软件学报 ISSN 1000-9825, CODEN RUXUEW
[doi: 10.13328/j.cnki.jos.007165] [CSTR: 32375.14.jos.007165]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于 Rényi 差分隐私的图卷积协同过滤推荐算法*

王 锟[1,3], 王 永[1,2], 刘金源[1], 邓江洲[2]

[1](重庆邮电大学 计算机科学与技术学院, 重庆 400065)

[2](重庆邮电大学 经济管理学院, 重庆 400065)

[3](西南政法大学 教育信息技术中心, 重庆 401120)

通信作者: 王永, E-mail: wangyong1@cqupt.edu.cn

摘 要: 近年来, 图卷积网络作为一种强大的图嵌入技术在推荐系统领域得到广泛应用. 主要原因是推荐系统中大多数信息可以建模为图结构, 而图卷积网络是一种基于图结构的深度学习模型, 有助于挖掘图数据中用户和项目之间的潜在交互, 从而提高推荐系统的性能. 由于推荐系统的建模通常需要收集和处理大量的敏感数据, 因此可能会面临隐私泄露的风险. 差分隐私是一种具有坚实理论基础的隐私保护模型, 已被广泛应用于推荐系统中解决用户隐私泄露的问题. 目前基于差分隐私的研究主要是面向独立同分布的数据模型. 然而, 在基于图卷积网络的推荐系统中, 数据之间关联性强且不具有独立性, 这使得现有方法难以对其进行有效的隐私保护处理. 为解决该问题, 提出基于 Rényi 差分隐私的图卷积协同过滤推荐算法 RDP-GCF, 旨在保护用户与项目交互数据安全的前提下, 实现隐私性和效用性之间的平衡. 该算法首先利用图卷积网络学习用户/项目的嵌入向量; 然后, 采用高斯机制对嵌入向量进行随机化处理, 同时基于采样的方法放大隐私预算, 减少差分噪声注入量, 以提升推荐系统的性能; 最后, 通过加权融合的方式得到用户/项目的最终嵌入向量, 并应用于推荐任务. 在 3 组公开数据集上进行实验验证. 结果表明, 与现有同类方法相比, 所提算法能更好地实现隐私保护与数据效用之间的平衡.

关键词: 推荐系统; 协同过滤; 图卷积网络; 隐私保护; Rényi 差分隐私

中图法分类号: TP18

中文引用格式: 王锟, 王永, 刘金源, 邓江洲. 基于Rényi差分隐私的图卷积协同过滤推荐算法. 软件学报. http://www.jos.org.cn/1000-9825/7165.htm

英文引用格式: Wang K, Wang Y, Liu JY, Deng JZ. Graph Convolutional Collaborative Filtering Recommendation Algorithm Based on Rényi Differential Privacy. Ruan Jian Xue Bao/Journal of Software (in Chinese). http://www.jos.org.cn/1000-9825/7165.htm

Graph Convolutional Collaborative Filtering Recommendation Algorithm Based on Rényi Differential Privacy

WANG Kun[1,3], WANG Yong[1,2], LIU Jin-Yuan[1], DENG Jiang-Zhou[2]

[1](College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

[2](School of Economics and Management, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

[3](Education Information Technology Center, Southwest University of Political Science and Law, Chongqing 401120, China)

Abstract: Recently, graph convolutional network (GCN), as a powerful graph embedding technology, has been widely applied in the field of recommendation. The main reason is that most of the information in recommender systems can be modeled as graph-structured data, and GCN, as a deep learning model that operates on graph structures, helps to explore the potential interactions between users and items in graph-structured data, to enhance the performance of the recommender systems. Since the modeling of recommender systems usually needs to collect and process a large amount of sensitive data, it may face the risk of privacy leakage. Differential privacy, as a privacy protection model with a solid theoretical foundation, has been widely used in recommender systems to solve the problem of personal

* 基金项目: 国家自然科学基金 (62272077); 重庆市自然科学基金 (cstc2021jcyj-msxmX0557); 教育部人文社科规划项目 (20YJAZH102)
收稿时间: 2023-07-19; 修改时间: 2023-10-02; 采用时间: 2024-02-06; jos 在线出版时间: 2024-11-01

privacy leakage. Currently, the research based on differential privacy is mainly oriented to independent and identically distributed data models. However, data within GCN-based recommender systems is highly correlated and not independent, making the existing privacy protection methods less effective. To solve the problem, this study proposes a graph convolutional collaborative filtering recommendation algorithm based on Rényi differential privacy (RDP-GCF for short), aiming to achieve a balance between privacy protection and utility while ensuring the security ofuser-item interaction data. The algorithm first utilizes GCN techniques to learn the embedding vectors for users and items. Then, the Gaussian mechanism is used to randomize the embedding vectors, and a sampling-based method is used to amplify the privacy budget and minimize the injection of differential noise, thereby improving the performance of the recommender system. Lastly, the final embedding vectors of the users and items are obtained by a weighted fusion and applied to the recommendation tasks. The proposed algorithm is validated through experiments on three publicly available datasets. The results show that compared to existing similar methods, the proposed algorithm more effectively achieves a balance between privacy protection and data utility.

Key words: recommender system; collaborative filtering; graph convolutional network (GCN); privacy-preserving; Rényi differential privacy

随着互联网信息量爆炸式的增长, 信息过载问题日益突出. 为了缓解这一问题, 推荐系统被广泛部署于各种网络应用中, 向用户提供个性化的推荐服务. 在众多推荐方法中, 协同过滤是目前应用最广泛的算法之一, 它基于历史交互数据来学习用户和项目的嵌入向量, 用于预测用户对项目的偏好, 并将结果以个性化列表的形式推荐给用户. 在推荐系统领域, 大多数信息可以表示成图结构, 相较于列表、矩阵等传统表示方式, 图数据可以更好描述实体之间的相互作用和影响. 图神经网络是一种专门处理图数据的神经网络模型, 在图表示学习方面表现优异, 能充分利用图结构信息来丰富实体的嵌入表示[1]. 因此, 使用图神经网络作为推荐系统的核心算法已成为当前的一个研究热点. 其中, 图卷积网络 (graph convolutional network, GCN) 是推荐系统中最常用的一种图神经网络模型. 最近的研究表明[2−6], 相较于传统推荐方法, 基于 GCN 的推荐系统能够充分利用图的拓扑结构信息, 挖掘用户与项目之间的潜在交互, 从而显著提升推荐系统的性能.

尽管基于 GCN 的推荐系统能够有效提升用户的使用体验和满意度, 但是同时, 也面临着严峻的隐私安全问题[7,8]. 首先, 为了提高基于 GCN 的推荐系统的性能, 通常需要收集各种信息, 并将其构建为图结构数据用于训练模型. 然而, 在收集的信息中可能包含大量用户的敏感信息, 如兴趣爱好、习惯、社交关系等. 在不可信的集中式服务器环境下, 这些敏感信息可能会因数据泄露或未授权的使用, 而严重威胁到用户的隐私安全[9]. 其次, 攻击者可通过对发布的模型和预测结果进行分析, 发起训练集重构[10]、成员推断[11]和链接推理等攻[12,13], 从而导致用户个人隐私泄漏. 因此, 在开发基于 GCN 的推荐系统时, 必须高度重视并采取措施保障用户的隐私安全.

本文旨在使用差分隐私[14]技术保护推荐系统中用户与项目交互数据的安全. 差分隐私是一种具有坚实理论基础的隐私保护模型. 它为定义和保护隐私提供了严格的数学框架, 可以有效防止攻击者通过模型输出辨别某一条数据是否被用于训练机器学习模型, 从而保护训练集中每个个体的隐私. 近年来, 差分隐私技术被广泛应用于推荐系统领域, 以解决用户的隐私保护问题. 通常, 研究人员通过采用不同方式, 在算法中添加噪声来保护用户数据的隐私. 主要包括生成噪声数据[15]、在模型训练中对参数或梯度添加噪声[16,17]、对目标函数添加噪声[18]以及在模型的输出结果上添加噪声[19]等. 目前, 这些方法主要是针对传统独立同分布的数据模型进行隐私保护处理. 然而, 在基于 GCN 的推荐系统中, 由于节点之间存在复杂的关联性, 不同数据记录之间不再相互独立, 这使得它们无法直接应用于此类场景[9,20]. 为此, 最近一些研究者开始探索将差分隐私技术引入图神经网络模型中, 以保护用户的隐私. Olatunji 等人[21]设计了一种满足节点级差分隐私的学习算法. 但该方法需要使用公共图数据来协助训练, 导致其应用场景受到较大的限制. Daigavane 等人[22]将梯度扰动[23]方法拓展到图神经网络模型, 提出了一种保护节点隐私信息的方法, 使攻击者无法准确判断某个特定节点是否出现在原始图中. 但是, 该方法仅适用于 1 层的图神经网络模型, 因此无法利用来自高阶邻域的信息. Wu 等人[12]基于原始图的邻接矩阵, 采用输入扰动机制设计了一种边差分隐私算法. 然而, 由于邻接矩阵通常是高维且稀疏, 直接添加噪声可能会严重损害数据的可用性. 因此该方法只适用于隐私预算较大的情况. 综合来看, 现有的方法还没有达到预期的效果, 主要面临两个方面的问题: (1) 面向传统深度学习模型的差分隐私方法, 无法满足基于图神经网络的推荐系统隐私保护需求, 需要针对其特点开发适用的差分隐私算法; (2) 如何合理平衡隐私保护和模型效用之间的关系.

针对上述问题, 本文提出了一种基于 Rényi 差分隐私 (Rényi differential privacy, RDP)[24]的图卷积协同过滤推

荐算法 RDP-GCF. 首先, 将用户与项目的交互信息表示为图结构, 并利用 GCN 模型学习用户/项目的嵌入向量; 然后, 采用高斯机制对嵌入向量进行随机化处理, 以保护训练数据的安全; 最后, 通过加权融合的方式得到用户/项目的最终嵌入表示. 与现有方法相比, 本文所提方法采用基于特征参数的扰动技术, 可以在保留原始图结构特征的情况下实现边差分隐私. 此外, 在 RDP 框架下通过应用随机采样进行隐私放大, 降低了因添加噪声而带来的负面影响, 有利于提高推荐系统的性能. 本文主要贡献包括以下 3 个方面.

(1) 提出了一种新的基于差分隐私的图卷积协同过滤推荐算法. 其采用扰动特征参数的方法, 向用户/项目嵌入向量中添加高斯噪声, 可以防止攻击者通过对嵌入向量进行影响分析, 推断出用户与项目之间的关联性, 保障了交互数据的安全. 同时, 该算法无需对图数据进行输入扰动, 保留了原始图结构的特征, 能够有效提高数据的可用性.

(2) 在 Rényi 差分隐私的框架下, RDP-GCF 算法通过随机采样的方式进行隐私放大, 减少了差分噪声的注入量, 从而在相同隐私保证的情况下, 提高了推荐系统的性能.

(3) 通过理论证明了 RDP-GCF 算法满足边差分隐私, 并且分析出它的隐私预算上界. 在 3 组公开数据集上进行了实验验证. 结果表明, 与现有同类方法相比, 所提算法能更好地实现隐私保护与数据效用的兼顾.

本文第 1 节介绍国内外相关研究工作. 第 2 节介绍本文所需的基础知识并对研究问题进行描述. 第 3 节提出基于 Rényi 差分隐私的图卷积协同过滤推荐算法, 并对其详细描述. 第 4 节对本文方法的隐私性进行理论分析. 第 5 节通过实验验证所提方法的有效性. 最后, 第 6 节总结全文并展望后续研究工作.

1 相关工作

1.1 基于图神经网络的推荐系统

随着图神经网络技术的迅速发展, 基于图神经网络的推荐算法已成为当前一个热门的研究方向. Wang 等人[2]提出了一种基于图神经网络的协同过滤推荐算法 NGCF, 它将用户与商品的交互数据构建为图结构, 并使用图神经网络来实现信息的传递和融合, 从而提高推荐算法的表达能力和准确性. He 等人[4]提出了一种轻量级图卷积网络模型 LightGCN, 主要应用于协同过滤推荐任务. 该方法通过简化传统图卷积网络的更新操作, 可以有效减少计算的复杂性, 从而提高推荐系统的效率和性能. 为了解决大规模的实时推荐问题, Ying 等人[5]提出了一种基于 GCN 的推荐算法. 该算法通过使用动态卷积、小批量计算等技术, 使 GCN 具有更高的效率和可伸缩性, 从而能够适应大规模图数据集的推荐任务. Zhang 等人[3]针对序列推荐提出了一种动态图神经网络模型 DGSR. 该方法通过一个动态图连接不同的用户序列, 在考虑时间和序列信息的基础上, 探索用户和项目之间的交互行为, 以提高推荐系统的性能. Liu 等人[6]将图神经网络应用于基于知识图谱的推荐系统中, 提出了一种上下文图注意力网络 CGAT. 该算法考虑了用户对实体的个性化偏好, 利用图注意力机制捕捉实体在知识图谱中的上下文信息, 从而有效解决数据稀疏和冷启动问题. 上述研究表明, 基于图神经网络的推荐系统能够更好地挖掘实体之间潜在的关联关系, 并生成蕴含丰富信息的嵌入表示, 因此在提升推荐系统的性能方面展现出了良好的应用前景.

1.2 基于差分隐私的推荐系统

差分隐私是一种具备坚实理论基础和强隐私保证的安全机制, 近年来被广泛应用于推荐系统中保护用户的个人隐私信息. Hua 等人[18]提出一种差分隐私矩阵分解方法, 该方法通过向目标函数添加噪声进行扰动, 从而保护用户个体隐私. 鲜征征等人[16]将隐式因子模型 SVD++和差分隐私相结合, 提出了基于梯度扰动、目标函数扰动以及输出结果扰动的 3 种隐私保护模型. Liu 等人[25]提出一种快速、可证明的差分隐私矩阵分解算法 FDPMF, 它通过对用户评分矩阵进行降维和扰动, 在提高计算效率的同时, 保证了用户数据的隐私性. Shin 等人[26]提出了一种基于本地差分隐私的矩阵分解算法, 通过对输入数据进行随机扰动, 并将扰动后的数据发送给推荐系统进行聚合, 以满足差分隐私保护. 上述这些研究主要专注于针对独立同分布的数据模型设计满足差分隐私的方法. 但是在基于图神经网络的推荐系统中, 数据之间通常存在复杂的关联性, 因此它们难以直接应用于此类场景.

最近, 一些工作开始探索将差分技术应用于图神经网络模型中解决隐私保护问题, 并取得了一定进展.

Olatunji 等人[21]通过将 PATE[19]框架应用于图神经网络, 提出了一种满足节点级差分隐私的图神经网络算法. 它利用公共图数据训练教师图神经网络模型, 为每个查询节点进行隐私标记, 以免泄露敏感信息. 然而, 该方法对公共数据集依赖度较高, 具有很大的局限性. 为了抵御属性推理攻击, Zhang 等人[27]提出了一种基于差分隐私的图卷积网络模型 GERAI, 它采用本地差分隐私机制, 创新性地设计了一种双阶段加密范式, 对用户的敏感特征和模型优化过程实施隐私保护. 然而, 该方法不能用于将图结构作为敏感信息的应用场景. Daigavane 等人[22]通过将 DP-SGD[23]算法和隐私放大技术相结合, 提出了一种基于节点级差分隐私的图神经网络模型, 能够用于保护单个节点的隐私信息. 但是, 该方法仅适用于 1 层的图神经网络模型, 因此无法利用来自高阶邻域的信息, 而这对于图神经网络在许多应用中实现高精度是至关重要的. Wu 等人[12]基于原始图的邻接矩阵, 采用输入扰动机制设计了一种满足边差分隐私的保护方案. 但由于邻接矩阵通常是高维且稀疏, 直接添加噪声会极大破坏数据的可用性. 因此该方法只适用于隐私预算相对较大的情况.

综上所述, 现有基于差分隐私的图神经网络推荐算法, 主要存在对图数据扰动困难以及数据隐私性和效用性难以平衡等问题. 针对这些问题, 本文提出了一种基于 Rényi 差分隐私的图卷积协同过滤推荐算法. 该算法采用特征参数的扰动方法对节点嵌入进行随机化处理, 可以在保持原始图结构的情况下, 提供边差分隐私的保证. 此外, 该算法在 RDP 框架下, 通过随机采样进行隐私放大, 能有效减少差分噪声带来的负面影响, 从而在保护数据隐私的同时, 尽可能提高推荐系统的性能.

2 基础知识及问题描述

2.1 差分隐私

差分隐私是机器学习中常用的一种隐私保护技术, 它提供了严格的数学定义和坚实的理论基础. 其核心思想是: 设计随机算法, 保证任意个体的数据无论是否在数据集中, 对算法的输出结果几乎没有影响. 相关的定义如下.

定义 1. (ε, δ)-差分隐私[14]. 给定随机算法 M, $Range(M)$ 为 M 的输出域. 对于任意仅相差一条数据的相邻数据集 $D, D' \subset \mathcal{D}$, 若 M 在 D 与 D'上的任意输出结果 $S \in Range(M)$ 满足以下不等式, 则称 M 满足 (ε, δ)-差分隐私:

$$\Pr[M(D) \in S] \leqslant e^{\varepsilon} \times \Pr[M(D') \in S] + \delta \tag{1}$$

其中, $M(D)$ 和 $M(D')$ 分别代表算法 M 在数据集 D、D'上的输出; ε 为隐私预算, 用于衡量隐私保护级别; δ 表示隐私泄露的概率.

在本文中, 使用了传统差分隐私的一种拓展定义, 称为 Rényi 差分隐私. RDP 是 (ε, δ)-差分隐私的广义化, 它使用 Rényi 散度[24]来量化隐私泄露的程度, 能够提供更加精细和严格的隐私分析.

定义 2. Rényi 差分隐私. 给定随机算法 M, 设实数 $\alpha > 1$, 对于任意仅相差一条数据的相邻数据集 $D, D' \subset \mathcal{D}$, 若 M 在 D 与 D'上的任意输出结果满足下列不等式, 则称 M 满足 (α, ε)-RDP:

$$D_{\alpha}(M(D) \| M(D')) \leqslant \varepsilon \tag{2}$$

其中, $D_{\alpha}(P \| Q)$ 表示概率分布 P 和 Q 之间 α 阶的 Rényi 散度. 为方便起见, 本文将 RDP 的隐私参数 ε 记为 $\varepsilon_M(\alpha)$, 表示算法 M 满足 $(\alpha, \varepsilon_M(\alpha))$-RDP. 相关研究表明, 与 (ε, δ)-差分隐私相比, RDP 更适合于分析和量化各种隐私机制及其组合的隐私保护强度. 它提供了一种易于操作且定量准确的方法, 用于跟踪隐私机制在执行过程中的累积隐私损失, 可以更精确地表示其组合的隐私保证, 从而获得严格的总隐私损失约束. 目前, RDP 框架已成为隐私保护算法设计的重要理论基础和分析工具. 例如, Wang 等人[28]提出了一种基于 RDP 的子采样隐私分析机制 (analytical moments accountant), 该机制通过精细地跟踪和分析算法在整个训练过程中的隐私损失, 可以提供比传统强组合定理[23]更准确的总隐私损失估计. 这对于隐私算法的设计和优化具有重要意义, 使其可以在隐私保护与数据效用之间取得更好的平衡.

定义 3. 边级相邻图[12]. 给定图 $G_1 = (V_1, E_1)$, $G_2 = (V_2, E_2)$, 且 $V_1 = V_2$. 如果 G_1 可以通过在 G_2 中添加或删除一条边而得到, 即满足 $|E_1 \oplus E_2| = 1$, 则称 G_1 和 G_2 边级相邻图.

定义 4. L_2 边级敏感度. 给定一个函数 $f: G \to \mathbb{R}^d$, G_1 和 G_2 为任意两个边级相邻图, 则函数 $f: G \to \mathbb{R}^d$ 的 L_2

边级敏感度函数表示如下:

$$\Delta f = \max_{G_1,G_2} \|f(G_1) - f(G_2)\|_2 \tag{3}$$

其中, $f(G_1)$ 和 $f(G_2)$ 分别代表函数 f 在图 G_1、G_2 上的输出, $\|\cdot\|_2$ 是欧几里得范数.

定义 5. 边级 (α, ε)-RDP. 给定随机算法 M 和实数 $\alpha > 1$, 对于任意两个边级相邻图 G_1 和 G_2, 若 M 在图 G_1、G_2 上的任意输出结果满足以下不等式, 则称算法 M 满足边级 (α, ε)-RDP:

$$D_\alpha(M(G_1)\|M(G_2)) \leqslant \varepsilon \tag{4}$$

其中, $M(G_1)$ 和 $M(G_2)$ 分别代表算法 M 在图 G_1、G_2 上的输出.

定理 1. RDP 的序列组合性质[24]. 给定一个数据集 D 和 n 个随机化算法 $\{M_1, M_2, \ldots, M_n\}$, 算法 $M_i (1 \leqslant i \leqslant n)$ 满足 (α, ε_i)-RDP, 则这 n 个算法构成的序列组合 $M(M_1(D), \ldots, M_n(D))$ 满足 $\left(\alpha, \sum_{i=1}^{n} \varepsilon_i\right)$-RDP.

2.2 基于采样的隐私放大理论

子采样是设计和分析差分隐私机制的基本工具. Balle 等人[29]研究利用子采样技术增强差分隐私保证的问题, 并在此基础上提出了子采样隐私放大定理. 该定理的基本思想是, 从原始数据集中通过随机子采样选取一个子集, 并在该子集上应用差分隐私机制来保护数据隐私. 由于隐私保护机制仅能访问随机选取的子集, 而不是整个原始训练数据集, 这样便可以有效减少敏感个体数据的暴露风险, 从而实现增强隐私保证的效果. 对于不同的隐私机制, 通常需要根据实际情况调整采样策略和采样比例, 以获得更好的数据隐私性与效用性之间的平衡. 本节对基于 RDP 框架的子采样高斯机制进行介绍.

定理 2. 高斯机制[28]. 给定任意一个函数 $f: D \to \mathbb{R}^d$, 其 L_2 敏感度的大小为 1. 当 $\alpha > 1$ 时, 高斯机制 $G_\sigma(f(D)) = f(D) + (n_1, \ldots, n_d)$ 满足 (α, ε)-RDP, 其中 $(n_1, \ldots, n_d)$ 表示添加的噪声向量, n_i 是取自正态分布 $N\left(0, \frac{\alpha}{2\varepsilon}\right)$ 的独立同分布随机变量.

高斯机制是实现 RDP 的一种常用方法. 在 RDP 框架下, Mironov 等人[30]基于子采样隐私放大定理, 提出了一种采样高斯机制. 该机制通过将随机采样技术与高斯机制相结合来放大 RDP 的隐私约束, 从而可以获得更严格的隐私损失界限. 具体而言, 假设训练数据集 D 中的样本数量为 N, 以采样概率 $q = b/N$ 从数据集 D 随机选择一个大小为 b 的子集 d_s, 然后在该子集 d_s 上应用高斯机制 $G_\sigma(\cdot)$ 进行随机扰动. 这种针对随机选取的子集 d_s 应用的高斯机制被称为采样高斯机制, 记为 $G_\sigma \circ subsample$. 该机制比在整个训练集 D 上运行的高斯机制 $G_\sigma(\cdot)$ 提供了更强的隐私保证.

定理 3. 采样高斯机制的隐私成本上界[28]. 给定数据集 D, 采样概率 q 以及高斯机制 $G_\sigma(\cdot)$. 如果对于任意整数 $\alpha \geqslant 2$, 高斯机制 $G_\sigma(\cdot)$ 满足 $(\alpha, \varepsilon_g(\alpha))$-RDP. 则 $G_\sigma \circ subsample$ 满足 $\left(\alpha, \varepsilon_g^s(\alpha)\right)$-RDP, 其中,

$$\varepsilon_g^s(\alpha) \leqslant \frac{1}{\alpha - 1} \log\left(1 + q^2 \binom{\alpha}{2} \min\left\{4\left(e^{\varepsilon_g(2)} - 1\right), e^{\varepsilon_g(2)} \min\left\{2, \left(e^{\varepsilon_g(\infty)} - 1\right)\right\}\right\} + \sum_{j=3}^{\alpha} q^j \binom{\alpha}{j} e^{(j-1)\varepsilon_g(j)} \min\left\{2, \left(e^{\varepsilon_g(\infty)} - 1\right)^j\right\}\right) \tag{5}$$

2.3 问题描述

假定推荐系统中有 m 个用户和 n 个项目, 它们构成的集合分别为 $U = \{u_1, \ldots, u_i, \ldots, u_m\}$ 和 $P = \{p_1, \ldots, p_j, \ldots, p_n\}$. 用户-项目交互矩阵 $R = \{r_{i,j}\}_{m \times n}$ 是由用户 $u_i \in U$ 与项目 $p_j \in P$ 的交互信息组成. 本文将用户-项目交互矩阵 R 表示成图数据结构. 具体而言, 以用户集 U 和项目集 P 作为图的节点, 用户与项目的交互信息作为链接集合 E, 构建一个无向的用户-项目二部图 $G = (U, P, E)$, 其中, 对于任意链接 $e_{ij} = (u_i, p_j) \in E$, 表示用户 u_i 与项目 p_j 之间存在交互行为. 图 G 的结构信息可用邻接矩阵 $A = [A_{i,j}]_{m \times n}$ 表示, 如果 $A_{i,j} = 1$ 表示已观测到 u_i 与 p_j 之间的交互, 否则 $A_{i,j} = 0$. 如此, 推荐问题便可转化为图的链接预测问题. 下面为简化表达, 将用户和项目统称为节点.

在基于图卷积网络的推荐系统中, 首先是根据交互矩阵 R 构建用户-项目二部图 G, 并利用 GCN 模型学习节点的嵌入向量. 然后, 基于学到的嵌入向量生成用户与未交互项目之间的偏好分数. 最后, 将偏好分数最高的前 N 个项目推荐给用户, 以提供个性化的推荐服务. 然而, 由于个性化推荐涉及用户的历史交互数据, 直接发布推荐模

型的预测结果, 可能导致敏感信息的泄露. 因此, 需要采取隐私保护措施, 尽可能降低敏感信息泄露的风险, 以确保用户交互数据的安全. 本文要解决的关键问题是: 如何设计一种满足边级 (α, ε)-RDP 的图卷积协同过滤推荐算法, 来保护用户-项目图 G 中链接信息的安全, 并在隐私性和可用性之间获得良好平衡.

3 解决方案

3.1 算法的整体框架

本文提出的 RDP-GCF 算法的整体框架如图 1 所示, 主体部分由 3 个层模块组成. 模块 1 是隐私图卷积层, 它是一个多层结构, 其中的每一层都包括 3 个步骤: 图卷积计算、随机扰动、L_2 范数归一化. 该层主要负责从用户-项目二部图中学习节点的嵌入向量, 并对其进行差分隐私处理, 最后的输出是经过噪声处理后的节点嵌入向量; 模块 2 是融合层, 其主要负责将上游模块 (隐私图卷积层) 中得到的各层节点嵌入向量进行特征聚合, 以生成每个节点最终的综合嵌入表示; 模块 3 是预测层, 负责根据所得到的节点最终嵌入向量来预测用户与项目的偏好分数, 该预测分数为用户与项目嵌入的内积形式.

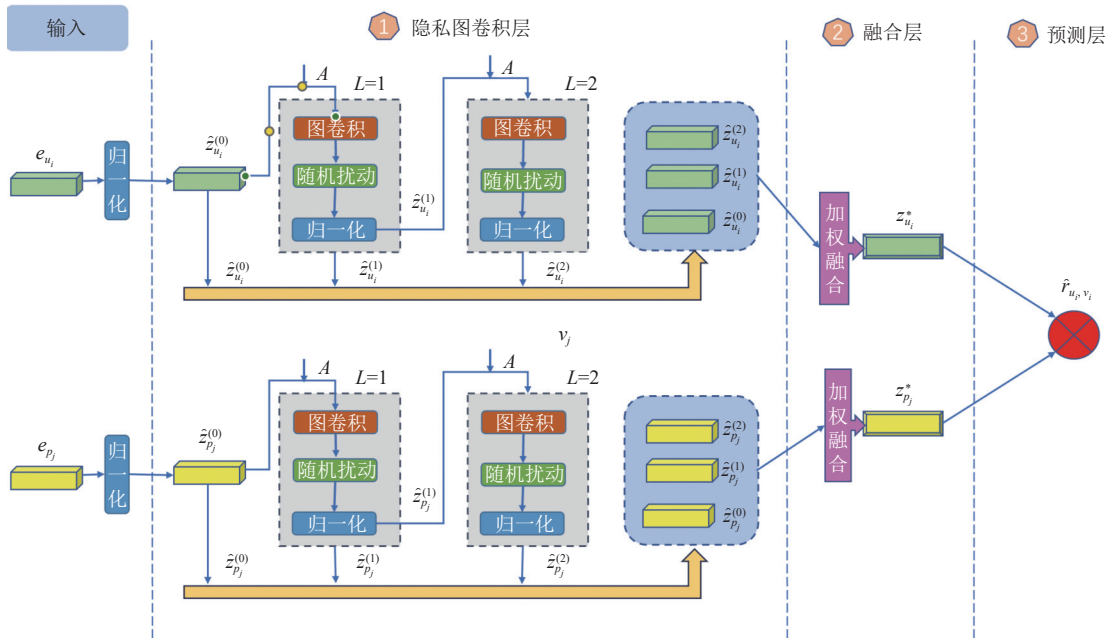

图 1 基于 RDP 的图卷积协同过滤推荐算法的整体框架

在 RDP-GCF 的整体框架中, 关键是利用隐私图卷积层来实现保护用户数据的隐私. 具体而言, 在训练过程中, 通过向每一层输出的节点嵌入中添加与敏感度成比例的高斯噪声, 以满足边级 (α, ε)-RDP 的保证. 这样设计的原因, 主要有以下几个方面:

(1) 在 RDP-GCF 中, 每个隐私图卷积层通过图卷积函数对节点嵌入进行更新. 假设在输入的用户-项目交互图中添加 (移除) 一条边, 这实际上可以被视为从该边目标顶点的输入中, 添加 (删除) 了一个数据样本 (节点), 从而引起该目标顶点的嵌入更新发生改变. 因此, 可以利用差分隐私机制向输出的节点嵌入中添加噪声来掩盖因单条边的增 (删) 而产生的影响变化, 以防止攻击者推断出目标节点之间的链接关系.

(2) RDP-GCF 是利用 GCN 来学习节点的嵌入, 每个节点的嵌入向量都会受到其 L 阶邻域中所有节点的影响. 因此即使是用户数据的一个微小变化, 可能经过嵌入传播而放大, 从而对大量节点的嵌入更新产生影响, 并导致敏感度过高的问题. 这将给差分隐私方案的设计带来了很大困难. 因此, 本文提出了一种分层的隐私处理方式, 即每

次随机化处理是仅针对单层图卷积网络, 这样只需考虑一阶邻域节点所产生的影响, 从而可以有效降低敏感度的大小, 并有利于隐私保护方案的设计和实现.

(3) 为确保 RDP-GCF 满足边级 (α, ε)-RDP, 需要在每层的节点嵌入更新后, 使用高斯机制对其进行随机扰动. 因此, 当 GCN 的层数为 L, 训练迭代次数为 T 时, 高斯机制共需要执行 TL 次. 如果基于定理 1 跟踪累积隐私损失, 可能会过分高估该损失值. 为缓解该问题, 本文在 RDP 框架下使用采样高斯机制添加噪声, 并利用定理 3 对每次执行该机制所产生的隐私损失进行严格核算, 以精确控制模型训练过程中的累积隐私损失. 这样可以有效减少差分噪声的注入规模, 在实现保护隐私的同时提高模型的推荐质量. 下面详细介绍各层模块.

3.2 隐私图卷积层

隐私图卷积层是 RDP-GCF 算法中最重要的组成部分, 主要负责提取节点的嵌入向量以及保护训练数据的隐私. 为了实现该层的目标任务, 本文设计了一种分层扰动机制 (layered perturbation mechanism, LPM). 具体地, 该机制首先利用用户-项目二部图 G 逐层进行图卷积操作更新节点的嵌入向量, 然后使用高斯机制向每一层所生成的节点嵌入向量中添加噪声, 隐藏其中所包含的敏感信息, 以保护训练数据的安全. 假设最大的层数为 L, 下面以生成用户 u_i 和项目 p_j 的嵌入向量为例, 介绍 LPM 机制的工作过程.

(1) 图卷积计算

首先, 本文基于 LightGCN 模型, 在用户-项目二部图 $G=(U,P,E)$ 上执行图卷积计算. 假设图 G 的邻接矩阵为 A, 第 $(l-1)$ 层输出的用户和项目嵌入矩阵为 $\hat{Z}_U^{(l-1)}=\left[\hat{z}_{u_1}^{(l-1)},\ldots,\hat{z}_{u_m}^{(l-1)}\right]$、$\hat{Z}_P^{(l-1)}=\left[\hat{z}_{p_1}^{(l-1)},\ldots,\hat{z}_{p_n}^{(l-1)}\right]$. 将它们作为输入, 计算第 l 层用户 u_i 和项目 p_j 的嵌入向量 $z_{u_i}^{(l)}\in\mathbb{R}^d$, $z_{p_j}^{(l)}\in\mathbb{R}^d$, 其计算方法如公式 (6) 所示:

$$\begin{cases} z_{u_i}^{(l)}=\sum_{p_j\in N_{u_i}}\tilde{A}_{i,j}\cdot\hat{z}_{p_j}^{(l-1)}=\tilde{A}_{i,*}\cdot\hat{Z}_P^{(l-1)}, & \forall u_i\in U \\ z_{p_j}^{(l)}=\sum_{u_i\in N_{p_j}}\tilde{A}_{i,j}\cdot\hat{z}_{u_i}^{(l-1)}=\tilde{A}_{*,j}\cdot\hat{Z}_U^{(l-1)}, & \forall p_j\in P \end{cases} \tag{6}$$

其中, N_{u_i} 表示与用户 u_i 相邻的项目节点集合, N_{p_j} 表示与项目 p_j 相邻的用户节点集合. $\tilde{A}=\left[\tilde{A}_{i,j}\right]_{m\times n}$ 为 A 的对称归一化邻接矩阵[4], 其计算公式如下:

$$\tilde{A}=D_U^{-\frac{1}{2}}\cdot A\cdot D_P^{-\frac{1}{2}},\quad \tilde{A}_{i,j}=\frac{A_{i,j}}{\sqrt{\left|N_{u_i}\right|\left|N_{p_j}\right|}} \tag{7}$$

其中, $D_U^{-\frac{1}{2}}\in\mathbb{R}^{m\times m}, D_P^{-\frac{1}{2}}\in\mathbb{R}^{n\times n}$ 分别表示用户和项目的对角度矩阵.

(2) 随机扰动

接着, 采用高斯机制向生成的节点嵌入向量中添加噪声进行随机扰动, 以确保训练数据集的隐私性, 其计算方法如公式 (8) 所示:

$$\begin{cases} \hat{z}_{u_i}^{(l)}=z_{u_i}^{(l)}+N\left(0,\sigma^2 I_d\right), & \forall u_i\in U \\ \hat{z}_{p_j}^{(l)}=z_{p_j}^{(l)}+N\left(0,\sigma^2 I_d\right), & \forall p_j\in P \end{cases} \tag{8}$$

其中, $N(0,\sigma^2 I_d)$ 表示 d 维高斯分布的概率密度函数, σ^2 表示高斯分布的方差.

(3) L_2 范数归一化

最后, 由于在随机扰动步骤中向每个节点的嵌入向量添加了高斯噪声, 这导致无法估计其对后续图卷积层输出的影响. 因此为了避免这种情况的发生, 类似于梯度裁剪[23], 需对随机扰动后的节点嵌入向量进行 L_2 范数归一化操作, 其计算方法如公式 (9) 所示:

$$\begin{cases} \hat{z}_{u_i}^{(l)}=\dfrac{\hat{z}_{u_i}^{(l)}}{\|\hat{z}_{u_i}^{(l)}\|_2}, & \forall u_i\in U \\ \hat{z}_{p_j}^{(l)}=\dfrac{\hat{z}_{p_j}^{(l)}}{\|\hat{z}_{p_j}^{(l)}\|_2}, & \forall p_j\in P \end{cases} \tag{9}$$

其中, $\|\cdot\|_2$ 表示欧几里得范数.

在 RDP-GCF 算法中, 隐私图卷积层模块, 通常是以多层叠加的方式组成, 并且每一层都将前一层输出的节点嵌入向量作为输入, 采用迭代的方式执行上述步骤, 直到达到设定的最大层数 L 为止. 这样就得到了每个用户 u_i 和项目 p_j 在各层的经过噪声处理后的嵌入向量, 分别为 $\left\{\hat{z}_{u_i}^{(1)},\ldots,\hat{z}_{u_i}^{(L)}\right\}$ 和 $\left\{\hat{z}_{p_j}^{(1)},\ldots,\hat{z}_{p_j}^{(L)}\right\}$.

LMP 机制的具体过程如算法 1 所示.

算法 1. LMP 机制算法.

输入: 用户-项目二部图 G 的邻接矩阵 A, 用户和项目的嵌入矩阵 $E_U \in \mathbb{R}^{m\times d}$ 、 $E_P \in \mathbb{R}^{n\times d}$, 最大层数 L, 高斯机制 $G_\sigma(\cdot)$ 的标准差 σ;

输出: 第 1–L 层的用户嵌入矩阵 $\{\hat{Z}_U^{(0)},\hat{Z}_U^{(1)},\ldots,\hat{Z}_U^{(L)}\}$ 和项目嵌入矩阵 $\{\hat{Z}_P^{(0)},\hat{Z}_P^{(1)},\ldots,\hat{Z}_P^{(L)}\}$.

1. 根据公式 (9) 分别对用户和商品的嵌入矩阵 E_U, E_P 进行 L_2 范数归一化处理得到 $\hat{E}_U$, $\hat{E}_P$;
2. $\hat{Z}_U^{(0)} \leftarrow \hat{E}_U, \hat{Z}_P^{(0)} \leftarrow \hat{E}_P$;
3. **for** $l \leftarrow 1$ **to** L **do**
4. **for** u_i **in** U, p_j **in** P **do**
5. $z_{u_i}^{(l)} \leftarrow \tilde{A}_{i,*}\cdot\hat{Z}_U^{(l-1)}$, $z_{p_j}^{(l)} \leftarrow \tilde{A}_{*,j}\cdot\hat{Z}_P^{(l-1)}$;
6. $\hat{z}_{u_i}^{(l)} \leftarrow z_{u_i}^{(l)} + N(\sigma^2 I_d)$, $\hat{z}_{p_j}^{(l)} \leftarrow z_{p_j}^{(l)} + N(\sigma^2 I_d)$;
7. $\hat{z}_{u_i}^{(l)} \leftarrow \dfrac{\hat{z}_{u_i}^{(l)}}{\|\hat{z}_{u_i}^{(l)}\|_2}$, $\hat{z}_{p_j}^{(l)} \leftarrow \dfrac{\hat{z}_{p_j}^{(l)}}{\|\hat{z}_{p_j}^{(l)}\|_2}$;
8. **end for**
9. **end for**
10. **return** 第 1–L 层的用户嵌入矩阵 $\{\hat{Z}_U^{(0)},\hat{Z}_U^{(1)},\ldots,\hat{Z}_U^{(L)}\}$ 和项目嵌入矩阵 $\{\hat{Z}_P^{(0)},\hat{Z}_P^{(1)},\ldots,\hat{Z}_P^{(L)}\}$.

3.3 融合层

融合层通过采取适当的方式对各层输出的节点嵌入向量进行聚合操作, 以综合利用所有层的表示信息, 从而改善模型的性能和鲁棒性. 本文方法根据预先设定的权重 α_l , 对第 l 层至第 L 层的节点嵌入向量采用加权平均的方式进行融合, 以得到每个用户和项目全局的嵌入向量如公式 (10) 所示:

$$\begin{cases} z_{u_i}^{*} = \sum_{l=0}^{L}\alpha_l \hat{z}_{u_i}^{(l)}, & \forall u_i \in U \\ z_{p_j}^{*} = \sum_{l=0}^{L}\alpha_l \hat{z}_{j}^{(l)}, & \forall p_j \in P \end{cases} \tag{10}$$

其中, α_l 表示第 l 层的权重参数.

3.4 预测层

预测层的任务是根据节点的嵌入向量计算每个用户与所有未交互项目之间的偏好分数. 本文采用用户与项目的嵌入内积进行计算, 因此用户 u_i 对项目 p_j 的偏好分数可表示为:

$$\hat{r}_{u_i,p_j} = z_{u_i}^{*\mathrm{T}}\cdot z_{p_j}^{*} \tag{11}$$

最后, 本文采用了 BPR (Bayesian personalized ranking)[31]方法, 基于正样本 (观察到的交互行为) 和负样本 (未观察到的交互行为) 构造成对排序损失函数, 并在此基础上, 使用 Adam[32]优化预测模型, 学习节点的嵌入向量. 其损失函数的定义如下:

$$L_{\mathrm{BPR}} = -\sum_{u_i=1}^{m}\sum_{p_j\in N_{u_i}}\sum_{p_n\notin N_{u_i}} \ln\sigma(\hat{r}_{u_i,p_j} - \hat{r}_{u_i,p_n}) + \lambda\|E\|^2 \tag{12}$$

其中, $\sigma(\cdot)$ 表示 Sigmoid 函数, λ 是正则化系数, Nu_i 表示与用户 u_i 有交互的项目集合, E 表示可训练的模型参数.

综合上述各节所述, 本文提出的方法如算法 2 所示.

算法 2. RDP-GCF 算法.

输入: 用户集合 U, 项目集合 P, 用户-项目交互矩阵 R, 最大层数 L, 最小批大小 B, 训练周期数 T;
输出: 用户和项目的最终嵌入矩阵 Z_U^*、 Z_P^*.

1. 基于用户-项目交互矩阵 R 构建用户-项目二部图 $G=(U,P,E)$;
2. 随机初始化用户和项目的嵌入矩阵 E_U、E_P;
3. **for** $t=1$ **to** T **do**
4. **for** 从训练集中随机抽样生成一个大小为 B 的训练批次 d_s **do**
5. 利用算法 1 生成用户和项目各层的嵌入矩阵 $\{\hat{Z}_U^{(0)},\hat{Z}_U^{(1)},\ldots,\hat{Z}_U^{(L)}\}$, $\{\hat{Z}_P^{(0)},\hat{Z}_P^{(1)},\ldots,\hat{Z}_P^{(L)}\}$;
6. 根据公式 (10) 执行层组合获得用户和项目的全局嵌入矩阵 Z_U^*, Z_P^*;
7. 根据公式 (11) 计算训练子集 d_s 中用户对项目的偏好程度 $\hat{R}_s$;
8. 根据公式 (12) 计算损失函数值, 并基于梯度下降法进行优化;
9. **end for**
10. **end for**
11. 返回所有用户和项目的最终嵌入矩阵 Z_U^*、 Z_P^*.

4 隐私保护分析

本节对提出的 RDP-GCF 方法进行隐私分析, 分析的过程包括 3 个部分: 首先, 计算函数的敏感度, 即分析输入数据中单个数据点的变化可导致函数输出的最大变化量; 其次, 基于敏感度的分析结果, 论证 LMP 机制满足边差分隐私; 最后, 根据 RDP 的序列组合性质, 进一步证明 RDP-GCF 方法满足边差分隐私.

定理 4. 给定 RDP-GCF 算法的图卷积函数 $Z_U=\tilde{A}\cdot X_P$ 、 $Z_P=\tilde{A}^{\mathrm{T}}\cdot X_U$, 其中, $\tilde{A}$ 表示归一化的邻接矩阵, $X_U=[x_{u_1},\ldots,x_{u_m}]$和$X_P=[x_{p_1},\ldots,x_{p_n}]$ 分别表示用户和项目的嵌入矩阵. 假设作为输入的 X_U 和 X_P 均已按行进行 L_2 范数归一化处理, 则对于任意两个边级相邻图 G 和图 G', 该图卷积函数的敏感度满足:

$$\Delta S_V=\begin{cases}\max\limits_{A,A'}\|Z_U-Z'_U\|_2=1, & V=U\\ \max\limits_{A,A'}\|Z_P-Z'_P\|_2=1, & V=P\end{cases} \tag{13}$$

证明: 设两个边级相邻图 $G=(U,P,E)$ 和 $G'=(U,P,E')$ 的邻接矩阵分别为 A、A'. 根据定义 3, 可得到相邻图 G 和图 G'之间仅相差一条边 e_{xy} , 即 $|E\oplus E'|=1$. 不失一般性, 假设 $e_{xy}\in E$ 且 $e_{xy}\notin E'$. 因此, 邻接矩阵 $A=[A_{i,j}]_{m\times n}$ 与 $A'=\left[A'_{i,j}\right]_{m\times n}$ 之间的关系满足:

$$\begin{cases}A_{i,j}=1, A'_{i,j}=0, & \text{if } i=x \text{ and } j=y\\ A_{i,j}=A'_{i,j}, & \text{else}\end{cases} \tag{14}$$

将 X_U 和 X_P 作为输入, 根据公式 (6), 使用图卷积操作对每个用户 u_i 和商品 p_j 的嵌入向量进行更新:

$$z_{u_i}=\tilde{A}_{i,*}\cdot X_P,\ \forall u_i\in U;\quad z_{p_j}=\tilde{A}_{*,j}\cdot X_U,\ \forall p_j\in P\ .$$

这里将更新后的用户和商品的嵌入矩阵分别记为: $Z_U=[z_{u_1},\ldots,z_{u_x},\ldots,z_{u_m}]$; $Z_P=\left[z_{p_1},\ldots,z_{p_y},\ldots,z_{p_n}\right]$.

接下来, 计算敏感度 ΔS_V, 当 $V=U$ 时:

$$\begin{aligned}\Delta S_V&=\max_{A,A'}\left\|Z_U-Z'_U\right\|_2=\max_{A,A'}\left\|[z_{u_1},\ldots,z_{u_x},\ldots,z_{u_m}]-[z'_{u_1},\ldots,z'_{u_x},\ldots,z'_{u_m}]\right\|_2\\&=\max_{A,A'}\left\|[(z_{u_1}-z'_{u_1}),\ldots,(z_{u_x}-z'_{u_x}),\ldots,(z_{u_m}-z'_{u_m})]\right\|_2\\&=\max_{A,A'}\left\|\left[\left(\tilde{A}_{1,*}\cdot X_P-\tilde{A}'_{1,*}\cdot X_P\right),\ldots,\left(\tilde{A}_{x,*}\cdot X_P-\tilde{A}'_{x,*}\cdot X_P\right),\ldots,\left(\tilde{A}_{m,*}\cdot X_P-\tilde{A}'_{m,*}\cdot X_P\right)\right]\right\|_2\ .\end{aligned}$$

根据公式 (6)、公式 (14) 进行化简得到:

$$= \max_{A,A'} \left\| \left[\left(\tilde{A}_{x,y} \cdot x_{p_y} - \tilde{A}'_{x,y} \cdot x_{p_y} \right) \right] \right\|_2 .$$

根据公式 (7) 得到:

$$\Delta S_U = \max_{A,A'} \left\| \left[\left(\frac{A_{x,y}}{\sqrt{|N_{u_x}||N_{p_y}|}} \cdot x_{p_y} - \frac{A'_{x,y}}{\sqrt{|N_{u_x}||N_{p_y}|}} \cdot x_{p_y} \right) \right] \right\|_2 \leqslant \max_{A,A'} \left\| \left[\left(A_{x,y} \cdot x_{p_y} - A'_{x,y} \cdot x_{p_y} \right) \right] \right\|_2 .$$

根据公式 (9), 则有:

$$\Delta S_V \leqslant \max_{A,A'} \| x_{p_y} \|_2 = 1 .$$

同理, 当 $V = P$ 时, 计算敏感度 ΔS_V 可得到:

$$\Delta S_V \leqslant \max_{A,A'} \| x_{u_x} \|_2 = 1 .$$

因此, 综上可得: $\Delta S_V = \begin{cases} \max\limits_{A,A'} \| Z_U - Z'_U \|_2 = 1, & V = U \\ \max\limits_{A,A'} \| Z_P - Z'_P \|_2 = 1, & V = P \end{cases}$.

引理 1. 给定隐私图卷积层的最大层数为 L, 高斯机制 $G_\sigma(\cdot)$ 满足 $(\alpha, \varepsilon_g(\alpha))$-RDP. 假设以概率 q 对训练集 D 进行随机采样得到子集 d_s, 并在子集 d_s 上应用算法 1 提出的 LPM 机制, 则对于任意整数 $\alpha \geqslant 2$, LPM 机制的输出满足边级 $(\alpha, \varepsilon_{\mathrm{LMP}}(\alpha))$-RDP, 其中,

$$\begin{aligned} \varepsilon_{\mathrm{LMP}}(\alpha) \leqslant & \frac{L}{\alpha - 1} \log \left(1 + q^2 \binom{\alpha}{2} \min \left\{ 4\left(\mathrm{e}^{\varepsilon_g(2)} - 1 \right), \mathrm{e}^{\varepsilon_g(2)} \min \left\{ 2, \left(\mathrm{e}^{\varepsilon_g(\infty)} - 1 \right) \right\} \right\} \right. \\ & \left. + \sum_{j=3}^{\alpha} q^j \binom{\alpha}{j} \mathrm{e}^{(j-1)\varepsilon_g(j)} \min \left\{ 2, \left(\mathrm{e}^{\varepsilon_g(\infty)} - 1 \right)^j \right\} \right) \end{aligned} \tag{15}$$

证明: 在算法 1 中, 步骤 6 采用高斯机制 $G_\sigma(\cdot)$ 向每个节点嵌入向量添加噪声 $N(0, \sigma^2 I_d)$. 高斯机制 $G_\sigma(\cdot)$ 满足 $(\alpha, \varepsilon_g(\alpha))$-RDP, 根据定理 2 和定理 4, 可得高斯机制的方差 $\sigma^2 = \dfrac{\alpha}{2\varepsilon_g(\alpha)}$. 由于是在随机选取的子集上应用 LMP 机制, 因此根据定理 3, 可获得在采样的情况下 $G_\sigma \circ subsample$ 满足 $\left(\alpha, \varepsilon_g^s(\alpha)\right)$-RDP, 且 $\varepsilon_g^s(\alpha)$ 满足:

$$\varepsilon_g^s(\alpha) \leqslant \frac{1}{\alpha - 1} \log \left(1 + q^2 \binom{\alpha}{2} \min \left\{ 4\left(\mathrm{e}^{\varepsilon_g(2)} - 1 \right), \mathrm{e}^{\varepsilon_g(2)} \min \left\{ 2, \left(\mathrm{e}^{\varepsilon_g(\infty)} - 1 \right) \right\} \right\} + \sum_{j=3}^{\alpha} q^j \binom{\alpha}{j} \mathrm{e}^{(j-1)\varepsilon_g(j)} \min \left\{ 2, \left(\mathrm{e}^{\varepsilon_g(\infty)} - 1 \right)^j \right\} \right) .$$

算法 1 中, 步骤 3 表示一共进行了 L 次随机扰动, 且每次添加噪声都是单独使用高斯机制 $G_\sigma(\cdot)$. 因此, 根据定理 1 可得, 对于任意整数 $\alpha \geqslant 2$, LMP 机制的输出能够满足边级 $(\alpha, \varepsilon_{\mathrm{LMP}}(\alpha))$-RDP, 且 $\varepsilon_{\mathrm{LMP}}(\alpha) = L\varepsilon_g^s(\alpha)$, 即:

$$\varepsilon_g^s(\alpha) \leqslant \frac{L}{\alpha - 1} \log \left(1 + q^2 \binom{\alpha}{2} \min \left\{ 4\left(\mathrm{e}^{\varepsilon_g(2)} - 1 \right), \mathrm{e}^{\varepsilon_g(2)} \min \left\{ 2, \left(\mathrm{e}^{\varepsilon_g(\infty)} - 1 \right) \right\} \right\} + \sum_{j=3}^{\alpha} q^j \binom{\alpha}{j} \mathrm{e}^{(j-1)\varepsilon_g(j)} \min \left\{ 2, \left(\mathrm{e}^{\varepsilon_g(\infty)} - 1 \right)^j \right\} \right) .$$

定理 5. 对于任意整数 $\alpha \geqslant 2$, RDP-GCF 算法满足边级 $\left(\alpha, \left\lceil \dfrac{T}{q} \right\rceil \varepsilon_{\mathrm{LMP}}(\alpha)\right)$-RDP, 其中 T 是算法训练周期数, q 为数据集的采样参数.

证明: 在 RDP-GCF 算法中, 步骤 3 说明整个训练过程一共需要完成 T 轮训练周期. 步骤 4–8 表示算法使用了小批次采样的训练方法来学习节点的嵌入向量, 因此每 1 轮训练周期的需要进行 $\left\lceil \dfrac{1}{q} \right\rceil$ 次的迭代, 且每次迭代都是单独使用 LMP 机制对随机选取的子集进行处理. 另外, 由引理 1 可知, 在采样的条件下, 单独使用 LMP 机制, 其输出满足 $(\alpha, \varepsilon_{\mathrm{LMP}}(\alpha))$-RDP.

综合上述分析, RDP-GCF 算法的整个训练过程总共需要进行 T 个周期完整批次的训练, 即可以看作是由 $\left\lceil \dfrac{T}{q} \right\rceil$ 个 LMP 机制构成的一个序列组合. 因此, 根据定理 1 可得, 对于任意整数 $\alpha \geqslant 2$, RDP-GCF 算法满足边级 $(\alpha, \varepsilon_{\mathrm{total}}(\alpha))$-RDP, 其中 $\varepsilon_{\mathrm{total}}(\alpha) = \left\lceil \dfrac{T}{q} \right\rceil \varepsilon_{\mathrm{LMP}}(\alpha)$.

5 实验与结果分析

本节设计了多组实验对本文方法进行评估, 主要目的是: (1) 验证 RDP-GCF 方法推荐的准确性; (2) 检测 RDP-GCF 方法在隐私与效用平衡方面的表现; (3) 分析重要参数对 RDP-GCF 方法推荐效果的影响. 本节组织如下: 第 5.1 节介绍本文的实验数据; 第 5.2 节介绍实验设置, 包括评价指标和对比方法; 第 5.3 节介绍实验的实现细节; 第 5.4 节给出实验结果, 并对实验结果进行详细分析.

5.1 数据集

为了验证本文方法的有效性, 本文选用了推荐领域中 3 个具有代表性的真实数据集来开展实验, 包括: MovieLens-1M、Gowalla 和 Yelp. 这些数据集分别涵盖了不同的应用场景, 并且规模和稀疏度也各不相同, 都被广泛用于推荐算法的性能评估. 同时, 它们也都在本文的对比方法中所使用, 这可以避免数据集对实验结果的影响, 有助于保证比较的客观性和公平性.

MovieLens-1M[33]是由 MovieLens 网站收集并提供的真实电影评级数据集, 主要包含用户对电影的评分数据、电影属性、标签以及用户人口统计特征等信息, 本文使用的是包含一百万条评分的版本, 并参考 Rendle 等人[31]的处理方式, 将其评分信息转化为隐式反馈再进行计算. Gowalla[2]是一个基于位置服务和社交网络分析的数据集, 它收集了用户在 Gowalla 社交网络平台上的位置签到记录和社交关系数据. Yelp[2]是一个常用的商业评论数据集. 该数据集包含了用户对商家的评论、评分、信任关系等, 其中评论和评分信息可以被用来分析用户和商家的偏好及特征. 表 1 给出了实验数据集的具体统计信息.

表 1 数据集的统计信息

数据集	用户数	商品数	关系数	密度
MovieLens-1M	6 040	3 952	1 000 209	0.044 68
Gowalla	29 858	40 981	1 027 370	0.000 84
Yelp	31 668	38 048	1 561 406	0.001 30

5.2 实验设置

5.2.1 评价指标

本文是一个典型的 Top-N 推荐问题. 因此, 实验采用了 Top-N 推荐系统中两种主要的评价指标, 即召回率 Recall@N 和归一化折损累积增益 NDCG@N. 其中, Recall@N 表示在用户真正感兴趣的项目中, 被正确预测到 Top-N 推荐列表中的比例, 可用来衡量推荐结果的查全率. NDCG@N 是基于排序结果的评价指标, 可用来衡量推荐结果的排序质量. 因此, 通过使用 Recall@N 和 NDCG@N 能够从推荐结果的查全率和排序性角度综合反映出推荐方法的有效性. 默认情况下, 本文中 N 的值设定为 20.

5.2.2 对比方法

为了全面客观地评估 RDP-GCF 方法的有效性, 本文选择了具有代表性的两组方法进行比较, 包括非隐私保护推荐方法 (BPRMF、NGCF、LightGCN) 和隐私保护推荐方法 (LapGraph、EdgeRand). 其中, BPRMF、NGCF、LightGCN 都是经典的推荐方法, 但未考虑隐私问题. 通过与它们进行对比, 可以评估本文方法是否能在满足边差分隐私的同时, 仍然保持良好的推荐性能. 而 LapGraph 和 EdgeRand 都是满足边差分隐私的方法, 区别是所使用的隐私机制不同. 通过与它们的对比, 可以检验本文方法在平衡隐私与效用方面的表现. 所用的对比方法如下.

(1) BPRMF[31]: 一种传统的矩阵分解模型, 通过最小化成对排序损失函数实现个性化推荐, 通常被用作推荐排序算法的基线.

(2) NGCF[2]: 一种典型的基于图神经网络的协同过滤推荐方法. 通过构建用户-项目交互图, 然后利用图神经网络提取图中高阶协同信息, 来学习用户和项目的嵌入.

(3) LightGCN[4]: 一种轻量级图卷积协同过滤方法, 是在 NGCF 方法基础上进行的改进. 该模型通过在交互图上线性传播来学习嵌入表示, 将不同传播层的嵌入信息的加权和作为最终嵌入, 并取消特征转换和非线性激活, 使其更

简洁和高效, 更适合于推荐. LightGCN 是 RDP-GCF 方法的基础模型, 同时也是本文在推荐性能方面的对比基线.

(4) LapGraph[12]: 一种满足边差分隐私的方法. 该方法基于输入扰动策略, 利用拉普拉斯机制向原始图的邻接矩阵中添加噪声, 并通过对扰动后的数据进行去噪, 来提高数据的可用性.

(5) EdgeRand[12]: 一种满足边差分隐私的方法. 该方法基于原始图的拓扑结构, 采用随机响应机制直接修改原始图, 即以概率 μ 随机翻转每条边的存在状态, 以保证图结构数据的安全.

需要说明的是, 本文中所有的隐私保护对比方法, 全部采用 LightGCN 作为基础模型并进行训练和比较. 这样设置可以消除因基础模型不同而对实验结果产生的影响, 有助于保证比较的客观性和公平性.

5.3 实验细节

本实验首先对数据集进行预处理. 对于所有的数据集均使用 10-core 设置 (即只保留至少有过 10 次交互的用户和项目), 以提高数据集的质量. 在此基础上, 对于数据集中的每个用户随机选择 80% 的历史交互数据作为训练集, 其余 20% 作为测试集. 同时, 从训练集中, 随机选择 10% 的交互数据作为验证集用于调整超参数. 最后, 基于训练集构建三元组, 将观察到的每个用户与项目的交互记录视为一个正例项目, 并使用负采样策略[2]生成对应的负例项目. 这种三元组形式可表示为{用户, 正例项目, 负例项目}.

实验中, RDP-GCF 方法的参数设置如下: (1) 模型的嵌入维度固定为 64, 使用 Xavier 方法[34]初始化嵌入参数; (2) 使用 Adam[32]优化器对模型进行优化, 并设置批大小为 1 024; (3) 超参数方面, 设置学习率为 1E–3, L_2 正则化系数 λ 为 1E–4; (4) 图卷积网络的层数 L 在{1, 2, 3, 4}中选择, 并将层组合系数 α_l 统一设置为 $1/(1+L)$. 在默认情况下, 本文方法的最大层数 L 设置为 3. 此外, 对于 BPRMF 和 NGCF 方法, 本文根据对应的论文进行参数初始化, 并微调使其达到最佳性能. 对于 LightGCN、LapGraph 以及 EdgeRand 这 3 种方法, 均使用与本文方法相同的参数设置, 这样有助于保证比较的公平性.

默认情况下, 所有模型都进行 300 个周期完整批次的训练, 并采取早期停止策略, 即如果连续 10 个周期内验证数据集上的 Recall@20 和 NDCG@20 指标没有增加, 则提取终止训练. 所有实验均使用不同的随机种子运行 5 次, 然后取评价指标的均值作为最终实验结果. 实验环境为 Intel Core™ i7-11700@2.5 GHz 处理器, 16 GB 内存, Windows 10 操作系统, 所有实验程序均在 PyTorch 上实现.

5.4 结果分析

5.4.1 与非隐私保护推荐方法的比较

为了验证 RDP-GCF 方法的整体推荐性能, 本文将其与 3 个具有代表性的非隐私保护推荐方法进行比较, 即 BPRMF、NGCF 和 LightGCN. 其中, RDP-GCF 方法的隐私预算 ε 设定为 5. 表 2 给出了各方法的具体实验结果, 本文方法和最优方法的评价指标采用加粗显示. 从表 2 中可得出以下结论.

表 2 在 3 个数据集上的各模型性能对比

数据集	评价指标	模型			
		BPRMF	NGCF	LightGCN	RDP-GCF (ε=5)
MovieLens-1M	Recall@20	0.204 3	0.220 1	**0.252 1**	**0.232 2**
	NDCG@20	0.307 4	0.354 1	**0.385 1**	**0.358 3**
Gowalla	Recall@20	0.129 1	0.156 9	**0.182 3**	**0.151 2**
	NDCG@20	0.110 9	0.132 7	**0.155 4**	**0.129 0**
Yelp	Recall@20	0.043 3	0.057 9	**0.063 9**	**0.050 5**
	NDCG@20	0.035 4	0.047 7	**0.052 5**	**0.041 1**

(1) 与传统的矩阵分解方法 BPRMF 相比, RDP-GCF 方法在两种评价指标上均有明显的提升. 具体而言, 在 3 个数据集上, 本文方法的 Recall@20 和 NDCG@20 指标比 BPRMF 方法分别平均提升了 15.8% 和 16.3%. 这是因为 BPRMF 是传统的矩阵分解方法, 只能利用用户与物品的相似性进行推荐, 缺乏考虑用户和项目之间更复杂的相关性, 因而导致其性能较差. 相比之下, RDP-GCF 方法是采样 LightGCN 作为基础模型, 能够利用图的拓扑结构深入挖掘用户与项目之间的高阶协同信息, 从而获得更准确的用户和项目的特征表示. 因此, 即便是受到差分噪声

的影响, 仍然可以显著提升推荐系统的性能.

(2) 与基于图神经网络的推荐方法相比, RDP-GCF 在两种评价结果上与 NGCF 的表现相近, 但要略逊于表现最好的 LightGCN 方法. 这主要有两个原因: 第一, RDP-GCF 方法虽然是采用 LightGCN 作为基础模型, 但在训练过程引入了差分隐私机制, 即需要通过向模型参数注入噪声来保护训练数据的隐私. 由于这种差分噪声的注入会不可避免对模型的训练产生负面影响, 因此导致其性能相较于 LightGCN 方法有所下降; 第二, 本文方法设计的隐私保护机制能够有效控制噪声添加的规模, 对于推荐性能的影响比较小, 从而可获得与 NGCF 方法相当的表现.

综上所述, RDP-GCF 方法在满足边差分隐私的前提下, 仍然能够充分发挥图神经网络在嵌入学习中的优势, 使其推荐性能具有良好的竞争力.

5.4.2 与隐私保护推荐方法的比较

为了检验 RDP-GCF 方法在平衡隐私性和效用性方面的表现, 本节将它与现有同类型的隐私保护方法进行比较, 包括 LapGraph 和 EdgeRand. 两种对比方法都采用了基于原始图的输入扰动策略. 实验中, 对各隐私方法共设置了 4 种级别的隐私预算, 分别为 1、3、5、10. 图 2 以曲线的形式展示各方法在每种隐私预算 ε 下的实验结果, 从中可以得出以下结论.

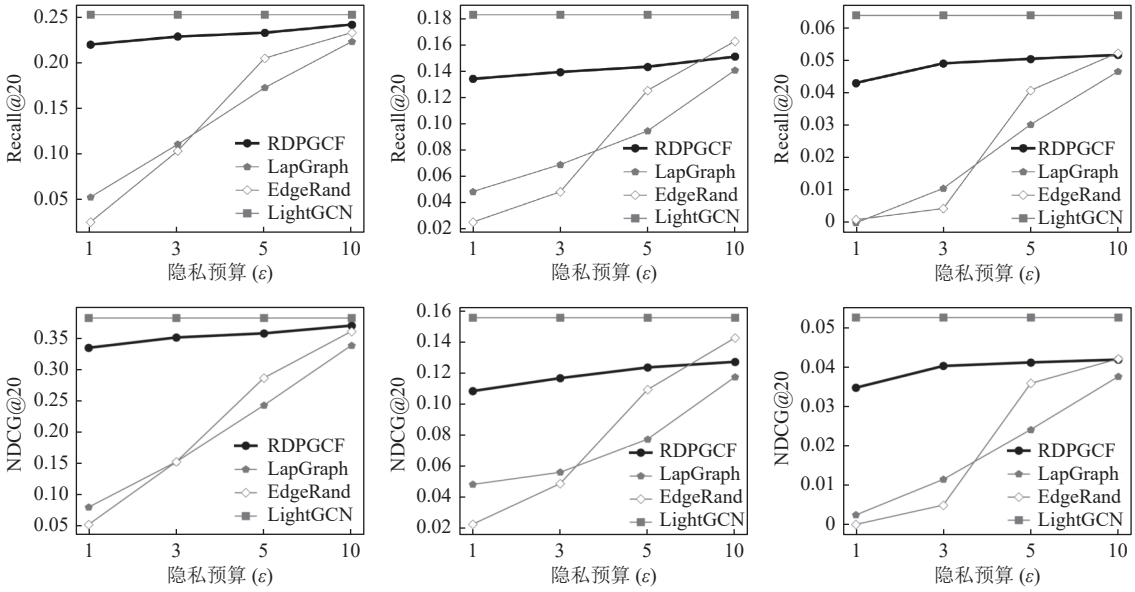

图 2 各模型在不同隐私预算下的性能表现

(1) 随着隐私预算 ε 增大, 各隐私方法的推荐性能均呈现上升趋势. 这与直观理解相符, 因为随着 ε 的增大, 各隐私方法的噪声注入量会相应地减少, 从而使得它们的推荐性能逐渐提高.

(2) 从整体来看, 在各隐私级别下, RDP-GCF 方法大多数时候的表现都要优于两个对比方法, 尤其是在隐私预算较低时, 其优势会更加明显. 例如, 当 $\varepsilon = 5$ 时, RDP-GCF 方法在 3 个数据集上的 Recall@20 和 NDCG@20 指标, 分别比最佳竞争对手 EdgeRand 高出 18.2% 和 17.8 %.

(3) 从评价指标的变化幅度分析发现, RDP-GCF 方法的评价指标变化幅度较小, 且始终能保持较高的性能水准. 相比之下, LapGraph 和 EdgeRand 方法随着隐私预算 ε 的变化, 其评估指标变化幅度非常大. 其主要原因可能是: 第一, LapGraph 和 EdgeRand 方法都是基于原始图的邻接矩阵, 采用各自的隐私保护机制直接进行随机化处理. 然而, 通常原始图的邻接矩阵是高维且稀疏的, 在隐私预算有限情况下, 由于直接向邻接矩阵添加了较多的噪声, 会严重破坏数据的可用性, 进而导致模型的推荐性能急剧下降. 因此, 这些方法具有较大的局限性, 一般仅适合在较高的隐私预算下使用; 第二, 本文方法对基于图神经网络的推荐系统具有更好的适用性. 一方面, 本文方法采

用了基于模型参数的扰动策略, 可以避免直接对原始图数据进行扰动; 另一方面, 它还利用采样高斯机制进行隐私放大, 从而有效降低了差分噪声对模型性能的影响. 因此, 即使在隐私预算有限的情况下, 本文方法仍然可以保证具有较高且相对稳定的推荐性能.

综上可知, 在各种隐私级别下, 本文方法的推荐性能通常要明显优于两个对比方法, 且始终能够保持相对稳定的推荐效果. 这充分表明, RDP-GCF 方法能够获得更好的隐私与效用之间的平衡.

5.4.3 重要参数对推荐效果的影响

本节, 分析了 RDP-GCF 方法中两种重要参数对推荐效果的影响, 包括图卷积的层数 L 和节点的嵌入维度 d. 实验中, 本文方法的隐私预算 ε 设定为 5, 并通过调整相应的参数, 观察其推荐性能的变化.

(1) 分析图卷积层数 L 对 RDP-GCF 方法的影响. 表 3 展示了本文方法在不同的 L ($L \in \{1,2,3,4\}$) 设置下的实验结果. 从中可以看出, 初始阶段随着层数 L 增大, RDP-GCF 方法的性能得到了一定提升. 然而, 当 L 增加到一定程度后, 模型性能便趋于稳定. 这是因为, 基于图神经网络的推荐模型, 随着层数 L 的增加, 可以更好挖掘图的结构特征, 从而提高模型的泛化能力, 并最终达到稳定的状态. 但是, 当超过一定限制, 再继续增大 L 时, 容易导致过拟合的问题, 模型的性能便不再提升, 甚至可能会下降. 与此同时, 随着层数 L 的增加, 也将注入更多差分噪声, 这对模型的性能会产生更大的负面影响.

表 3 图卷积层数的大小对 RDP-GCF 推荐性能的影响

最大层数	MovieLens-1M		Gowalla		Yelp	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
$L=1$	0.2287	0.3530	0.1453	0.1195	0.0491	0.0400
$L=2$	0.2258	0.3502	0.1449	0.1193	0.0505	0.0409
$L=3$	**0.2322**	**0.3583**	**0.1469**	**0.1233**	**0.0512**	**0.0409**
$L=4$	0. 2314	0.3575	0.1429	0.1164	0.0505	0.0401

(2) 分析节点嵌入维度 d 对本文方法的影响. 表 4 展示了本文方法在不同的嵌入维度 d ($d \in \{16,32,64,128\}$) 设置下的实验结果. 从中可以看出, 随着嵌入维度 d 增加到 64, RDP-GCF 方法性能的提升幅度较大. 之后, 当 d 继续增大, 模型的性能提升幅度减小并趋于稳定. 其原因在于: 维度较大的嵌入可以编码更多的信息, 则表达能力更强, 但是嵌入维度过高也可能会产生过拟合的问题, 导致模型的推荐性能受到影响. 上述实验结果说明, 在选择图卷积层参数 L 和嵌入维度参数 d 时, 需要综合考虑各种影响因素, 以满足推荐方法在平衡隐私保护和性能上的要求.

表 4 嵌入维度的大小对 RDP-GCF ($\varepsilon = 5$) 推荐性能的影响

嵌入维度	MovieLens-1M		Gowalla		Yelp	
	Recall@20	NDCG@20	Recall@20	NDCG@20	Recall@20	NDCG@20
$d=16$	0.2067	0.3235	0.1271	0.1081	0.0462	0.0375
$d=32$	0.2265	0.3489	0.1369	0.1159	0.0492	0.0400
$d=64$	**0.2322**	**0.3583**	**0.1469**	**0.1233**	**0.0512**	**0.0409**
$d=128$	0.2362	0.3608	0.1508	0.1317	0.0514	0.0412

6 总 结

现有基于差分隐私的推荐方法大多存在对图数据扰动困难以及数据隐私性和效用性难以平衡的问题. 针对该问题, 本文提出了一种基于 RDP 的图卷积协同过滤算法 RDP-GCF. 该算法不需要对图数据训练集进行输入扰动, 可以在保留原始图数据结构特征的情况下实现边差分隐私. 这有效缓解了因引入过量噪声, 而导致数据可用性严重降低的问题. 此外, 在 RDP 框架下通过应用随机采样进行隐私放大, 减少了差分噪声的注入量, 在保证隐私防御效果的前提下, 提升了推荐系统的性能. 实验结果表明, 本文所提出的方法在多个评价指标下的表现均优于现有方法, 能够更好地实现隐私保护和推荐性能的平衡. 在后续的工作中主要考虑以下两个方面: (1) 如何提高差分隐私算法的效率, 减少计算和存储开销, 以便更好地满足大规模图结构数据下的推荐应用需求; (2) 如何在分布式存储

场景下, 提出可行的差分隐私方法, 在保证基于图神经网络的推荐系统可用性的同时, 保护用户数据的安全.

References:

[1] Ge Y, Chen SC. Graph convolutional network for recommender systems. Ruan Jian Xue Bao/Journal of Software, 2020, 31(4): 1101−1112 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5928.htm [doi: 10.13328/j.cnki.jos.005928]

[2] 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]

[3] Zhang MQ, Wu S, Yu XL, Liu Q, Wang L. Dynamic graph neural networks for sequential recommendation. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(5): 4741−4753. [doi: 10.1109/TKDE.2022.3151618]

[4] He XN, Deng K, Wang X, Li Y, Zhang YD, Wang M. LightGCN: Simplifying and powering graph convolution network for recommendation. In: Proc. of the 43rd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. New York: ACM, 2020. 639−648. [doi: 10.1145/3397271.3401063]

[5] Ying R, He RN, Chen KF, Eksombatchai P, Hamilton WL, Leskovec J. 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]

[6] Liu Y, Yang SS, Xu YH, Miao CY, Wu M, Zhang JY. Contextualized graph attention network for recommendation with item knowledge graph. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(1): 181−195. [doi: 10.1109/TKDE.2021.3082948]

[7] Dai EY, Zhao TX, Zhu HS, Xu JJ, Guo ZM, Liu H, Tang JL, Wang SH. A comprehensive survey on trustworthy graph neural networks: Privacy, robustness, fairness, and explainability. arXiv:2204.08570, 2022.

[8] Jiang HL, Pei J, Yu DX, Yu JG, Gong B, Cheng XZ. Applications of differential privacy in social network analysis: A survey. IEEE Trans. on Knowledge and Data Engineering, 2023, 35(1): 108−127. [doi: 10.1109/TKDE.2021.3073062]

[9] Liu YH, Chen H, Liu YX, Zhao D, Li CP. State-of-the-art privacy attacks and defenses on graphs. Chinese Journal of Computers, 2022, 45(4): 702−734 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2022.00702]

[10] Duddu V, Boutet A, Shejwalkar V. Quantifying privacy leakage in graph embedding. In: Proc. of the 17th MobiQuitous EAI Int'l Conf. on Mobile and Ubiquitous Systems: Computing, Networking and Services. Darmstadt: ACM, 2020. 76−85. [doi: 10.1145/3448891.3448939]

[11] Zhang ZK, Chen M, Backes M, Shen Y, Zhang Y. Inference attacks against graph neural networks. In: Proc. of the 31st USENIX Security Symp. Boston: USENIX Association, 2022. 4543−4560.

[12] Wu F, Long YH, Zhang C, Li B. LINKTELLER: Recovering private edges from graph neural networks via influence analysis. In: Proc. of the 2022 IEEE Symp. on Security and Privacy (SP). San Francisco: IEEE, 2022. 2005−2024. [doi: 10.1109/SP46214.2022.9833806]

[13] He XL, Jia JY, Backes M, Gong NZ, Zhang Y. Stealing links from graph neural networks. In: Proc. of the 30th USENIX Security Symp. Berkeley: USENIX Association, 2021. 2669−2686.

[14] Dwork C, Roth A. The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science, 2014, 9(3−4): 211−407. [doi: 10.1561/0400000042]

[15] Xie LY, Lin KX, Wang S, Wang F, Zhou JY. Differentially private generative adversarial network. arXiv:1802.06739, 2018.

[16] Xian ZZ, Li QL, Huang XY, Lv W, Lu JY. Collaborative filtering via SVD++ with differential privacy. Control and Decision, 2019, 34(1): 43−54 (in Chinese with English abstract). [doi: 10.13195/j.kzyjc.2017.0961]

[17] Adesuyi TA, Kim BM. A layer-wise perturbation based privacy preserving deep neural networks. In: Proc. of the 2019 Int'l Conf. on Artificial Intelligence in Information and Communication (ICAIIC). Okinawa: IEEE, 2019. 389−394. [doi: 10.1109/ICAIIC.2019.8669014]

[18] Hua JY, Xia C, Zhong S. Differentially private matrix factorization. In: Proc. of the 24th Int'l Conf. on Artificial Intelligence. Buenos: AAAI Press, 2015. 1763−1770.

[19] Papernot N, Abadi M, Erlingsson Ú, Goodfellow IJ, Talwar K. Semi-supervised knowledge transfer for deep learning from private training data. arXiv:1610.05755, 2016.

[20] Wu ZQ, Hu J, Tian YP, Shi WC, Yan J. Privacy preserving algorithms of uncertain graphs in social networks. Ruan Jian Xue Bao/Journal of Software, 2019, 30(4): 1106−1120 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5368.htm [doi: 10.13328/j.cnki.jos.005368]

[21] Olatunji IE, Funke T, Khosla M. Releasing graph neural networks with differential privacy guarantees. arXiv:2109.08907, 2021.

[22] Daigavane A, Madan G, Sinha A, Thakurta AG, Aggarwal G, Jain P. Node-level differentially private graph neural networks. arXiv:2111.15521, 2021.

[23] Abadi M, Chu A, Goodfellow I, McMahan HB, Mironov I, Talwar K, Zhang L. Deep learning with differential privacy. In: Proc. of the 2016 ACM SIGSAC Conf. on Computer and Communications Security. Vienna: ACM, 2016. 308–318. [doi: 10.1145/2976749.2978318]

[24] Mironov I. Rényi differential privacy. In: Proc. of the 30th IEEE Computer Security Foundations Symp. (CSF). Santa Barbara: IEEE, 2017. 263–275. [doi: 10.1109/CSF.2017.11]

[25] Liu ZQ, Wang YX, Smola A. Fast differentially private matrix factorization. In: Proc. of the 9th ACM Conf. on Recommender Systems. Vienna: ACM, 2015. 171–178. [doi: 10.1145/2792838.2800191]

[26] Shin H, Kim S, Shin J, Xiao XK. Privacy enhanced matrix factorization for recommendation with local differential privacy. IEEE Trans. on Knowledge and Data Engineering, 2018, 30(9): 1770–1782. [doi: 10.1109/TKDE.2018.2805356]

[27] Zhang SJ, Yin HZ, Chen T, Huang Z, Cui LZ, Zhang XL. Graph embedding for recommendation against attribute inference attacks. In: Proc. of the 2021 Web Conf. Ljubljana: ACM, 2021. 3002–3014. [doi: 10.1145/3442381.3449813]

[28] Wang YX, Balle B, Kasiviswanathan SP. Subsampled Rényi differential privacy and analytical moments accountant. In: Proc. of the 22nd Int'l Conf. on Artificial Intelligence and Statistics. Naha: AISTATS, 2019. 1226–1235.

[29] Balle B, Barthe G, Gaboardi M. Privacy amplification by subsampling: Tight analyses via couplings and divergences. In: Proc. of the 32nd Int'l Conf. on Neural Information Processing Systems. Montréal: Curran Associates Inc., 2018. 6280–6290.

[30] Mironov I, Talwar K, Zhang L. Rényi differential privacy of the sampled Gaussian mechanism. arXiv:1908.10530, 2019.

[31] 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.

[32] Kingma DP, Ba JL. Adam: A method for stochastic optimization. In: Proc. of the 3rd Int'l Conf. on Learning Representations. San Diego: ICLR, 2015.

[33] 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]

[34] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: Proc. of the 13th Int'l Conf. on Artificial Intelligence and Statistics. Sardinia: AISTATS, 2010. 249–256.

附中文参考文献:

[1] 葛尧, 陈松灿. 面向推荐系统的图卷积网络. 软件学报, 2020, 31(4): 1101–1112. http://www.jos.org.cn/1000-9825/5928.htm [doi: 10.13328/j.cnki.jos.005928]

[9] 刘宇涵, 陈红, 刘艺璇, 赵丹, 李翠平. 图数据上的隐私攻击与防御技术. 计算机学报, 2022, 45(4): 702–734. [doi: 10.11897/SP.J.1016.2022.00702]

[16] 鲜征征, 李启良, 黄晓宇, 吕威, 陆寄远. 基于差分隐私和 SVD++的协同过滤算法. 控制与决策, 2019, 34(1): 43–54. [doi: 10.13195/j.kzyjc.2017.0961]

[20] 吴振强, 胡静, 田堉攀, 史武超, 颜军. 社交网络下的不确定图隐私保护算法. 软件学报, 2019, 30(4): 1106–1120. [doi: 10.13328/j.cnki.jos.005368]

王锟(1983－), 男, 博士生, 高级工程师, 主要研究领域为推荐系统, 数据隐私保护.

刘金源(1985－), 男, 博士, 讲师, 主要研究领域为网络管控, 网络攻击与防御, 数据隐私保护.

王永(1977－), 男, 博士, 教授, 博士生导师, CCF 高级会员, 主要研究领域为数据隐私保护, 推荐系统, 信息管理.

邓江洲(1993－), 男, 博士, 讲师, 主要研究领域为推荐系统, 决策优化.