 LinkCG 可以总结为 2 个部分：

1)局部有向完全图的构建及叠加. 如图 2(a)所示，为了让用户更快地链接到未交互的更高阶的项目，LinkCG 首先利用每个用户的历史交互项目信息，构建与用户对应的局部有向完全图以表示项目间的隐式关联关系，直接地链接 2 个项目，将用户和项目对关联关系的影响表示为权重赋值到链接上. 再通过所有用户对应的局部有向完全图的叠加获得项目间隐式关联关系的全局图 F，得到最终项目间隐式关联关系的权重表示.

2)异构图的构建及链接预测. 如图 2(b)所示，构建好项目间关系全局图 F 之后，为了获得用户到未交互项目的链接路径，将用户-项目交互图 G 和项目间隐式关联关系全局图 F 组合为异构图，用户从历史交互项目出发链接到未交互项目，通过路径权重确定用户对项目的偏好程度. 例如，1 个可能的链接传播路径为：$u_1 \leftrightarrow v_1 \rightarrow v_2 \rightarrow v_7 \cdots$，其中，$u_1 \leftrightarrow v_1$ 是已观察到的用户-项目交互关系，$v_1 \rightarrow v_2 \rightarrow v_7 \cdots$ 是构建的项目间的隐式关联关系全局图 F 中的链接关系. 这样就得到了 1 条从用户 u_1 到它未交互过的项目的链接路径.

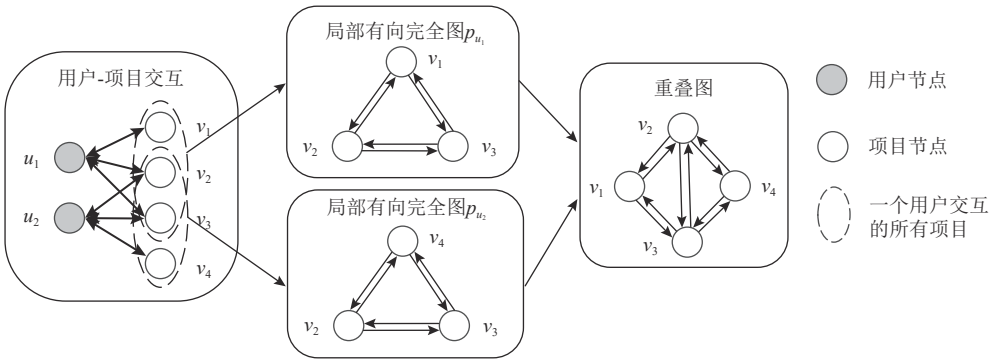
2.3 完全图的构建及叠加

若项目 v_i 和项目 v_j 被多个用户同时交互，可以认为项目 v_i 和项目 v_j 存在较强的隐式关联关系. 基于此想法，从用户历史交互的项目出发，构建局部有向完全图来表示项目间的局部隐式关联关系. 例如，在图 2(a)中，用户 u_1 所交互的项目为 v_1，v_2，v_3，构建一个只包含 v_1，v_2，v_3 的局部有向完全图 p_{u_1}.

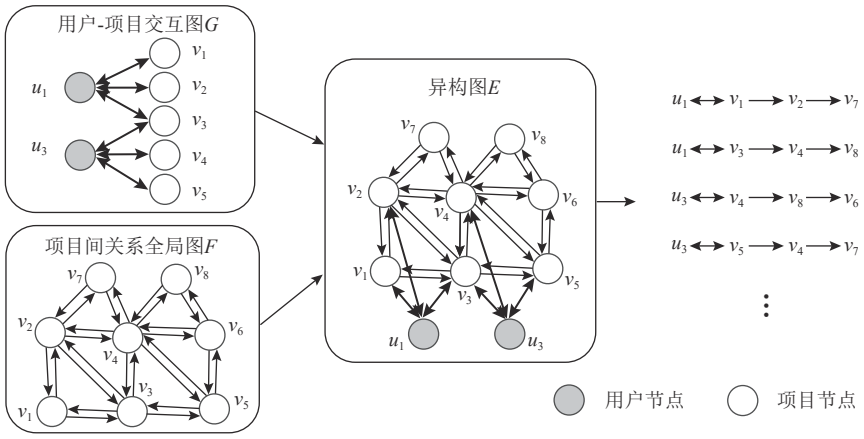
在用户角度，项目之间关联关系的紧密性与用户交互项目的数量有关，用户交互的项目数量越少，项目间的隐式关联关系越紧密，反之则越疏远；其次，2 个项目被越多用户同时交互，则 2 个项目间的隐式关联关系越紧密. 因此，可将交互图 G 中用户 u 的度与局部有向完全图 p_u 中的有向边权重联系起来，再将所有用户对应的局部有向完全图进行叠加. 在项目角度，1 个项目被越多用户交互，能传达出用户偏好的信息就越少，和文献 [29] 观点类似，在多个文档中都出现的术语能够提供的信息量较小. 因此，本文提出将交互图 G 中项目 v_1 的度与局部有向完全图 p_u 中的节点 v_1 的出边权重关联起来. 综上，用户 u_x 对应的局部有向完全图 p_{u_x} 中项目 v_i 到项目 v_j 的有向边权重函数定义为：

$$y\left(v_i \rightarrow v_j\right)=\frac{1}{d_{v_i}^{\alpha}} \times \frac{1}{d_{u_x}^{\beta}}, \tag{3}$$

其中 d_{u_x} 是用户 u_x 在用户-项目交互 2 部图 G 中的度，即用户交互的项目个数，d_{v_i} 是项目 v_i 在用户-项目交互 2 部图 G 中的度. 局部有向完全图的叠加方式如图 2(a)所示，用户 u_1 和用户 u_2 对应的局部有向完全图 p_{u_1} 和 p_{u_2} 叠加，相同节点间的有向边合并，权重定义为所合并的有向边的权重和. 所有局部有向完全图叠加得到表示项目间隐式关联关系的全局图 F，在全局图 F 中只考虑 2 个节点之间的 1 跳关系和 2 跳关系. 原因如下：首先，任意 2 个项目之间有直接相连的边表示这 2 个项目同时被 1 个或 1 个以上的用户交互，即 2 个项目之间具有直接的隐式关联关系. 在图 2(a)中，项目 v_2 和项目 v_3 同时被用户 u_1 和用户 u_2 交互，项目 v_2 和项目 v_4 同时被用户 u_2 交互，则若用户交互了项目 v_2，就有可能对项目 v_3 或项目 v_4 感兴趣. 其次，项目 v_1 和



(a) 局部有向完全图的构建及叠加



(b) 异构图的构建及链接预测

Fig. 2 The frame of LinkCG

图 2 LinkCG 框架

项目v_4之间虽然没有直接相连的边(直接的隐式关联关系)，但项目v_1却通过项目v_2和项目v_3与项目v_4间接相连，有$v_1 \to v_2 \to v_4$和$v_1 \to v_3 \to v_4$两条从项目v_1到项目v_4的 2 跳链接路径，即项目v_1和项目v_4之间有间接的隐式关联关系，若用户u与项目v_1交互，则用户u很可能对项目v_2和项目v_3感兴趣，继而因项目v_2和项目v_3的关系对项目v_4感兴趣.

在后续章节中的消融实验结果表明，2 跳关系的添加可以提升推荐的准确性. 在项目间的隐式关联关系全局图F中，项目节点间的有向边权重由 2 个项目节点所在的所有局部有向完全图中的权重求和得到，有向边权重函数计算为：

$$\hat{y}(v_i \to v_j) = \sum_{\langle v_i,\, v_j\rangle \in p_{u_x}} \left(\frac{1}{d_{v_i}^{\alpha}} \times \frac{1}{d_{u_x}^{\beta}} \right), \tag{4}$$

其中图p_{u_x}是用户u_x所交互的项目构成的局部有向完全图，$\langle v_i, v_j\rangle$是p_{u_x}中从项目v_i节点到项目v_j节点的有向边，d_{u_x}是用户u_x在用户-项目交互 2 部图G中的度即用户交互的项目个数，d_{v_i}是项目v_i在用户-项目交互 2 部图G中的度. 最终链接$v_i \to v_j$的权重由v_i和v_j被同时交互的所有用户的度以及项目v_i的度决定，即 LinkCG 仅包含 2 个控制节点度对有向边权重影响的超参数α和β，后面的消融实验结果表明这 2 个超参数的取值调整对提高本文方法推荐性能的必要性.

2.4 异构图的构建及链接预测

如图 2(b)所示，将项目间隐式关联关系全局图F和用户-项目交互 2 部图G组成异构图E，其中用户-项目交互 2 部图G中边的权重为 1，从而获得用户到未交互的项目的链接路径. 下一步，将式(4)转换成矩阵形式，计算用户到未交互项目的链接评分. 用d_{v_i}表示项目v_i的度，d_{u_j}表示用户u_j的度，则项目之间的 1 跳关联关系权重矩阵计算为：

$$\boldsymbol{O} = \left(\boldsymbol{A}\boldsymbol{M}^{\mathrm{T}}\right)(\boldsymbol{B}\boldsymbol{M}), \tag{5}$$

$$\boldsymbol{A} = \mathrm{diag}\left(d_{v_1}^{-\alpha}, d_{v_2}^{-\alpha}, \cdots, d_{v_m}^{-\alpha}\right), \boldsymbol{A} \in \mathbb{R}^{m\times m}, \tag{6}$$

$$\boldsymbol{B} = \mathrm{diag}\left(d_{u_1}^{-\beta}, d_{u_2}^{-\beta}, \cdots, d_{u_n}^{-\beta}\right), \boldsymbol{B} \in \mathbb{R}^{n\times n}, \tag{7}$$

其中矩阵$\boldsymbol{M}$是用户-项目交互矩阵，O_{ij}即为隐式关联关系全局图F中$v_i \to v_j$的边的权重. 上文提到，在隐式关联关系全局图F中，将项目之间的 1 跳关系和 2

跳关系相加,得到最终的项目之间有向边的权重矩阵为

$$\boldsymbol{T}=\boldsymbol{O}+\boldsymbol{O}^2, \tag{8}$$

则用户到项目的链接评分矩阵为

$$\boldsymbol{S}=\boldsymbol{MT}, \tag{9}$$

其中S_{ij}表示从用户u_i到项目v_j的链接打分,也表示用户u_i对项目v_j的偏好程度.

3 实验评估

在3个真实世界的数据集上验证本文提出的LinkCG的有效性. 在本节中,首先说明确保公平性和可重复性的实验设置以及评估方案;其次,展示LinkCG与基线在2个数据集上的实验结果对比;最后,通过实验验证LinkCG在社交关系数据方面的适用性、2个超参数对LinkCG的影响以及LinkCG在Gowalla数据集中的有效性.

3.1 数据集

本文使用2个从真实应用中收集的不同规模的数据集进行实验,数据集的统计数据如表1所示.

Table 1 Statistics of the Datasets

表1 数据集的统计

数据集	用户数	项目数	用户-项目交互数	用户-用户-项目分享数
Beidian	3 773	4 544	39 252	9 358
Beibei	149 361	30 486	1 089 266	522 264

1)Beidian. 由Lin等人[30]分享的数据集. Beidian是一个电子商务平台,这个数据集包含表示用户购买行为的用户-项目型的数据和表示用户分享行为的用户-用户-项目型的数据.

2)Beibei. 由Zhang等人[31]分享的数据集. Beibei是一个与拼多多类似的母婴产品电子商务平台,这个平台里团购是最受用户欢迎、最流行的购买产品方式. 数据集包含用户购买行为的用户-项目型数据以及用户-用户-项目型的团购数据.

来自2个电子商务平台的数据集反映了2个极具代表性的现实世界场景. 值得注意的是,采集的数据保留了交互较少的用户并对其进行推荐,使得本文方法能够有更加广泛且有价值的应用.

3.2 评估方法

本文将数据集分为训练集、验证集和测试集. 为了评估模型的整体性能,我们采用了2个常用的指标,即召回率(recall rate, Recall)和归一化折损累计增益(normalized discounted cumulative gain, NDCG),在推荐系统中定义为:

1)Recall@K. 在推荐系统中, Recall@K表示根据函数打分结果,为用户所推荐的前k名推荐列表中包含正样例的比例.

2)NDCG@K. NDCG@K相比Recall@K考虑得更全面,还考虑到了正样例的排名位置. 特别地,当只为用户推荐排名中的前1名项目时Recall@K和NDCG@K是等价的.

为了证明LinkCG的有效性且充分地利用数据集信息,除了选择2个基于图神经网络的推荐模型外,还选择了一些先进的社交推荐模型作为性能比较的基线.

1)DiffNet[32]. 相比于以往的只考虑静态社交关系的社交推荐模型,它是一种用户间关系迭代影响的模型.

2)LightGCN[6]. 提出简化并增强基于协同过滤的图卷积网络技术,去掉图卷积中的特征转换和非线性激活2个部分,仅保留最重要的组件-邻域聚合,具有更容易被训练、更好的泛化能力及更有效等优点.

3)LR-GCCF[25]. 提出了一种专门为获得用户-项目交互的协同信号而设计的残差网络结构,可以缓解数据稀疏性问题以及邻域聚合操作中的过平滑问题.

4)SocialMF[33]. 是一个影响广泛的基于矩阵分解的社交推荐模型. 该模型中每个用户的特征向量依赖于其在社交网络中的直接邻居的特征向量,能够处理信任和信任传播的传递性.

5)MHCN[22]. 是一种基于多通道超图卷积网络的社交推荐模型. 该方法利用超图建模复杂的社交关系,并通过聚合多个通道学习到的嵌入优化用户表示.

6)Consisrec[34]. 是一个影响广泛的基于图神经网络的社交推荐模型. 该方法有效解决了社交推荐中用户之间上下文不一致以及关系不一致问题,能够为用户找到一致邻居,消除推荐噪音.

7)ESRF[35]. ESRF利用自动编码器有选择地使用社交关系,解决社交关系中常见的噪音问题.

8)Mult-VAE[36]. 将变分自编码扩展到隐式反馈的协同滤波,成功应用于推荐系统并取得更好的性能.

9)LinkProp[28]. LinkProp是通过非深度学习模型进行链接预测的最先进模型之一,仅利用3个超参数控制链接权重进行链接预测任务.

3.3 主要实验结果

和现有工作[2, 32]一样,本文采用抽样指标来评估所有方法的性能,为测试集中的每一个正样例项目随机采样100个负样例项目. 对于top-K指标,我们参

照文献 [37] 中广泛使用的范围，设置 $K \in \{1, 3, 5, 10\}$. 所有基线模型在 Beidian 数据集上的嵌入维度设置为 16，学习率为 0.01. 在 Beibei 数据集的实验设置参照文献 [38]，在 Beibei 数据集上完成了 3 个基线实验 LightGCN[6]，LR-GCCF[25]，Consisrec[34]，并从文献 [38] 中选取了 SocialMF[33] 和 MHCN[22] 这 2 个基线实验结果作为对比. 本文 LinkCG 模型和基线在 2 个数据集上的性能比较结果分别展示在表 2 和表 3 中，尽管 LinkCG 是只包含 2 个超参数的简单模型，在 2 个数据集上的 2 个性能评价指标都优于基线，只在 Beibei 数据集上的 Recall@10 略低于 Consisrec[34] 的性能. 这是由于现有的基于图的推荐模型主要依靠邻节点聚合的方式优化中心节点的嵌入表示，随着神经网络层数的增加容易产生过平滑现象，难以进行较深的学习. 而 LinkCG 直接对节点之间的链接进行建模，通过对用户所交互的项目构建局部有向完全图来表示项目间的隐式关联关系，搭建一条从用户到未交互的项目之间的链接，将相同边的权重叠加得到节点间关联关系的强度，不需要邻节点聚合，避免了过平滑现象. 此外，LinkCG 中将高阶关系作为对低阶关系的增强或对低阶关系不足的补充，有效地缓解了数据稀疏性问题.

Table 2 Performance Comparison of LinkCG and Benchmark Methods in Beidian Dataset

表 2 Beidian 数据集中 LinkCG 与基线方法的性能对比

方法	Recall@1	Recall@3	Recall@5	Recall@10	NDCG@3	NDCG@5	NDCG@10
LightGCN	0.165 5	0.315 4	0.389 7	0.502 2	0.252 0	0.282 6	0.318 1
LR-GCCF	0.169 9	0.299 5	0.370 2	0.477 3	0.245 5	0.274 5	0.309 2
DiffNet	0.150 2	0.287 7	0.372 1	0.503 8	0.233 1	0.271 2	0.322 9
MHCN	0.154 3	0.286 8	0.374 0	0.504 0	0.232 2	0.270 7	0.321 1
Consisrec	0.152 5	0.282 9	0.364 3	0.490 5	0.227 9	0.261 4	0.301 6
ESRF	0.140 9	0.271 3	0.348 6	0.480 0	0.218 5	0.254 4	0.304 6
LinkCG（本文）	**0.183 9**	**0.326 2**	**0.395 4**	**0.509 8**	**0.266 4**	**0.294 7**	**0.331 7**

注：黑体数值表示最优值.

Table 3 Performance Comparison of LinkCG and Benchmark Methods in Beibei Dataset

表 3 Beibei 数据集中 LinkCG 与基准方法的性能对比

方法	Recall@1	Recall@3	Recall@5	Recall@10	NDCG@3	NDCG@5	NDCG@10
LightGCN	0.137 4	0.267 7	0.345 2	0.465 6	0.212 5	0.244 3	0.283 1
LR-GCCF	0.135 3	0.265 3	0.343 1	0.466 8	0.210 1	0.242 0	0.281 9
SocialMF	0.123 5	0.250 9	0.329 7	0.453 1	0.196 8	0.229 2	0.269 0
MHCN	0.109 3	0.227 4	0.302 0	0.422 9	0.177 2	0.207 8	0.246 8
Consisrec	0.126 5	0.263 4	0.346 4	**0.476 6**	0.205 3	0.239 4	0.281 4
LinkCG（本文）	**0.145 4**	**0.275 9**	**0.352 0**	0.475 5	**0.220 6**	**0.251 9**	**0.291 7**

注：黑体数值表示最优值.

3.4 消融实验

3.4.1 超参数消融分析

LinkCG 是仅包含 2 个超参数的非深度学习模型，因此通过损失函数来进行梯度下降的常规优化方法不适用于 LinkCG. 考虑到 LinkCG 包含的超参数较少，本文首先对超参数空间进行量化并通过粗略的网格搜索次优超参数对，再在次优超参数对附近加减 0.2 的范围以 0.05 为间隔搜索最优超参数对. 表 4 展示了在拟合模型时超参数的粗略搜索值以及在 2 个数据集上搜索到的最优超参数值，显然，搜索的超参数对总数来说是相当少的. 本文首先在数据验证集上搜索最优的模型超参数对，再使用这些设置在测试集中进行链接预测，最终得到一个用户-项目打分矩阵 $\boldsymbol{S}$，根据打分对用户未交互项目进行排序，观察正样例项目的排名情况. 根据搜索结果，在 Beidian 数据集的验证集上搜索到的最优超参数对值为：$\{\alpha=0.6, \beta=1.05\}$，在 Beibei 数据集的验证集上搜索到的最优超参数对值为 $\{\alpha=0.6, \beta=0.3\}$. 表 5 显示了 LinkCG 排除超参数时的性能，将所排除的参数值设置为 0，其中 LinkCG-α 表示将 β 设置为 0 并搜索 α 的最优超参数值，

Table 4 Parameters of LinkCG

表 4 LinkCG 参数

数据集	参数搜索值	最优参数对
Beidian	0.2, 0.4, 0.6, 0.8,1.0,1.2,1.4,1.6	{α=0.6, β=1.05}
Beibei		{α=0.6, β=0.3}

LinkCG-β表示将α设置为 0 并搜索β的最优超参数值. 表 5 中，排除所有超参数的模型 LinkCG*性能表现最差，使用 2 个超参数可以提供最强的性能结果. 此外，本文通过记录 LinkCG 在 Beidian 数据集上不同超参数对情况下的性能，观察 LinkCG 对超参数的敏感性，由于 NDCG 相对于 Recall 还考虑了正样例的排名情况，所以本文选取 NDCG@10 作为观察性能的指标. 由图 3 可以看出，LinkCG 在超参数变化过程中性能变化缓和并且存在一定规律，观察到次优超参数对为{α=0.6, β=1.0}，下一步在次优超参数对附近加减 0.2 的范围进行离散化采样，搜索到最优超参数对.

Table 5 Ablation Experiment of Two Hyper-parameters in LinkCG

表 5 LinkCG 上的 2 个超参数的消融实验

数据集	方法	Recall@5	Recall@10	NDCG@5	NDCG@10
Beidian	LinkCG*	0.313 9	0.410 4	0.229 5	0.260 7
	LinkCG-α	0.318 0	0.417 5	0.232 3	0.264 5
	LinkCG-β	0.388 8	0.491 2	0.286 8	0.320 0
	LinkCG（本文）	**0.395 4**	**0.509 8**	**0.294 7**	**0.331 7**
Beibei	LinkCG*	0.346 8	0.465 0	0.246 9	0.285 0
	LinkCG-α	**0.353 2**	0.473 5	0.251 4	0.290 2
	LinkCG-β	0.344 9	0.467 7	0.246 6	0.286 1
	LinkCG（本文）	0.352 0	**0.475 5**	**0.251 9**	**0.291 7**

注：黑体数值表示最优值.

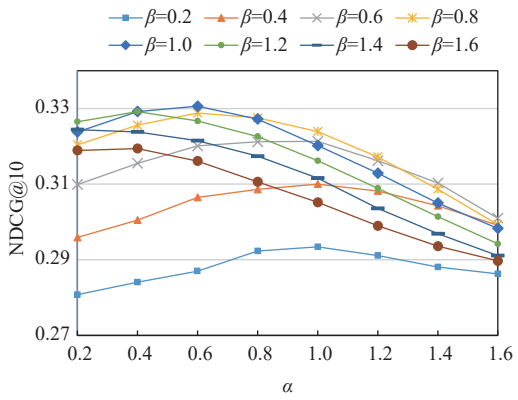


Fig. 3 Effect of parameter changes on performance in Beidian

图 3 Beidian 数据集中参数变化对性能的影响

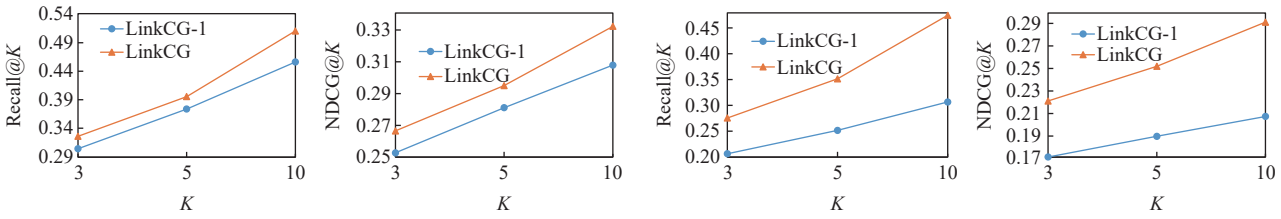
3.4.2 稀疏性问题

值得注意的是，本文使用的 2 个数据集没有根据用户历史交互项目数量的多少对用户进行筛选，依然保留了与项目交互较少的用户，而与项目交互较少的用户通过所交互的项目链接到的 1 跳未交互项目较少，不足以表达用户的偏好. 因此，本文通过获取项目间关系全局图F中的 2 跳关系信息，如式(8)所示，作为项目间 1 跳关系的补充或增强，为用户找到更高阶的未交互项目信息，提升推荐准确性. 为验证项目间的 2 跳关系可以丰富用户的偏好信息，本文通过实验对比了仅包含项目间 1 跳关系的 LinkCG 模型(LinkCG-1)与包含项目间 1 跳关系以及 2 跳关系的模型 LinkCG 的推荐性能，如图 4 所示. 图 4 中，在 Beibei 数据集上加入项目间 2 跳关系信息后，Recall@10 从 0.307 0 提升到 0.475 5，NDCG@10 从 0.206 7 提升到 0.291 7，说明加入项目间 2 跳信息能够丰富用户偏好信息，有效提升推荐性能.

为进一步验证 LinkCG 中加入项目间 2 跳关系有助于缓解数据稀疏性问题，本文根据所交互的项目数量将用户分成 2 组，将交互项目数量低于 10 的用户划分为交互稀疏用户，交互项目数量大于等于 10 的用户划分为交互不稀疏用户，并在不同的组中评估 LinkCG-1 与 LinkCG 的性能，如图 5 所示. 从图 5 可以看出，LinkCG 相比于 LinkCG-1 在用户交互数小于 10 的数据集中的性能提升要远大于在用户交互数大于等于 10 的数据集中的性能提升，这证明了 LinkCG 在处理数据稀疏性问题上的有效性.

3.4.3 适用性分析

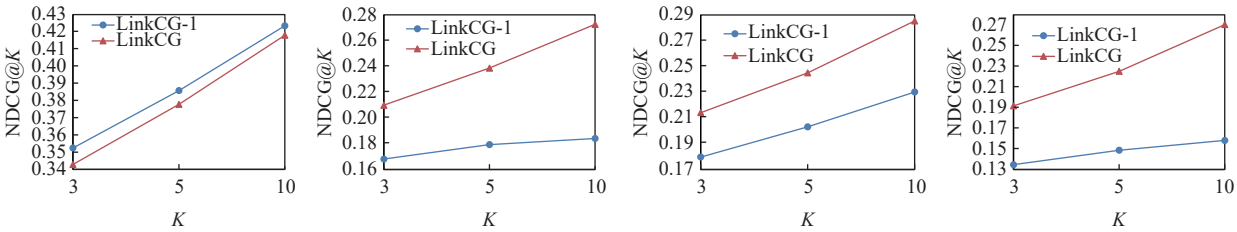
本文通过将社交关系数据构建为异构图的形式以挖掘隐含的用户偏好信息并添加到现有模型中，分析本文方法的适用性. 具体的，对于用户-用户-项目型的社交关系数据，构建一个用户-项目分享图，对分享了同一个项目的用户构建局部有向完全图表示用户之间局部的隐式关联关系，用户节点间的有向边权重由用户与项目节点在用户-项目分享图中的度决定，由于在 2 个数据集中，Beidian 数据集的用户-用户-项目型数据表示 2 个有社交关系的用户分享项目的行为，而 Beibei 数据集的用户-用户-项目型数据表示 2 个用户拼团购买项目的行为，用户之间不一定有社交联系，而且拼团购买行为与用户-项目交互行为重复，所以本文只在 Beidian 数据集上对用户-用户-项目型数据进行处理，将利用用户社交关系得到的用户到项目的链接预测打分添加到现有用户-项目打分矩阵$\boldsymbol{S}$中，并将添加了社交数据信息的 LinkCG 模型(LinkCG+S)与基线中的社交推荐方法进行比较，实验结果如图 6 所示. 图 6 的结果表明，LinkCG 在用户历史交互项目中能够获得足够丰富的用户偏好信息，社交关系信息的添加只能稍微提升推荐性能，用



(a) Beidian数据集 (b) Beibei数据集

Fig. 4 Performance comparison of LinkCG with and without two-hop relationship between items

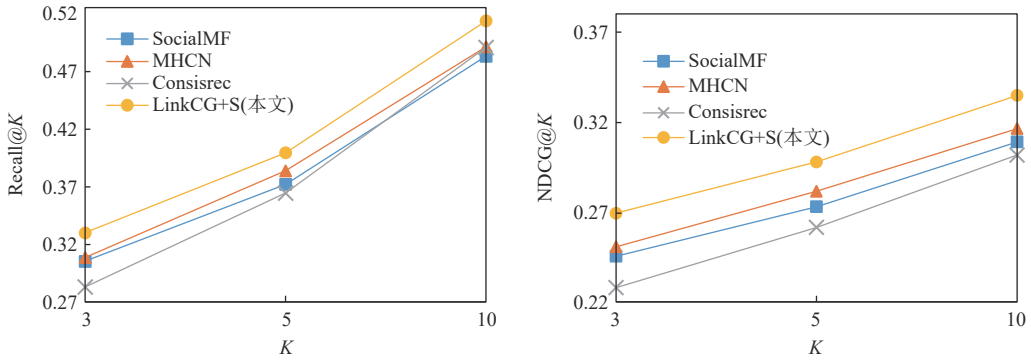
图 4 LinkCG 有无项目间 2 跳关系的性能对比



(a) Beidian (交互数大于等于10) (b) Beidian (交互数小于10) (c) Beibei (交互数大于等于10) (d) Beibei (交互数小于10)

Fig. 5 Performance comparison of LinkCG with and without two-hop relationship between items in different sparse datasets

图 5 LinkCG 有无项目间 2 跳关系在不同稀疏程度数据集中的性能对比



(a) Recall性能指标 (b) NDCG性能指标

Fig. 6 Performance comparison of LinkCG's social relationship variant and social recommendation models

图 6 LinkCG 的社交关系变体与社交推荐模型的性能对比

户的偏好主要由历史交互项目体现，同时证明了 LinkCG 的适用性. 与 LinkCG 类似设置 2 个超参数控制节点度对有向边权重的影响. 将每个项目对应的局部有向完全图叠加得到用户之间全局隐式关联关系图. 再结合用户-项目交互图 G，从而得到一条用户-用户-项目的链接路径，为用户找到可能偏好的项目，其中用户-用户是分享了相同项目而具备的用户间隐式关联关系，用户-项目是已观察到的用户-项目交互行为.

3.4.4 有效性分析

为进一步验证 LinkCG 方法的有效性，本文还在广泛使用的基准数据集 Gowalla 上进行了实验. Gowalla 数据集包含 29 858 个用户以及 40 981 个项目，与 Beibei 和 Beidian 数据集不同的是，Gowalla 数据集仅保留了至少有 10 次交互的用户和项目. 采用与 LinkProp[28] 方法完全相同的实验设置，LinkCG 与基线方法在 Gowalla 数据集上的实验性能比较结果如表 6 所示. 实验结果表明，尽管 LinkCG 仅包含 2 个超参数，在 Gowalla 数据集上的推荐性能仍然优于当前

Table 6 Ablation Experiment of Two Hyper-Parameters in LinkCG

表 6 LinkCG 上 2 个超参数的消融实验

数据集	方法	Recall@20	NDCG@20
Gowalla	NGCF	0.156 0	0.132 4
	Mult-VAE	0.164 1	0.133 5
	LightGCN	0.182 3	0.155 3
	LinkProp	0.181 4	0.147 7
	LinkCG（本文）	0.184 0	0.155 4

先进的深度学习推荐方法. 同时, LinkCG 在性能上也超过了使用 3 个超参数的非深度学习方法 LinkProp[28], 这主要是因为 LinkCG 能够通过项目间 2 跳关系获取用户更多的偏好信息, 提升了推荐性能.

4 总 结

本文通过用户的历史交互项目之间的隐式关联关系解决链接预测问题, 提出了一种利用异构图进行链接预测的方法 LinkCG. 本文方法结合用户-项目交互 2 部图 G 中用户与项目的度赋予项目之间有向边可调节权重, 能够准确地捕捉项目间隐式关联关系的紧密程度并避免过平滑问题, 项目间链接的建立也缩短了用户到未交互项目的距离, 使得用户能够更快地链接到未交互的高阶项目. 同时, 通过获取项目间的 2 跳关系的方式缓解了数据稀疏性问题, 为交互较少的用户探索到更多可能偏好的项目. LinkCG 是只包含 2 个超参数的非深度学习模型, 在 3 个公开数据集上相比于复杂的先进模型实现了最佳性能, 本文的消融实验也表明了 LinkCG 中 2 个超参数可调节的必要性. 此外, 我们将 LinkCG 应用于社交关系数据上, 获取更多的用户偏好信息来提升推荐性能, 进一步证明 LinkCG 的适用性及有效性, 表明了 LinkCG 能够从用户历史交互的项目中获取足够丰富的偏好信息, 通过研究项目之间的隐式关联关系能够有效地为用户找到最可能偏好的项目.

作者贡献声明: 张惠鹞负责设计模型、提出实验方案及撰写论文; 黄钦阳负责完成对比实验并提出实验优化方案及修订论文; 胡诗彦负责实验指导及论文修改; 杨青负责修订及完善论文; 张敬伟负责提出研究和实验方向、对论文提出修改意见以及设计论文框架.

参 考 文 献

[1] Covington P, Adams J, Sargin E. Deep neural networks for YouTube recommendations[C] //Proc of the 10th ACM Conf on Recommender Systems. New York: ACM, 2016: 191−198

[2] Ren Hao, Liu Baisong, Sun Jinyang, et al. A time and relation-aware graph collaborative filtering for cross-domain sequential recommendation[J]. Journal of Computer Research and Development, 2023, 60(1): 112−124(in Chinese)
(任豪, 刘柏嵩, 孙金杨, 等. 基于时间和关系感知的图协同过滤跨域序列推荐[J]. 计算机研究与发展, 2023, 60(1): 112−124)

[3] Huang Ruoran, Cui Li, Han Chuanqi. Feature-over-field interaction factorization machine for sparse contextualized prediction in recommender systems[J]. Journal of Computer Research and Development, 2022, 59(7): 1553−1568 (in Chinese)
(黄若然, 崔莉, 韩传奇. 推荐系统中稀疏情景预测的特征-类别交互因子分解机[J]. 计算机研究与发展, 2022, 59(7): 1553−1568)

[4] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems[J]. Computer, 2009, 42(8): 30−37

[5] Wang Xiang, He Xiangnan, Wang Meng, et al. Neural graph collaborative filtering[C] //Proc of the 42nd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2019: 165−174

[6] He Xiangnan, Deng Kuan, Wang Xiang, et al. LightGCN: Simplifying and powering graph convolution network for recommendation[C] //Proc of the 43rd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2020: 639−648

[7] Fan Wenqi, Liu Xiaorui, Jin Wei, et al. Graph trend networks for recommendations[J]. arXiv preprint, arXiv: 2108.05552, 2021

[8] Huang Zan, Li Xin, Chen H. Link prediction approach to collaborative filtering[C] //Proc of the 5th ACM/IEEE-CS Joint Conf on Digital Libraries. New York: ACM, 2005: 141−142

[9] Chiluka N, Andrade N, Pouwelse J. A link prediction approach to recommendations in large-scale user-generated content systems[C] // Proc of the 33rd European Conf on IR Research. Berlin: Springer, 2011: 189−200

[10] Salton G, McGill M. Introduction to Modern Information Retrieval[M]. New York: McGraw-Hill, 1983

[11] Leicht E A, Holme P, Newman M E J. Vertex similarity in networks[J]. Physical Review E, 2006, 73(2): 026120

[12] Baluja S, Seth R, Sivakumar D, et al. Video suggestion and discovery for YouTube: Taking random walks through the view graph[C] //Proc of the 17th Int Conf on World Wide Web. New York: ACM, 2008: 895−904

[13] Eksombatchai C, Jindal P, Liu J Z, et al. Pixie: A system for recommending 3+ billion items to 200+ million users in real-time[C] //Proc of the 27th World Wide Web Conf. New York: ACM, 2018: 1775−1784

[14] Perozzi B, Al-Rfou R, Skiena S. Deepwalk: Online learning of social representations[C] //Proc of the 20th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2014: 701−710

[15] Chen C M, Tsai M F, Lin Y C, et al. Query-based music recommendations via preference embedding[C]//Proc of the 10th ACM Conf on Recommender Systems. New York: ACM, 2016: 79−82

[16] Yang J H, Chen C M, Wang C J, et al. HOP-rec: High-order proximity for implicit recommendation[C]// Proc of the 12th ACM Conf on Recommender Systems. New York: ACM, 2018: 140−144

[17] Chen C M, Wang C J, Tsai M F, et al. Collaborative similarity embedding for recommender systems[C]// Proc of the 28th World Wide Web Conf. New York: ACM, 2019: 2637−2643

[18] Tang Jian, Qu Meng, Wang Mingzhe, et al. Line: Large-scale information network embedding[C]//Proc of the 24th Int Conf on World Wide Web. New York: ACM, 2015: 1067−1077

[19] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[C/OL]//Proc of the 5th Int Conf on Learning Representations. 2017[2022-10-21].https://openreview.net/forum?id=SJU4ayYgl

[20] Velickovic P, Cucurull G, Casanova A, et al. Graph attention networks[C/OL]//Proc of the 6th Int Conf on Learning Representations. 2018[2022-10-21].https://openreview.net/forum?id=rJXMpikCZ

[21] Xu K, Hu Weihua, Leskovec J, et al. How powerful are graph neural networks[C/OL]//Proc of the 7th Int Conf on Learning Representations. 2019[2022-10-21].https://openreview.net/forum?id=ryGs6iA5Km

[22] Yu Junliang, Yin Hongzhi, Li Jundong, et al. Self-supervised multi-channel hypergraph convolutional network for social recommendation[C]//Proc of the 30th Web Conf. New York: ACM, 2021: 413−424

[23] Chang Jianxin, Gao Chen, Zheng Yu, et al. Sequential recommendation with graph neural networks[C]//Proc of the 44th Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2021: 378−387

[24] Wang Ziyang, Wei Wei, Cong Gao, et al. Global context enhanced graph neural networks for session-based recommendation[C] //Proc of the 43rd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2020: 169−178

[25] Chen Lei, Wu Le, Hong Richang, et al. Revisiting graph based collaborative filtering: A linear residual graph convolutional network approach[C] //Proc of the 32nd AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2020: 27−34

[26] Liben-Nowell D, Kleinberg J. The link prediction problem for social networks[C]//Proc of the 12th Int Conf on Information and Knowledge Management. New York: ACM, 2003: 556−559

[27] Zhang Muhan, Li Pan, Xia Yinglong, et al. Revisiting graph neural networks for link prediction[J]. arXiv preprint, arXiv: 2010.16103, 2020

[28] Fu H M, Poirson P, Lee K S, et al. Revisiting neighborhood-based link prediction for collaborative filtering[C] // Proc of the 31st Web Conf. New York: ACM, 2022: 1009−1018

[29] Anand R, Jeffrey D U. Mining of Massive Datasets[M]. Cambridge, UK: Cambridge University Press, 2011

[30] Lin T H, Gao Chen, Li Yong. Cross: Cross-platform recommendation for social e-commerce[C] //Proc of the 42nd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2019: 515−524

[31] Zhang Jun, Gao Chen, Jin Depeng, et al. Group-buying recommendation for social e-commerce[C] //Proc of the 37th IEEE Int Conf on Data Engineering (ICDE). Piscataway, NJ: IEEE, 2021: 1536−1547

[32] Wu Le, Sun Peijie, Fu Yanjie, et al. A neural influence diffusion model for social recommendation[C] //Proc of the 42nd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2019: 235−244

[33] Jamali M, Ester M. A matrix factorization technique with trust propagation for recommendation in social networks[C] //Proc of the 4th ACM Conf on Recommender systems. New York: ACM, 2010: 135−142

[34] Yang Liangwei, Liu Zhiwei, Dou Yingtong, et al. Consisrec: Enhancing GNN for social recommendation via consistent neighbor aggregation[C] //Proc of the 44th Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2021: 2141−2145

[35] Yu Junliang, Yin Hongzhi, Li Jundong, et al. Enhance social recommendation with adversarial graph convolutional networks[J]. arXiv preprint, arXiv: 2004.02340, 2020

[36] Sachdeva N, Manco G, Ritacco E, et al. Sequential variational autoencoders for collaborative filtering[C] //Proc of the 12th Int Conf on Web Search and Data Mining. New York: ACM, 2019: 600−608

[37] Chae D K, Kim J, Chau D H, et al. Ar-cf: Augmenting virtual users and items in collaborative filtering for addressing cold-start problems[C]//Proc of the 43rd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2020: 1251−1260

[38] Zhu Zirui, Gao Chen, Chen Xu, et al. Inhomogeneous social recommendation with hypergraph convolutional networks[J]. arXiv preprint, arXiv: 2111.03344, 2021



Zhang Huijuan, born in 1996. Master. Her main research interests include big data analysis and recommendation systems.

张惠鹃, 1996 年生. 硕士. 主要研究方向为大数据分析、推荐系统.



Huang Qinyang, born in 1996. Master. His main research interest includes artificial intelligence.

黄钦阳, 1996 年生. 硕士. 主要研究方向为人工智能.



Hu Shiyan, born in 1998. Master. Her main research interests include recommendation systems and big data analysis.

胡诗彦, 1998 年生. 硕士. 主要研究方向为推荐系统、大数据分析.



Yang Qing, born in 1976. Master, professor. Her main research interest includes intelligent information process and recommendation.

杨　青, 1976 年生. 硕士, 教授. 主要研究方向为智能信息处理、推荐.



Zhang Jingwei, born in 1977. PhD, professor. Member of CCF. His main research interests include big data management, analysis and smart applications.

张敬伟, 1977 年生. 博士, 教授. CCF 会员. 主要研究方向为大数据管理、分析和智能应用.