软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software, 2024,35(8):3698−3720 [doi: 10.13328/j.cnki.jos.007124]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于负数据库的隐私保护图神经网络推荐系统 *

赵冬冬[1,2], 徐 虎[1], 彭思芸[1], 周俊伟[1]

[1](武汉理工大学 计算机与人工智能学院, 湖北 武汉 430070)
[2](武汉理工大学 重庆研究院, 重庆 401135)
通信作者: 周俊伟, E-mail: junweizhou@msn.com

摘 要: 图数据是一种特殊的数据形式, 由节点和边组成. 在这种数据中, 实体被建模为节点, 节点之间可能存在边, 表示实体之间的关系. 通过分析和挖掘这些数据, 人们可以获得很多有价值的信息. 因此, 对于图中各个节点来说, 它也带来了隐私信息泄露的风险. 为了解决这个问题, 提出了一种基于负数据库(NDB)的图数据发布方法. 该方法将图数据的结构特征转换为负数据库的编码形式, 基于此, 设计出一种扰动图(NDB-Graph)的生成方法. 由于 NDB 是一种保护隐私的技术, 不显式存储原始数据且难以逆转, 故发布的图数据能确保原始图数据的安全. 此外, 由于图神经网络在图数据中关系特征处理方面的高效性, 被广泛应用于对图数据的各种任务处理建模, 例如推荐系统, 还提出了一种基于 NDB 技术的图神经网络的推荐系统来保护每个用户的图数据隐私. 基于 Karate 和 Facebook 数据集上的实验表明, 与 PBCN 发布方法相比, 所提方法在大多数情况下表现更优秀. 例如: 在 Facebook 数据集上, 度分布最小的 $L1$ 误差仅为 6, 比同隐私等级下的 PBCN 方法低约 2.6%; 最坏情况约为 1 400, 比同隐私等级下的 PBCN 方法低约 46.5%. 在基于 LightGCN 的协同过滤实验中也表明, 所提出的隐私保护方法具有较高的精度.

关键词: 图数据; 隐私保护; 负数据库; 推荐系统; 图神经网络
中图法分类号: TP311

中文引用格式: 赵冬冬, 徐虎, 彭思芸, 周俊伟. 基于负数据库的隐私保护图神经网络推荐系统. 软件学报, 2024, 35(8): 3698–3720. http://www.jos.org.cn/1000-9825/7124.htm

英文引用格式: Zhao DD, Xu H, Peng SY, Zhou JW. Privacy-preserving Graph Neural Network Recommendation System Based on Negative Database. Ruan Jian Xue Bao/Journal of Software, 2024, 35(8): 3698−3720 (in Chinese). http://www.jos.org.cn/1000-9825/7124.htm

Privacy-preserving Graph Neural Network Recommendation System Based on Negative Database

ZHAO Dong-Dong[1,2], XU Hu[1], PENG Si-Yun[1], ZHOU Jun-Wei[1]

[1](School of Computer Science and Artificial Intelligence, Wuhan University of Technology, Wuhan 430070, China)
[2](Chongqing Research Institute, Wuhan University of Technology, Chongqing 401135, China)

Abstract: Graph data is a kind of data composed of nodes and edges, which models the entities as the nodes, nodes may be connected by edges, and edge indicates a relationship between entities. By analyzing and mining these data, a lot of valuable information can be revealed. Meanwhile, it also brings risks of privacy information disclosure for every entity in the graph. To address this issue, a graph data publishing method is proposed based on the negative database (NDB). This method transforms the structural characteristics of the graph

* 基金项目: 国家自然科学基金(61806151); 湖北省重点研发计划(2022BAA050); 海南省重点研发计划(ZDYF2021GXJS014); 重庆市自然科学基金(cstc2021jcyj-msxmX0002)
本文由“系统与网络软件安全”专题特约编辑向剑文教授、陈厅教授、王浩宇教授、罗夏朴教授、杨珉教授推荐.
收稿时间: 2023-09-11; 修改时间: 2023-10-30, 2023-12-15; 采用时间: 2023-12-22; jos 在线出版时间: 2024-01-05
CNKI 网络首发时间: 2024-04-03

data into the encoding format of a negative database. Based on this, a generation method for perturbed graphs (NDB-Graph) is designed. Since NDB is a privacy-preserving technique that does not explicitly store the original data and is difficult to reverse, the published graph data ensures the security of the original graph data. Besides, due to the high efficiency of graph neural network in relation feature processing in graph data, it is widely used in various task processing modeling on graph data, such as recommendation system. a graph neural network recommendation system is also proposed based on NDB technology to protect the privacy of graph data for each user. Compared with publishing method PBCN, the proposed method outperforms it in most cases in experiments on the Karate and Facebook datasets. For example, on Facebook datasets, the smallest $L1$-error of degree distribution is only 6, which is about 2.6% lower than the PBCN method under the same privacy level, the worst case is about 1 400, which is about 46.5% lower than the PBCN method under the same privacy level. The experiment of collaborative filtering based on LightGCN also demonstrates that the proposed privacy protection method has high precision.

Key words: graph data; privacy preservation; negative database; recommendation system; graph neural network

近年来, 随着网络的快速发展, 各种形式的数据研究和应用变得广泛. Facebook、微博、Twitter 等社交应用中积累了大量的图数据, 这些数据构成了社交网络. 此类数据由节点和边组成, 将实体建模为图中的节点, 将实体之间的关系建模为连接节点的边, 定义为 $G=(V,E)$, 其中, V 为节点的集合, E 是关系边的集合. 从基本的图算法到这些图的数据挖掘, 人们可以获得很多有用的信息, 并将其应用到各种场景中, 社交网络服务提供商可能会将积累的社交网络数据发布给企业、研究人员等第三方. 例如: 通过图数据进行图聚类、链接预测、节点分类等, 他们可以研究社会结构, 促进营销活动或实现更准确的广告活动. 除了基本的图算法之外, 图神经网络(GNN)[1]的开发和应用也非常热门. 近年来, 由于 GNN 获取图结构特征的强大能力, 它们也被用于处理图数据的许多任务和应用中, 推荐系统就是这些任务应用之一. 同时, 这些数据的使用不可避免地会涉及到用户的隐私以及其中包含的敏感信息. 例如: 一些不想公开的敏感个人信息或社会关系信息均可能会被泄露, 即使节点是匿名的, 但具有一定背景知识的外部人员也可以重新识别目标节点. 因此, 保护图数据及相关应用的隐私是非常有必要的.

在发布数据时, 一些发布者倾向于将数据匿名化, 并用随机数而不是标签来发布. 然而, 这种简单的匿名化远远不足以保护隐私, 它无法保护图数据的结构隐私. 传统的图数据隐私保护方法大致可分为基于图编辑的匿名化方法和基于差分隐私模型的图扰动方法[2,3]. 基于图编辑的方法可以分为随机图编辑、概率图编辑、k-匿名和通过聚类的泛化方法[2]. k-匿名是指通过某些方法编辑节点和边来发布数据, 使每条记录与其他 k–1 条记录具有相同的标识符属性值, 以减少链接攻击造成的隐私泄露. 但它也会改变原始数据的结构和特征, 降低数据的可用性. 差分隐私[4]属于扰动方法的一种, 它是 Dwork 提出的一种隐私保护方法, 具有严格的统计模型, 可以保证包含某人记录的数据集与不包含该人记录的数据集几乎相同. 图数据中的差分隐私一般分为边差分隐私和点差分隐私[5]. 然而, 差分隐私在图数据中的应用面临着边或节点对噪声过于敏感的问题, 主要体现在图结构过于复杂而无法满足差分隐私. 因此, 如何有效平衡差分隐私提供的安全性和数据效用, 是该技术在图数据中应用的一大挑战.

负数据库(NDB)[6,7]是受人工免疫系统负选择机制启发的一种新的数据表示形式和新的隐私保护技术, 它由 Esponda 等人首先提出, 具有无需解密即可进行计算和数据库操作的优点. 同时, 每个 NDB 都可以转换为 SAT 实例, 解决 SAT 问题是 NP 困难的, 因此, NDB 是很难被逆转的. 负数据库技术的特性使得负数据库非常适合数据的隐私保护, 也广泛应用于密码认证[8]、生物信息识别[9]、数据发布[10]、数据挖掘[11]等领域. 为了解决图结构数据的现有发布方法中的步骤复杂性、发布数据的低效用性以及隐私保护效果具有局限性的问题, 本文将该技术应用到图数据的隐私保护中, 并设计了相应的编码方案, 这不同于传统的图结构数据发布中使用的匿名和扰动方法. 该算法可用于估算图结构数据的一些基本属性特征. 本文还提出了一种基于负数据库技术的用户可自定义隐私需求的图神经网络推荐系统: 首先, 提出了一种适用于将每个用户 user-item 交互图的转化为负数据库的编码方式; 其次, 提出了一种用户可自定义隐私保护程度的联邦学习形式的推荐系统模型, 该模型中, 服务器端将来自不同用户的负数据库 *item* 序列数据推算出扰动交互图, 并聚合为总的扰动图进行学习, 然后把得到的对应模型参数返回用户端计算出最终推荐结果.

© 中国科学院软件研究所 http://www.jos.org.cn

本文的贡献可以概括如下:

(1) 与传统图数据发布中使用的匿名和扰动方法不同, 本文提出了一种负表示方法, 该方法使用合适的编码方法将图数据转换为负数据库, 以保护图数据的隐私;

(2) 提出了一种基于负数据库的扰动图 NDB-Graph 的生成方法, 该方法可用于估计图数据的属性, 例如度分布、三角形数、聚类系数和平均最短路径长度;

(3) 提出了一种合适的编码方法, 每个用户可自定义隐私需求来将自身的用户和项目(user-item)交互图转换为 NDB 以及基于 NDB 的 GNN 推荐系统的隐私保护框架模型.

本文第 1 节介绍图数据隐私保护的相关工作. 第 2 节介绍负数据库技术和负数据库生成算法的背景知识. 第 3 节提出基于 NDB 的图神经网络推荐系统, 并讨论基于该方法的安全性. 第 4 节展示实验结果. 第 5 节给出总结.

1 相关工作

1.1 保护隐私的图数据发布

与表格形式的数据不同, 由于节点的连通性[12], 为表格数据提出的隐私保护方法不能直接应用于图数据. 除了单元数据的准标识符之外, 攻击者还可以通过图数据中的边、子图或邻居信息来重新识别它们. 因此, 如果节点只是简单地匿名化, 仍然存在安全风险. k-匿名可以有效防止外部链接攻击, 应用于图数据的 k-匿名模型通常会确定地删除或添加图上的节点和边以提供匿名性, 例如 k-度匿名[13]构造图中的每个节点至少与其他 $k-1$ 节点具有相同的度数, 并且 k-同构匿名[14]使得发布的图包含 k 对相同构的子图等. 基于敏感属性的实际隐私保护需求, Ren 等人[15]提出了社交网络的个性化(α,β,l,k)-匿名模型.

基于聚类和泛化的图发布方法将图抽象为超节点和超边以满足 k 匿名性. 与 k 匿名和随机化方法不同, 它们不修改图结构. 例如, Campan 等人[16]引入了边泛化方法并量化了结构信息损失的度量方式, 他们提出了一种 SaNGreeA 算法, 通过最小化泛化信息和结构信息的损失来构建 k-匿名聚类网络. Tassa 等人[17]提出一种基于序列聚类的社交网络匿名算法, 通过聚类生成匿名信息, 比 SaNGreeA 算法具有更好的效用. Jiang 等人[18]基于结构相似性和属性相似性, 将图中的节点和边聚类成至少包含 k 个节点的超级节点, 他们隐藏超节点中每个节点的属性值, 并只发布其泛化的属性值. Zhou 等人[19]则是基于网络节点特征综合距离来划分超级节点. 但由于这类方法只发布具有超点和超边的聚类图, 不可避免地很大程度上降低了发布数据的效用.

近年来, 有许多研究人员将差分隐私应用于图数据: Ahmed 等人[20]采用随机投影方法, 将图的邻接矩阵的每一行投影到低维空间, 然后向矩阵中添加随机噪声来满足差分隐私, 他们通过添加低噪声来保持数据的效用; Mülle 等人[21]提出了适用于图聚类的相邻图的定义, 并基于该定义提出了一种通过边概率化保证单边差分隐私的隐私集成图聚类方法(PIG)算法; Liu 等人[22]提出了一种基于社区划分和局部差分隐私相结合的社交网络数据匿名发布模型 DP-LUSN, 他们的方法不仅确保了强隐私性, 还有效保留了网络的结构特征; Chen 等人[23]引入了一个附加参数来衡量图数据中可能的内部相关性, 他们自适应地识别邻接矩阵的密集区域, 然后用指数机制重建邻接矩阵以进行发布; Ma 等人[24]将场景指定为社区检测, 他们在 Louvain 算法的模块化中添加噪声, 将社区转化为 HRG 模型, 在其边缘连接概率中添加拉普拉斯噪声, 以保护社区的隐私; Huang 等人[25]提出了一种基于聚类和添加噪声的 PBCN 方法.

基于图编辑的匿名化方法和基于差分隐私模型的图扰动方法都存在一定的局限性和缺点. 例如: 随机图编辑方法[26]忽略了图数据集中所有节点应有的平等安全性, 只为随机节点提供安全性; 匿名技术针对不同的需求有不同的匿名方式; 基于超节点和超边的聚类和泛化的方法大大降低了原有网络的效用; 应用差分隐私于图数据也面临着如何平衡隐私预算和数据效用的挑战.

表 1 是相关符号及释义.

© 中国科学院软件研究所 http://www.jos.org.cn

表 1 相关符号及释义

符号	释义
V	原始图数据中的点集
E	原始图数据中的边集
$\bar{V}$	扰动图数据中的点集
$\bar{E}$	扰动图数据中的边集
K	NDB_s 中每个记录中的确定位个数
P_i	选择生成第 i 种类型的 NDB_s 的概率参数
q_i	选择每个属性中第 i 位的概率参数
r	决定 NDB_s 记录数量的参数
s	图结构数据生成负数据库前的对应编码串
$V_{\max}$	数据集中最大节点序号
L	数据集中节点数目转换为二进制形式的最大位数
P_{diff_i}	每个属性中的第 i 位不同于原始串相应位的概率
P_{same_i}	每个属性中的第 i 位同于原始串相应位的概率
$\Pr(Node_w)$	第 i 个节点等于扰动节点时的最大概率
P_{equal}	推算出的扰动图与原始图相等的概率
$e_u^{(k)}$	神经网络中第 k 层的用户嵌入
$e_i^{(k)}$	神经网络中第 k 层的项目嵌入
$\bar{a}_u$	第 u 个 user 对应 *item* 序列扰动向量
$\bar{N}_u$	节点 u 的扰动邻居节点集

1.2 基于图神经网络的隐私保护推荐系统

推荐系统可以帮助用户从复杂信息中提取他们可能感兴趣的内容, 其中, user-item 之间的交互信息可以建模为图数据. 图神经网络具有处理关系和结构化数据的优势[27,28], 由于其处理图数据中的连接关系的高效性, 近年来被广泛应用于图数据的各种建模任务中. 基于 GNN 的方法在推荐系统中也变得新颖且流行. 根据图神经网络的类型, 目前的相关应用大致可以分为基于图卷积网络、图注意力网络、门控图神经网络以及其他图神经网络架构的推荐系统模型. 其中, 基于 GCN 的模型使用嵌入传播来迭代聚合邻域嵌入, 通过堆叠传播层, 每个节点可以访问更高阶邻居的嵌入并为推荐提供更多特征信息.

图神经网络推荐系统中的隐私保护方法通常可以分为 3 类: 联邦学习框架、统计方法和加密方法[29]. 联邦学习是一种分布式学习框架, 通过在本地设备上训练模型并结合各种隐私保护[30]机制来增强隐私保护. 它可以通过在每一轮训练中从一组用户中进行抽样、聚合并将更新的模型传播给其他用户, 从而降低集中式学习模型中用户隐私泄露的风险. 统计方法通过移除特征、混淆、添加差分隐私噪声等方式来隐藏用户行为隐私, 这种方法在保护了数据隐私的同时, 也不可避免地降低了推荐系统的准确性. 在加密方法中, 同态加密可以在没有任何隐私开销的情况下进行计算, 但代价是消耗大量的计算资源.

在此领域近年来的研究工作中, Qiu 等人[31]提出了一种基于差分隐私的联邦图神经网络模型 DP-FedRec 来保护隐私, 解决了联邦系统中数据非独立但同分布的问题. Liu 等人[32]设计了一个名叫 FeSoG 的框架, 通过注意力机制和聚合处理异质性以保持个性化. 他们使用本地数据来推断用户嵌入, 提出的模型将伪标签技术与项目采样相结合以保护隐私并增强训练. Liu 等人[33]提出了一种在用户嵌入模型参数传输到服务器时,采用同态加密的隐私保护神经图协同过滤模型 FTL-NGCF. Shin 等人[34]基于局部差分隐私开发了一种新的矩阵分解算法, 以确保 user-item 交互图和评分的隐私性. 各个用户将其数据随机化以满足差分隐私, 并将处理后的数据发送到服务器. 最后, 推荐模型在扰动数据的集合上进行计算. Fang 等人在文献[35]中提出了一种基于可变自动编码器(VAE)的差分隐私推荐系统, 通过根据用户元数据计算用户级别的优先级来优化 VAE 模型, 并在计算过程中添加噪声以保护用户数据. Gao 等人[36]针对协同过滤提出了一个通用的差分隐私局部协同过滤推荐框架, 解决了数据收集和推荐过程中获得最终推荐结果的隐私问题. 此外, Wu 等人[37]提出了一种联邦框架 FedPerGNN, 该框架通过局部差分隐私(LDP)和伪交互项目采样方法来保护 use-item 交互信息, 同时利用了不同用户的高阶信息来缓解信息隔离问题, 可以较好地保护用户隐私.

© 中国科学院软件研究所 http://www.jos.org.cn

2 基础知识

2.1 负数据库

无论是基于图编辑的匿名化方法和基于差分隐私模型的图扰动方法, 都存在着一定的局限和不足. 例如: 随机图编辑忽略了图数据集中所有数据应有平等的安全性的问题, 它只为随机的用户提供安全性; 匿名化技术针对不同的数据, 有不同的匿名化方法, 简单的匿名化方法可能不足以提供足够的隐私保护, 复杂的匿名化方法又可能导致步骤过于繁琐; 基于聚类和泛化的方法只发布具有超点和超边的聚类图, 大大降低了原始网络的效用; 在图数据中应用的差分隐私方法也面临着边或点对噪音过于敏感的问题, 主要体现在图的结构形式复杂、隐私预算与数据效用难以权衡.

负数据库[7]作为一种新型的数据表示和一种受人工免疫系统负面选择机制启发的隐私保护新技术, 它与传统表示方法最大的区别在于, 它总是存储与实际信息不相符的信息, 因此很适合应用于信息隐藏. 在现有研究中, 负数据库基本都是基于二进制串组成的. 负数据库具有在不需要解密的情况下能够进行计算和数据库操作的优点, 例如传统数据库上的一些操作例如选择、插入、更新、删除等, 也可以用于负数据库. 每个负数据库都可以被转换成一个 SAT 实例, 求解 SAT 问题被证明是 NP 困难的, 因此使得负数据库就很难被逆转. Esponda 等人于也证明了, 对负数据库求逆恢复出原始数据是 NP 困难问题. 因此, 难以逆转的特点使得负数据库能用来对数据进行隐私保护.

负数据库的具体定义如下: 假设 U 是一个包含 m 位二进制字符串的全集, DB 是包含一些 m 位二进制字符串的传统数据库, 则 U-DB 中的信息就是 NDB. 由于 U-DB 的大小往往很大, 为了节省 NDB 的存储空间, 一般使用通配符*来表示可以是 0 或 1 的位. NDB 由*, 0 和 1 组成. 表 2 显示了 NDB 的一个简单示例. 关于负数据库生成算法已经有很多文献研究, 一些算法可以生成难以逆向的负数据库, 如 q-hidden 算法[38]、K-hidden 算法[39]以及更细粒度的负数据库生成算法 QK-hidden 算法[11].

表 2 一个简单的 NDB 实例

DB	U-DB	NDB
001, 010	000, 011, 100, 101, 110, 111	011, 1**, 000

2.2 *QK*-hidden算法

QK-hidden 算法可以在负数据库生成时, 对不同类型记录的概率更自由地控制, 具体表现为: 该算法首先以概率参数 $p_1,\ldots,p_K$ 来选择生成第 i 种类别(记录有 i 位与隐藏串不同)的负数据库记录, 然后根据概率参数 $q_1,\ldots,q_L$ 更加细粒度地决定每个属性位上第 i 位的是否取反(L 是属性的最大长度). 并按 $1/L$ 的概率将其他 $K-i$ 位设置为与 s 相同, 再将记录的其他 $m-K$ 位全部置为*, 最后将生成的记录加入 NDB_s 中. 不断重复上述步骤, 直到 NDB_s 中串的数量等于 $m\times r$. 下面的算法 1 是 QK-hidden 算法的具体步骤, 其中, 参数 K 决定了每条 NDB 记录有 K 个确定位; 参数 r 决定了生成的负数据的大小, 即等于数据串长 $m\times r$.

参数设置需要满足 $\sum_{i=1}^{K}(K-2i)\times p_i>0$, 以保证生成的 NDB 很难被局部搜索策略逆转[40].

算法 1. QK-hidden 算法.

Input: 一个 m 位的二进制串 s, 属性的最大位数 L, 常数 r, K 以及概率参数$\{p_1,\ldots,p_k\}$, $\{q_1,\ldots,q_L\}$;

Output: NDB_s.

1) $NDB_s\leftarrow\varnothing$;
2) $P_0\leftarrow 0, P_1\leftarrow p_1,\ldots,P_i\leftarrow p_1+\ldots+p_i$, 其中, $0\leqslant i\leqslant K$;
3) $N=m\times r$;
4) **While** ($|NDB_s|<N$) **do**:
5) $rnd\leftarrow random(0,1)$;
6) 得到满足 $P_{i-1}\leqslant rnd\leqslant P_i$ 的 i

© 中国科学院软件研究所 http://www.jos.org.cn

7) 初始化一个 m-bit 的串 v;
8) **While** (i>0) **do**:
9) 得到满足 $q_1+\ldots+q_{a-1}\leqslant random(0,1)\leqslant q_1,\ldots+q_a$ 的 a;
10) 随机选择 v 的一个属性位上未确定的第 a 位设置为与原始串 s 相反;
//若选中位上已确定则重新选择
11) 随机选择 v 的其他 $K-i$ 位设置为与 s 相同, 再将 v 的其他 $m-K$ 位设置为*;
12) $i\leftarrow i-1$;
13) $NDB_s\leftarrow NDB_s\cup v$;
14) **Return** NDB_s;

3 基于负数据库的隐私保护图神经网络相关方法

本文主要的研究将围绕图结构数据发布方法以及图神经网络的推荐系统中的隐私保护问题展开, 研究目标是一种基于负数据库的图结构的数据隐私保护发布方法以及应用到图神经网络推荐系统中的隐私保护方法. 本节将通过以下两个部分展开.

首先, 为了解决图结构数据的现有发布方法中的步骤复杂性、发布数据的低效用性以及隐私保护效果局限性的问题, 本节提出了一种基于负数据库的图结构数据隐私保护发布方法. 在图结构数据的隐私保护中, 如何有效地平衡数据的安全性和可用性是一大挑战. 为了解决这个问题, 本节针对图结构数据的发布问题提出了一种负表示图数据发布方案. 该方案包括一种适用于负数据库生成算法的图数据编码方法, 通过组合所有边的节点对来表示图形数据集. 该方法将图结构数据集转换为 m 位二进制隐藏字符串, 不仅节省了空间存储, 还不需要额外的度数信息. 同时, 提出了一种从 NDB_s 中计算扰动图 NDB_{Graph} 的方法 NDB-Graph, 由原始图数据生成的 NDB_s, 经过本文提出的 NDB-Graph 算法来生成扰动图 $NDB_{Graph}(\overline{V},\overline{E})$. 以及度分布、三角形数、平均聚类系数、平均最短路径长度等特征属性的估算方法, 相比起其他方法, 具有易实施、数据可用性较高的特点.

其次, 针对基于图神经网络的推荐系统中用户数据获取阶段中用户隐私泄露风险、模型计算阶段中不可信的数据计算处理过程中的隐私泄露风险和最终推荐结果生成阶段的用户喜好趋势的隐私泄露风险, 本文提出了一种基于负数据库的自定义隐私需求联邦推荐系统框架. 基于此, 本文还提出了自定义隐私需求负数据库参数生成算法和神经网络中的扰动交互图聚集算法. 具体来说, 该系统中用户通过自身隐私保护需求参数 sec_{need} 对应生成不同 q 参数下的 NDB_s 后上传给服务器端; 然后, 服务器端每收到一个用户的数据后计算其扰动图, 并将其聚合到总的 user-item 交互图中进行训练; 最后, 用户下载对应的 user 嵌入和所有 item 嵌入计算最终推荐结果.

3.1 负表示的图数据编码方法

负数据库生成的原始数据需要用二进制表示. 考虑到将图结构数据集转换为 m 位二进制隐藏字符串, 可对于具有 v 个节点和 e 条边的图的邻接矩阵, 由于它是一个仅由 0 和 1 组成的矩阵, 所以矩阵的每一行可以被直接连接以形成隐藏字符串 s. 然而, 图的邻接矩阵的空间复杂度是 $O(v^2)$. 例如: 在 1 000 个节点的情况下, 图形的表示需要 10^6 位. 即使无向图只需存储一半的数据, 表示它们仍然需要很大的空间. 此外, 真实的图数据集通常非常稀疏, 如果使用邻接矩阵来表示稀疏图, 矩阵中大多数位置为 0, 这将会导致很大的空间浪费. 对于邻接表, 若用二进制表示其所有边将占用 $2e\times L$ (无向图)或 $e\times L$ (有向图)的存储成本, 其中, L 为节点最大序号的二进制位数. 邻接表的空间复杂度为 $O(v+e)$, 当图比较稠密时, 仍然会导致较大的存储开销. 而且, 该方案在计算负数据库数据时每个节点的边数不确定, 在后期扰动图生成计算中需要额外的度数信息, 这将增加额外的隐私信息(度数信息)披露. 节点对编码方法在图数据集表示中具有显著优势, 其主要优势体现在空间复杂度、度数信息和计算效率方面: 首先, 该方法以高度紧凑的方式存储图的连接信息, 避免了邻接矩阵和

© 中国科学院软件研究所 http://www.jos.org.cn

邻接表中的空间浪费, 使得其空间复杂度线性于边的数量, 适用于大型或稀疏图; 其次, 节点对编码方法无需存储节点度数信息, 从而降低了隐私信息泄露的风险, 尤其在隐私敏感的应用中具备优势; 此外, 它在高效查询操作方面表现出色, 适用于不需要节点度数信息的图算法. 总之, 节点对编码方法是一种通用、高效、隐私友好的图数据表示方式, 特别适用于大规模、稀疏或需保护隐私的图数据集. 因此, 最好的选择是通过组合所有边的节点对来表示图形数据集.

算法 2 是一个负表示的图数据编码算法, 其具体步骤如下.

step 1: 将数据集的所有边序列都连接在一起, 这些边序列由两个连接节点的序号组成;

step 2: 将每个节点都转换为二进制, 为了实现这一点, 需要确定最大位数 L, 即在所有节点中, 取所有节点中二进制位数的最大值;

step 3: 将每个节点的序号转换为二进制, 如果某个节点二进制位数小于最大位数, 则在前面填充 0, 直到位数等于最大位数;

step 4: 得到表示整个图形的隐藏字符串 s, 这是用于最终输入负数据库生成算法的图结构数据编码.

算法 2. 负表示的图数据编码算法.

Input: 已匿名处理节点的图 *graph*, V 表示图中的节点,节点最大位数 L;

Output: 表示图形的隐藏字符串 s.

1) s=*null*;
2) 将图中所有由边连接的节点组合成节点对 $G(V,V')$;
3) 根据最大节点值 $\max_v$ 计算出节点最大位数 L;
4) **For** 节点对 $G(V,V')$:
5) 计算 V, V'的二进制串 bin_v, $bin_{v'}$;
6) **if** ($len(bin_v)<L$) bin_v 前补($L-(len(bin_v))$)个 0;
7) **if** ($len(bin_{v'})<L$) $bin_{v'}$前补($L-(len(bin_{v'}))$)个 0;
8) $s+=bin_v$;
9) $s+=bin_{v'}$;
10) **Return** s;

现以一个简略图为例, 图 1 左侧是一个具有 5 个节点的图, 每个节点的最大位数 L 为 3 ($2^2<5<2^3$). 最终, 该图将被编码为图 1 右侧的字符串 s, 字符串 s 中每 3 位(每 L 位)对应于 *QK*-hidden 算法中隐藏串的属性位, 这里称之为节点位. 通过这个算法, 我们可以将图形数据以一种紧凑而有效的方式编码, 以便进一步的处理和分析.

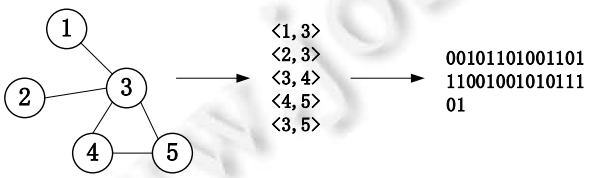

图 1 一个基于负数据库的图编码实例

3.2 扰动图生成算法NDB-Graph

本文所提出的方法主要针对图结构信息, 即节点和边的隐私保护. 由于现实生活中许多数据集中往往包含大量节点, 每个节点序号的最大位数 L 很大, 因此通过使用第 2 节中介绍的细粒度 *QK*-hidden 负数据库生成算法来更精准地通过负数据库生成出扰动图, 以准确地计算每个节点.

当用户发布了 *QK*-NDB 形式的图结构数据后, 便可以通过公式(1)[11]从中计算出每个节点位的第 i 位与编码字符串的对应位不同的概率 P_{diff_i}, 以及对应位相同的概率为 $P_{same_i}=1-P_{diff_i}$, P_{diff_i} 代表属性的第 i 位与隐藏串对应位不同的概率, P_{same_i} 代表属性的第 i 位与隐藏串对应位相同的概率. 编码字符串 s 的第 j 位为 0 的概率

© 中国科学院软件研究所 http://www.jos.org.cn

$\Pr(s_j=0)$可以通过公式(2)[11]计算, 其中, $n0_j$是第 j 位上为“0”的总记录数, $n1_j$是第 j 位上为“1”的总记录数, 第 j 位为 1 的概率相应地等于 $1-\Pr(s_j=0)$:

$$P_{diff_i}=\frac{\sum_{a=1}^{K} j\times p_a\times q_i}{\sum_{a=1}^{K} a\times p_a\times q_i+\sum_{a=1}^{K}(K-a)\times p_a\times \frac{1}{L}} \tag{1}$$

$$\Pr(s_j=0)=\frac{(P_{diff_i})^{n1_j}\times(P_{same_i})^{n0_j}}{(P_{diff_i})^{n1_j}\times(P_{same_i})^{n0_j}+(P_{diff_i})^{n0_j}\times(P_{same_i})^{n1_j}} \tag{2}$$

编码后的图结构数据字符串 s 通过 QK-hidden 算法生成 NDB_s 后, 首先计算出 NDB_s 中每个位上的“0”和“1”的数量, 再根据公式(1)和公式(2)计算每个位为 0 或 1 的概率后, 对于字符串 s 的第 n 个节点位, 其为节点 x 的概率 $\Pr(s_n^{\text{int}}=x)$ (第 n 个节点位的十进制值为 x)就可以通过公式(3)计算出, 其中, x_i^{bin} 表示二进制形式的 x 的第 i 位的值:

$$\Pr(s_n^{\text{int}}=x)=\prod \Pr(x_i^{bin}) \tag{3}$$

具体来说, 例如, $\Pr(s_1^{\text{int}}=2)=\Pr(\text{第1位为0})\times\Pr(\text{第2位为1})\times\Pr(\text{第3位为0})$ 是图 1 中第 1 个节点位为节点“2”时的概率. 每个节点位相当于分别连接每个边中的节点的序列号, 因此总的节点位个数为边数×2. 选择根据 NDB_s 计算的每个节点位上的最大概率值的十进制值 v 作为扰动图 NDB_{Graph} 的节点. 每条边节点的具体计算方法可以表示为公式(4):

$$Node_w=\{v\mid \forall x:\Pr(s_w^{\text{int}}=x)\leqslant \Pr(s_w^{\text{int}}=v)\} \tag{4}$$

其中, 第奇数位为边的开始节点, 它的下一个偶数位为边的结束节点. 综上所述, 边可以表示为

$$\bar{E}=\{(Node_w,Node_{w+1})\mid w=1,3,\ldots,2|\bar{E}|-1\} \tag{5}$$

基于负数据库生成扰动图 $NDB_{Graph}(\bar{V},\bar{E})$ 的具体算法见算法 3, 其中, s_w^{int} 表示编码后字符串 s 的第 w 个节点位的十进制数.

算法 3. NDB-Graph 算法.

Input: 由原始图数据生成的 NDB_s, 最大节点序号 $v_{\max}$, 节点位的最大长度 L;

Output: 扰动图 $NDB_{Graph}(\bar{V},\bar{E})$.

1) $m\leftarrow\frac{N}{r}$, $n\leftarrow\frac{m}{L}$, $w\leftarrow 1$;
2) 计算出 NDB_s 中的每一位为 0 和 1 的数量 $n0_j$, $n1_j$ (j=1,…,m);
3) P_{diff_i}, $P_{same_i}=1-P_{diff_i}\ (i=1,\ldots,L)$;
4) **While** ($w\leqslant n$) **do**:
5) $\Pr(s_w^{\text{int}}=x),\ x=0,\ldots,v_{\max}$; //计算每个节点位为所有节点的概率
6) $Node_w\leftarrow\arg\max(\Pr(s_w^{\text{int}}=x))$;
7) **If** $Node_w\notin\bar{V}$:
8) $\bar{V}\leftarrow\bar{V}\cup Node_w$;
9) **If** $\left(\frac{w}{2}\right)=0$ and $(Node_{w-1},Node_w)\notin\bar{E}$:
10) $\bar{E}\leftarrow\bar{E}\cup(Node_{w-1},Node_w)$;
11) $w\leftarrow w+1$;
12) **Return** $NDB_{Graph}(\bar{V},\bar{E})$;

对于直接从 NDB_s 中估计图的一些基本特征属性, 以下称为扰动属性, 本节列举介绍了其中的一些计算方法. 对于一个节点 v, 如果每个节点位上有一个概率值最大的节点等于 v, 则节点 v 的度数加 1. 公式(6)描述了对一个节点的扰动度数计算, 其中, 当 $f(x)$为真时, $|f(x)|$等于 1; 否则等于 0. 通过遍历所有节点的最大概率

© 中国科学院软件研究所 http://www.jos.org.cn

值，可以得到总的扰动度序列 $\overline{Deg}=(\deg_1^{NDB_{Graph}},\ldots,\deg_{v_{\max}}^{NDB_{Graph}})$：

$$\deg_v^{NDB_{Graph}}=\sum|\arg\max_x(\Pr(s_w^{\text{int}}=x))=v|,\ w=1,2,\ldots,2|\bar{E}| \tag{6}$$

扰动三角形数可以通过节点的 1 跳邻居之间的边数来计算，因此在 NDB_s 中，它等于公式(7)，其中：u_i，u_j 满足扰动边 $\bar{E}(v,u_i\text{或}u_j)$ 或 $\bar{E}(u_i\text{或}u_j,v)$ 存在：

$$Triangle_v^{NDB_{Graph}}=\sum|\bar{E}(u_i,u_j)| \tag{7}$$

聚类系数用于描述图中节点及其相邻节点之间的聚集情况．在社交网络中，它可以理解为一个人的两个朋友之间的联系概率．对于一个节点，可以通过它的邻接节点数和度数来计算它．扰动聚类系数的具体计算过程由公式(8)表示：

$$\overline{CC}_v=\frac{2\sum|\bar{E}(u_i,u_j)|}{\deg_v^{NDB_{Graph}}\times(\deg_v^{NDB_{Graph}}-1)} \tag{8}$$

类似地，本文提出的方法也适用于计算基于 NDB 的扰动最短路径长度．这里以最短路径的广度优先搜索算法(BFS)为例，其具体步骤是输入原始图的负数据库 NDB_s，计算节点 $\bar{V}$ 和边 $\bar{E}$．从 $\bar{V}$ 中的每个节点开始，逐层遍历所有相邻节点，并将访问顺序与队列存储，找到目标节点后再回溯，从而找到最短路径．

3.3 基于NDB的图神经网络推荐系统

受图神经网络强大的图数据学习能力的启发，近年来涌现了大量基于图神经网络的推荐模型，如 NGCF, LightGCN, GC-MC, GraphSAGE 等．图神经网络从用户-项目交互图中捕获协作信息的方法主要包括以下几个步骤：构建图、聚合邻居、更新信息和节点表示．应用于个性化服务的推荐系统的隐私泄露风险主要来自 3 个部分[29,41]：用户数据获取阶段、模型计算阶段和最终推荐结果生成阶段．针对上述风险，本文提出了一种基于负数据库的联邦图神经网络推荐模型．

该模型主要可以分为 3 个部分：第 1 部分是为每个用户生成 user-item 交互图并生成对应的负数据库，在这部分，不同的用户使用自定义的隐私需求参数 sec_{need} 来生成相应的 NDB 数据并发送给服务器；第 2 部分聚合所有用户的扰动交互图，在这部分中，服务器接收来自不同用户的 NDB，并聚合总的扰动 user-item 交互图矩阵进行计算；第 3 部分是本地推荐，由于最终推荐结果生成阶段可能存在隐私泄露风险，因此这部分最终结果生成在用户侧本地进行．具体来说，服务器端并不计算最终的推荐结果，而是将相应的用户嵌入以及最后一层计算出的所有项目嵌入提供给用户．模型框架如图 2 所示．

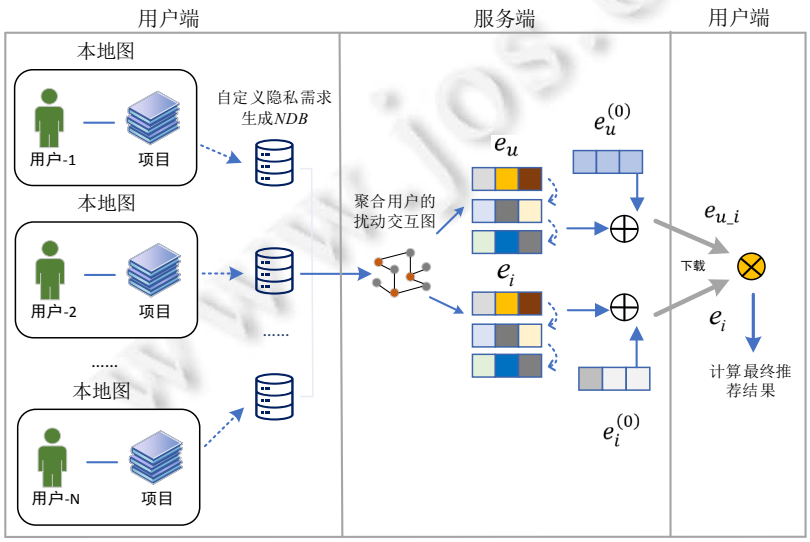

图 2 基于负数据库的自定义隐私需求推荐系统框架图

© 中国科学院软件研究所 http://www.jos.org.cn

3.4 自定义隐私需求的负数据库参数生成算法

user-item 交互可以被认为是二分图或两个同构图(即 user-user 图和 item-item 图). 在数据收集阶段, 不同于一般形式的图结构, 由每个用户发送个人的 item 序列给服务器, 因此预处理中, 仅需将该 item 序列转化为 NDB_s, 具体处理方法如图 3 所示: 首先得到最大 item 二进制位数 L; 然后对应地将每个用户的 item 序列转换为二进制形式, 若不足 L 位, 则补 0 直至 L 位, 其中, 每 L 位对应一个 item 节点位, 以下简称 item 位; 最后, 再通过算法 1 将此二进制串转换为负数据库.

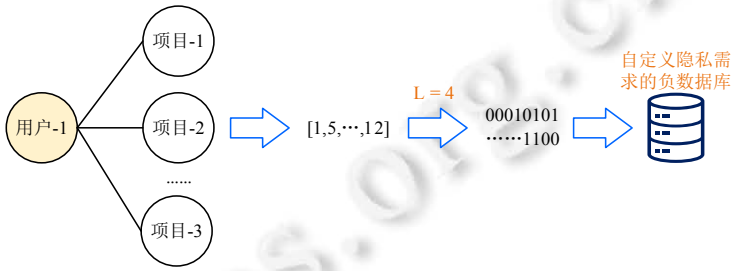

图 3 对于每个 user-item 交互图的预处理

由于每一个用户对于自己的数据可能有不同要求, 由文献[41]可知, QK-hidden 算法中每个参数 q_i 可以控制不同位的推算概率, 具体表现为 P_{diff} 越接近于 0.5, 该位的推算越不准确; 越远离 0.5, 该位的推算越准确. 因此, 每个用户的 item 序列转换为二进制串后, 本文提出使用一个隐私需求程度参数 $\sec_{need}\in(0,1)$(越接近于 0, 隐私保护程度越低; 反之越高)来控制 P_{diff} 值以及该用户生成 NDB_s 的参数 q_i, 具体表现为通过控制第 1 位 q 值来实现(在 item 序列形式的编码中, 第 1 位的权重最大). 根据公式(9)可以变形推导出 q_i 由 P_{diff_i} 计算出的公式如下:

$$q_i=\frac{P_{diff_i}\times\left[\frac{1}{L}\times\sum_{a=1}^{K-1}p_a\times(K-a)\right]}{(1-P_{diff_i})\times\sum_{a=1}^{K}a\times p_a} \tag{9}$$

其中, 隐私需求程度参数 $\sec_{need}$ 分别决定了第 1 位和最后 1 位参数 q, 第 2 位至第 $L-1$ 位(中间位固定为相同值)决定了剩余的总的隐私预算. 算法 4 展示了由隐私需求程度参数 $\sec_{need}$ 生成概率参数 q_1, q_L 的过程. 简单来说就是, 当隐私需求程度参数 $\sec_{need}$ 越靠近 1 时, 对应 P_{diff_1} 就越靠近 0.5. 用户根据生成的参数将原始数据转换为 NDB_s 后发送到服务器端, 由于每个用户的数据都是负数据库形式, 即使服务器端是不可信任的, 也在直观上保护了用户数据的隐私.

算法 4. QK-hidden 参数生成算法.

Input: 隐私需求程度参数 $\sec_{need}$, item 节点位的最大长度 L, K 以及概率参数 $p_1,\ldots,p_k$, $q_2,\ldots,q_{L-1}$ 位的值 h;

Output: 概率参数$\{q_1,q_L\}$.

1) 计算 $q_2,\ldots,q_{L-1}$ 总和最大值 $Q_{p2}=h\times(L-2)$以及 q_1 和 q_L 总和最大值 $Q_{p1}=1-h\times(L-2)$;

2) 通过公式(9)(以下简称为 $F(x)$)计算 $P_{diff}=0.5$ 时的 q'值;

3) **If** $x>q'$ or $Q_{p2}>1$:

4) **Return** false;

5) **Else**:

6) **If** $P_{diff}^{Q_{p1}}>0.5$:

7) $P_{diff_1}=0.5+|P_{diff}^{Q_{p1}}-0.5|\times(1-\sec_{need})$, $q_1=F(P_{diff_1})$, $q_L=Q_{p1}-q_1$;

8) **Else**:

9) $P_{diff_1}=0.5-|P_{diff}^{Q_{p1}}-0.5|\times(1-\sec_{need})$, $q_1=F(P_{diff_1})$, $q_L=Q_{p1}-q_1$;

© 中国科学院软件研究所 http://www.jos.org.cn

10) **Return** 概率参数$\{q_1,q_L\}$.

3.5 扰动交互图聚集算法

在服务器端, 用户每次上传自身 user-item 交互图转换后的 NDB_s 后, 服务器计算当前局部的扰动交互图, 并聚集这些局部扰动图加入到全局图中得到总的扰动图 $\overline{A}$.

算法 5 描述了服务器端扰动交互图的聚集过程. 首先, 推算出上一节中 user-item 编码串中每个 item 节点位上 1 到 item 数的所有概率值, 取最大概率值的 item 作为当前位的扰动 item; 其次, 对于每一个用户 u, 将其扰动 item 序列转换为邻接向量 $\overline{a}_u$ (形状为 1×*item* 总数). 具体来说, 每当得到一个 item 节点位的扰动 item 值, 对于该扰动 item 位值置为 1, 剩余所有位全部置为 0, 最后将该扰动邻接向量 $\overline{a}_u$ 加入总 user- item 交互图 $\overline{A}$ 邻接矩阵第 u 行. 当用户不是首次上传数据时, 将总 user-item 交互图 $\overline{A}$ 邻接矩阵第 u 行更新为新的扰动邻接向量 $\overline{a}_u$:

$$\overline{a}_u=\begin{cases}1, \text{ if } \arg\max x_x(\Pr(s_w^{\text{int}}=x))=item\ existed\\ 0, \text{ else}\end{cases} \tag{10}$$

$$\overline{A}=\begin{bmatrix}\overline{a}_1\\ \overline{a}_2\\ \cdots\\ \overline{a}_{u\in U}\end{bmatrix} \tag{11}$$

算法 5. 扰动交互图聚集算法.

Input: 用户发送的 NDB_s 数据以及负数据库生成相关参数;

Output: 扰动 user-ite*m* 交互图 $\overline{A}$.

1) (Server 端)初始化 $\overline{A}$ 为一个形状为 *user* 总数×*item* 总数的全零矩阵
2) (Server 端) **For** 每个参与者 u:
3) 估算每个用户 u 的数据串中所有 item 节点位上 1 到 item 数的概率值;
4) 将最大概率值的 item 节点位上 item 值作为当前扰动 item 值;
5) 根据公式(10)求出用户 u 对应的扰动邻接向量 $\overline{a}_u$;
6) 将 $\overline{A}$ 第 u 行替换为 $\overline{a}_u$;
7) **Return** $\overline{A}$.

本地推荐: 由于推荐结果在某种意义上来说是对用户未来行为的预估, 因此在服务器上集中生成推荐结果的方式同样面临着泄露用户偏好趋势的隐私风险. 所以在提出的模型中, 对于每一个用户最终的推荐结果由用户下载对应模型参数自行计算. 例如: 基于 LightGCN 网络模型, 最终的模型预测被定义为 user-item 最终嵌入表示的内积, 即最终每个用户 x 将可以下载相应的最终用户嵌入参数($e_x^{(k+1)}$)以及所有最终项目嵌入参数($e_{i\in I}^{(k+1)}$), 最后的本地结果如公式(12)所示:

$$\hat{y}_{xi}=e_x^T e_i \tag{12}$$

3.6 效率与安全性分析

与直接发布简单的匿名图数据相比, 以 NDB_s 形式发布可用图数据在一定程度上更加安全. 如果攻击者想要从 NDB_s 恢复原始数据, 他需要逆向 *QK*-NDB 需要求解相应的 *K*-SAT 实例. 从文献[11]中可以知道, 目前 *QK*-NDB 很难逆转, 攻击者很难恢复原始数据.

QK-hidden 算法的安全性主要受参数 K, r, p 和 q 的影响: 参数 K 决定了 NDB_s 中的每个记录有 K 个确定的位; 参数 r 控制 NDB 的大小, r 越大, 提供的原始数据信息就越多, 因此 r 太大也会让 NDB_s 容易逆转. 参数 p, q, K 决定了每个属性的第 i 位与原始串不同的概率 P_{diff_i}, P_{diff_i} 的值越接近 0.5, 该位的估算精度就越低, 相应的安全性也会越高. 因此, 可以根据用户的需求, 通过改变这些参数来平衡原始图结构数据的安全性和准确

© 中国科学院软件研究所 http://www.jos.org.cn

性. 但是, 攻击者可以通过累乘每个扰动节点的概率来计算扰动图 $NDB_{Graph}(\bar{V},\bar{E})$ 等于原始图的概率, 如公式(13)所示:

$$P_{equal}=\prod_{w=1}^{2|E|}\Pr(Node_w)=\prod_{w=1}^{2|E|}\max(\Pr(s_w^{\text{int}})) \tag{13}$$

当 P_{equal} 的值越小, 说明发布数据估算精度越低, 数据安全性就越高; 反之, 安全性越低.

- *QK*-hidden 算法效率的理论分析.

在算法 1 中: 由步骤 4)知道算法要生成 N 条负数据库记录, 所以循环的迭代次数为 N; 步骤 6)中找到合适 *rnd* 的时间复杂度是 O(K); 步骤 9)中找到合适的 a 使其满足该不等式的时间复杂度是 O(K); 而步骤 10)改变串形状的时间复杂度为 O(m). 算法的总体时间复杂度为 O(N×(K+m)). 其中, N 是要生成的负数据库的记录数量, K 是常数, m 是二进制串的位数. 初始化步骤的时间复杂度可以忽略不计. 此外, 还需要考虑随机数生成的开销, 但在大多数情况下, 随机数生成的时间复杂度可被视为常数.

- 常见图数据受到匿名攻击时的安全性概率分析:

图的节点再识别攻击是指攻击者根据先验知识, 从发布的匿名图中重新识别出一个特定的目标节点. 在基于度的重识别攻击中, 攻击者拥有目标节点 v 的度值 $\deg_v$ 作为背景知识, 目标是从匿名图中识别出节点 v. 通常, 基于度的攻击要求目标节点在原始图中具有度属性的唯一值, 否则目标节点不能被唯一识别. 在真实的社交网络数据集中, 度分布往往呈现长尾形式(即大部分的节点度数小, 只有少数节点度数大). 因此在真实数据集中, 具有识别条件的节点 v 的度 $\deg_v$ 通常较大. NDB_{Graph} 中每个节点的概率是 $\max(\Pr(s_w^{\text{int}}))$, 所以如果 NDB_{Graph} 中的一个节点 v 具有目标度值, 攻击者可以根据它的邻居节点 *Neighbor*(v)概率计算出这个节点是目标节点的概率, 即节点 v 被正确重新识别的概率可以通过公式(14)计算:

$$P_v^{equal}=\prod_{i}^{\deg_v}\Pr(Node_i),\ Node_i\in Neighbor(v) \tag{14}$$

同样, 对于其他攻击, 例如友谊攻击[42]、邻里攻击[43]等, 攻击者只能根据扰动图猜测出的目标节点或子图与原图相等的概率. 类似地, 可通过被识别节点的所有节点位上估算的概率累乘积计算出来, 用户可以使用不同的参数生成有不同安全需求的 NDB_s.

4 实验结果与分析

4.1 实验设置介绍

通过在四组真实图数据集和一组随机生成的图数据集上进行实验, 实验在不同的参数设置下估算了以下几个图结构数据集基于本文提出发布方法的度分布、三角形计数、平均聚类系数和平均最短路径属性, 并与原始数据进行了比较. 考虑到真实图数据集节点数、有向无向性、连通性以及和随机图的差异性, 本文选取了以下数据集来进行实验:

(1) Karate 数据集[44]: 美国某大学空手道部 34 名成员之间的社交关系网络, 包含 78 条无向边;

(2) Email-EU 数据集[45]: 欧洲某研究所核心成员之间的电子邮件联系关系构成的网络, 由 1 005 个节点和 25 571 条有向边组成;

(3) NetScience 数据集[45]: Newman 编制的网络中理论和实验科学家的协作网络, 由 1 589 个节点(科学家)和 2 742 条无向边组成;

(4) Facebook 数据集[45]: Facebook 一些用户的好友列表, 由 4 039 个节点(用户)和 88 234 条无向边组成;

(5) 随机生成图: 由随机 ER 图模型以 0.1 的概率生成的具有 500 个节点和 12 403 条边的无向图.

由于每个图形数据集的节点和边的数量不相等, 其编码字符串的节点位数也不同. 当设置负数据库的参数时, 使用相同的参数 K, r 和 3 组不同的参数 p. 具体地, K=3, r=15. 第 1 组 p: p_1=0.725, p_2=0.175, p_3=0.1; 第 2 组 p: p_1=0.85, p_2=0.1, p_3=0.05; 第 3 组 p: p_1=0.925, p_2=0.065, p_3=0.01. 以概率参数 p_i 选择生成第 i 种类型(记

© 中国科学院软件研究所 http://www.jos.org.cn

录有 i 位与隐藏串不同)的负数据库记录, 而要使 QK-hidden 算法生成的 NDB 很难被局部搜索策略逆转[40], 参数 p_i 的取值需要满足特定条件. 在 3 组参数 p 中, 根据每个数据集节点数与边数, 相应地设置第 1 位和最后一位的参数 q, 而其他位 q 是固定的. 通过引入控制属性位取反的概率的参数 $q_1,\ldots,q_L$ (其中, $q_i\in(0,1)$, $q_1+q_2+\ldots+q_L=1$), 我们可以更加灵活地控制每一位上的估算准确度和隐私度. 对于隐私度较高的属性位, 我们不希望它被估算出准确的值, 因此可以通过控制取反概率来降低这些位置上估算的准确度, 从而提高隐私度; 同样, 对于隐私度较低的属性位, 我们可以通过控制取反概率来提高这部分估算的准确度, 从而提高整体估算的准确度. 具体地, 在 Karate 数据集上, 节点位长 L=6, 固定 $q_1,\ldots,q_5$=0.1; 在 Email-EU 数据集上, 每个节点位长 L=10, 固定 $q_2,\ldots,q_9$=0.02; 在 NetScience 据集上, 节点位长 L=11, 固定 $q_2,\ldots,q_{10}$=0.02; 在 Facebook 数据集上, 节点位长 L=12, 固定 $q_2,\ldots,q_{11}$=0.02.

在随机生成图上, 节点位长 L=9, 固定了 $q_2,\ldots,q_8$=0.05. 表 3 列出了每组数据集实验中具体其余参数 q 设置, 为了表述方便, 以下小节中参数设置中, $P(m)Q(n)$表示对应的第 m 组参数 p 以及第 n 组参数 q. 对于参数 q 的每次变换, 每组实验重复 10 次, 并取平均值作为最终结果. 所有实验都是在一台 PC 上进行的, 该 PC 采用 AMD Ryzen5 4600U 处理器, 配有 Radeon Graphics 2.10 GHz CPU 和 16.0 GB RAM, 操作系统为 Windows 10.

表 3 q 参数设置

	Karate	Email-EU	NetScience	Facebook	随机生成图
$Q(1)$	q_1=0.20, q_6=0.40	q_1=0.20, q_{10}=0.64	q_1=0.20, q_{11}=0.62	q_1=0.18, q_{12}=0.62	q_1=0.30, q_9=0.35
$Q(2)$	q_1=0.30, q_6=0.30	q_1=0.40, q_{10}=0.44	q_1=0.40, q_{11}=0.42	q_1=0.38, q_{12}=0.42	q_1=0.40, q_9=0.25
$Q(3)$	q_1=0.40, q_6=0.20	q_1=0.60, q_{10}=0.24	q_1=0.60, q_{11}=0.22	q_1=0.58, q_{12}=0.22	q_1=0.50, q_9=0.15
$Q(4)$	q_1=0.50, q_6=0.10	q_1=0.80, q_{10}=0.04	q_1=0.80, q_{11}=0.02	q_1=0.78, q_{12}=0.02	q_1=0.60, q_9=0.05

4.2 属性评估实验结果与分析

本节通过实验评估基于 QK-hidden 算法生成的 NDB_{Graph} 的实用性. 首先, 本文评估估计的 NDB_{Graph} 与原始图之间的 $L1$-误差和 Kolmogorov-Smirnoff (KS)距离. $L1$-误差度量两个度分布直方图之间的累积差值. 对于长度 M 相等的两个分布, 它们之间的 $L1$-误差定义如下:

$$L1\text{-}error=\sum_{i=0}^{M-1}|D_i-D_i'| \tag{15}$$

其中, D_i 和 D_i' 分别表示分布 D_i 和 D_i' 的第 i 个值.

KS 距离可以用来反映两个分布之间的相似度, 它等于两个分布的累积分布函数 $CDF(i)$值的最大差值. KS 距离越小, 两个分布之间的相似度越高. 以下公式是两个分布 D 和 D'间的 KS 距离定义:

$$KS(D,D')=\max|CDF_D(i)-CDF_{D'}(i)| \tag{16}$$

图 4 和图 5 给出了 5 个数据集在不同参数设置下, NDB_{Graph} 和原始图之间度分布的 $L1$-误差和 KS 距离的估算结果. 其中, NetScience 数据集、Email-EU 数据集、Facebook 数据集的度分布 $L1$-误差或 KS 距离在最好的情况下几乎接近于 0, 节点数和边数最多的 Facebook 数据集的 $L1$-误差在最坏的情况下只有 1 400 左右. 因此可以认为, 负表示的图结构数据发布方法具有较小的数据精度损失和较高的度效用.

- 三角形计数

图 6 展示了 NDB_{Graph} 的三角形计数, 并将其与原始图形进行了比较. 在第 2 组和第 4 组参数 q 的变换下, 结果与原始结果非常接近. 对于 NetScience, Email-EU, Facebook 这 3 组真实数据集, 不同参数下的变化幅度都比较大. 一种可能的推测认为: 这是因为真实的图数据中每个节点(实体)本身的交流范围是有限的, 节点与节点之间的连接不是完全随机的, 而 NDB_{Graph} 推算过程中的随机性加上上述图数据集的规模较大, 导致其三角形计数的变化范围很大. 相比之下, 随机生成图三角形数的变化幅度就比较小.

© 中国科学院软件研究所 http://www.jos.org.cn

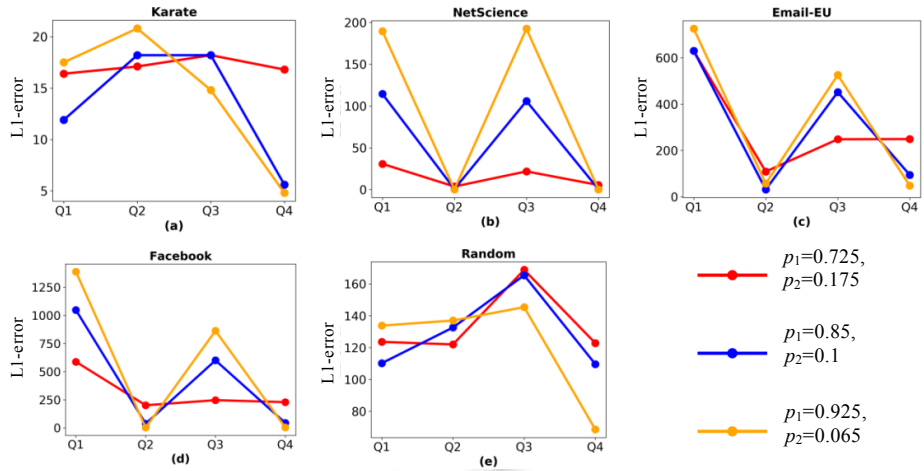

图 4　不同参数设置下度分布的 $L1$-误差

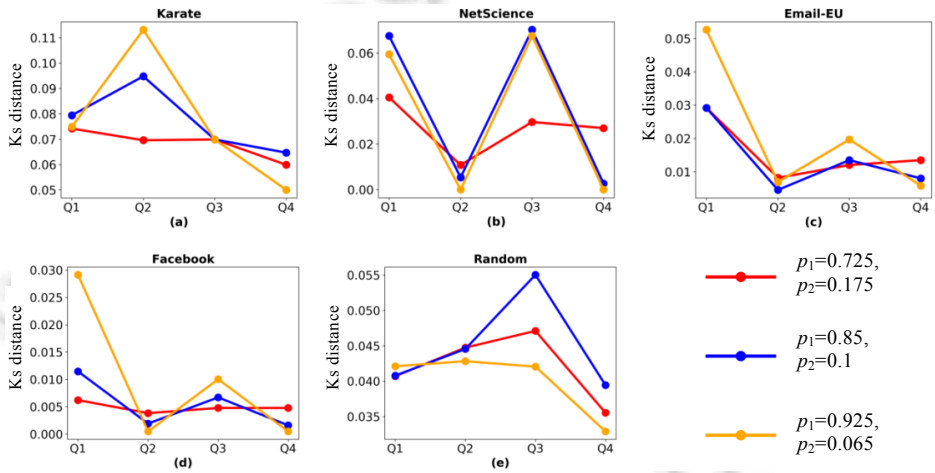

图 5　不同参数设置下度分布的 KS 距离

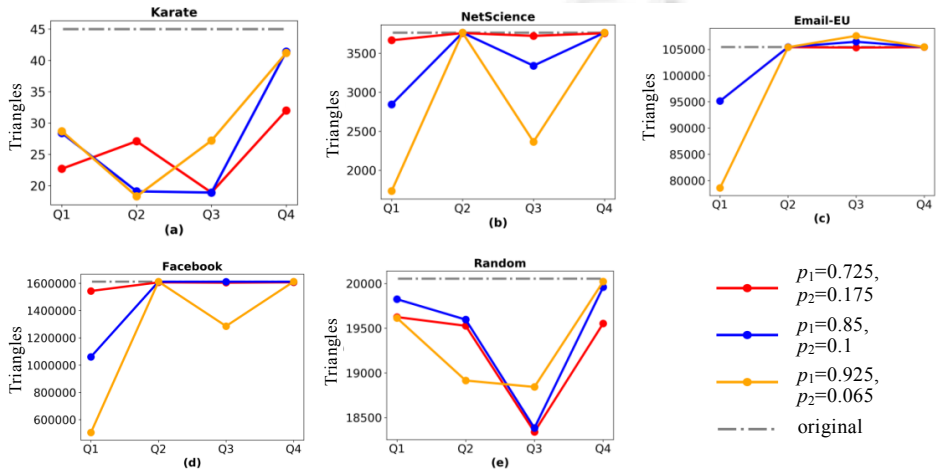

图 6　不同参数设置下三角形数

© 中国科学院软件研究所 http://www.jos.org.cn

- 平均聚类系数:

图 7 展示了 NDB_{Graph} 的平均聚类系数(以下简写为 CC), 可以注意到, 图中在 NetScience 数据集上添加了一条绿色虚线. 这是因为NetScience数据集具有一些孤立的离散点, 没有与它们连接的边, 所以生成NDB_{Graph}时有较大概率会无法直接推算得出这些孤立点, 但是可以根据最大顶点序号数 $v_{\max}$ 得到孤立点. 具体来说, 就是将 $v_{\max}-|\bar{V}|$ 个节点添加到 NDB_{Graph} 中. 这里的绿色虚线表示的是加上这些孤立点后的平均聚类系数, 灰色虚线是除去孤立点后的值, 实验中对扰动图 NDB_{Graph} 估算时没有包括孤立点. 总的来看, 平均聚类系数变化趋势与三角形数的表现差不多.

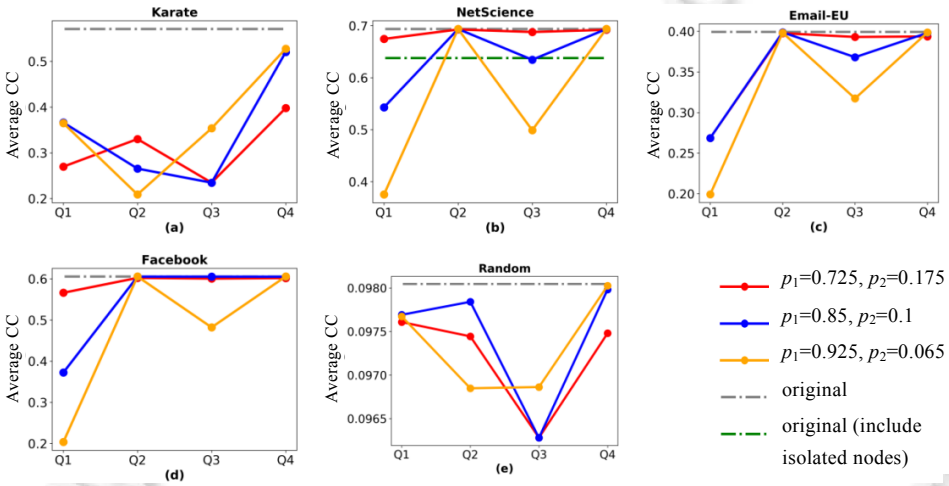

图 7 不同参数设置下平均聚类系数

- 平均最短路径长度

图 8 展示了 NDB_{Graph} 的平均最短路径, NetScience 数据集上的绿色虚线情况类似于平均聚类系数. 5 组实验结果(NetScience 数据集以未包含离散点结果为基准)中, Karate 数据集上平均最短邻径长度与原始数据相比最大相差 4.7%, NetScience 数据集上最大相差 9.9%, Email-EU 数据集最大相差 210%, Facebook 数据集上最大相差 27%以及随机数据集上最大相差 0.3%. 其中, Email-EU 数据集第 1 组 q 下和原始图相差较大. 可以认为: 这是由于它是一个非全连通图, 在这组 q 下较强的随机性导致了不同连通分量被连通了, 从而使得总的平均最短路径相差较大.

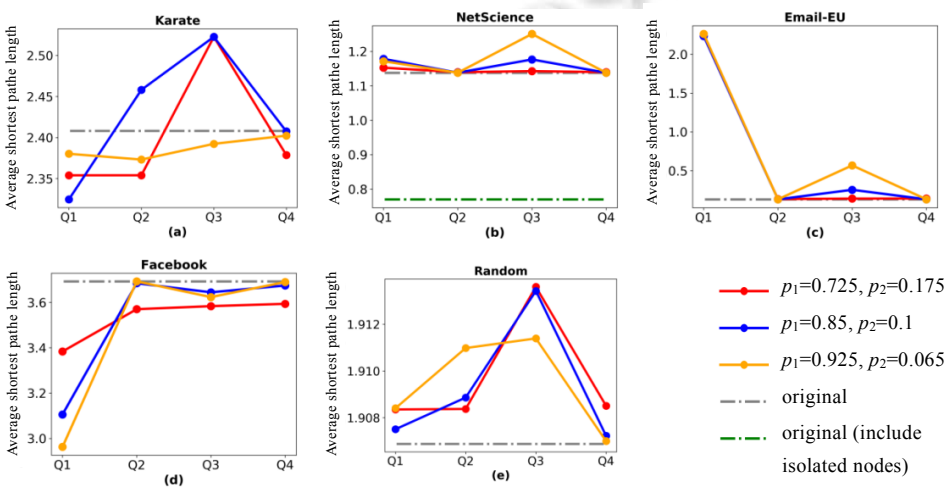

图 8 不同参数设置下平均最短路径长度

© 中国科学院软件研究所 http://www.jos.org.cn

- 参数 q 对实验的影响

注意到实验中, 以上属性在不同参数 q 变化下整体的变化似乎不太有规律. 例如在 NetScience, Email-EU, Facebook 数据集上第 2 组 q 参数、第 4 组 q 参数下, 无论是度分布、三角形数、平均聚类系数还是平均最短路径都非常接近原始数据, 而第 1 组 q 参数、第 3 组 q 参数下的结果相对比较远离原始数据结果, 其他数据集上的结果变化看上去也没有规律. 对此, 本文猜想为: 扰动图计算方法中只用到了最大的概率对应的十进制值, 即使最低位权重低, 但在图结构数据中即使差1都将导致这条边的端点改变, 此时两个点的度值也会改变, 因此除了 q_1, q_L 也对结果有一定影响. 举例来说, 表 4 列出了 Karate 以及 Facebook 两个数据集上每个位上对应的 P_{diff} 值与结果进行了对照比较, 可以看到每一组同样 p 参数下, q_1 对应 P_{diff} 值虽然与 0.5 差值不断增大, 但随着 q_L 对应的 P_{diff} 值与 0.5 差值的减小, 最终结果相对就越偏离原始数据. 例如: Karate 数据集在 $P(1)Q(3)$ 组中, 由于 q_6 对应的 P_{diff} 值接近 0.5, 导致这一组的以上属性结果变得较不精确; 而在第 $P(1)Q(4)$组中, 由于 q_6 对应的 P_{diff} 值又远离了 0.5, 导致最终的结果又变得比较精确. 同样, Facebook 数据集上的第 $P(1)Q(3)$组、$P(1)Q(4)$组中也有类似变化. 以上两组仅为举例描述, 其他组的实验结果变化趋势也和相应的 P_{diff} 值对应, 因此这也进一步验证了以上猜想.

表 4 Karate 以及 Facebook 数据集上每个位上对应的 P_{diff} 值

Dataset	Karate			Facebook		
Group	q_1 对应的 P_{diff} 值	q_2~q_5 对应的 P_{diff} 值	q_6 对应的 P_{diff} 值	q_1 对应的 P_{diff} 值	q_2~q_{11} 对应的 P_{diff} 值	q_{12} 对应的 P_{diff} 值
$P(1)Q(1)$	0.504	0.337	0.671	0.646	0.169	0.863
$P(1)Q(2)$	0.604		0.604	0.794		0.810
$P(1)Q(3)$	0.670		0.504	0.855		0.691
$P(1)Q(4)$	0.717		0.337	0.888		0.169
$P(2)Q(1)$	0.444	0.286	0.615	0.590	0.138	0.832
$P(2)Q(2)$	0.545		0.545	0.752		0.771
$P(2)Q(3)$	0.615		0.444	0.823		0.638
$P(2)Q(4)$	0.666		0.286	0.862		0.138
$P(3)Q(1)$	0.405	0.254	0.576	0.550	0.120	0.808
$P(3)Q(2)$	0.505		0.505	0.721		0.740
$P(3)Q(3)$	0.576		0.405	0.798		0.590
$P(3)Q(4)$	0.629		0.254	0.841		0.120

- 对比实验

为了更好地验证本文方法的有效性, 本文还与其他图数据隐私保护方法进行了对比实验. 图 9–图 11 是分别在 Karate 数据集、Facebook 数据集和 NetScience 数据集上与文献[25]中的 PBCN 方法进行比较的结果(在本文中提到的 5 个数据集中: Karate 数据集仅是包含数十名成员之间的社交关系网络, Facebook 数据集包含数千名成员、近十万条边的大数据集, NetScience 数据集包含数千名成员、数千条边的中等规模数据集, 作为代表, 选择了较小、较大和中等规模的的 3 个数据集来观察结果). PBCN 是基于聚类和噪声的隐私保护方法, 该方法可分为几个步骤: 度序列聚类、预处理、度序列扰乱、图重建和后预处理, 其主要是基于 K-Means 聚类的群构建和拉普拉斯机制的噪声分配来实现隐私保护. 为对两种方法有一个准确的测量标准, 本文使用文献[25]中基于邻接度的测量算法. 具体来说, 本文通过分别改变两种方法中的参数, 对具有相同隐私保护级别[25] P 的扰动图的上述属性进行比较, 并取每 10 次实验的均值作为参数变化的每组结果. PBCN 方法在现有方法中具有较高的准确率, 可以看出: 在相同的隐私保护级别 P 下, 本文的方法在 Karate 数据集、Facebook 数据集和 NetScience 数据集的度分布特征和最短路径长度方面表现优于 PBCN 方法, 且在 Facebook 数据集上的平均聚类系数也优于 PBCN 方法. 尽管本文提出的方法在 Karate 的平均聚类系数和 Facebook 的三角形计数方面不如 PBCN, 但它们仍然很接近原始数据.

© 中国科学院软件研究所 http://www.jos.org.cn

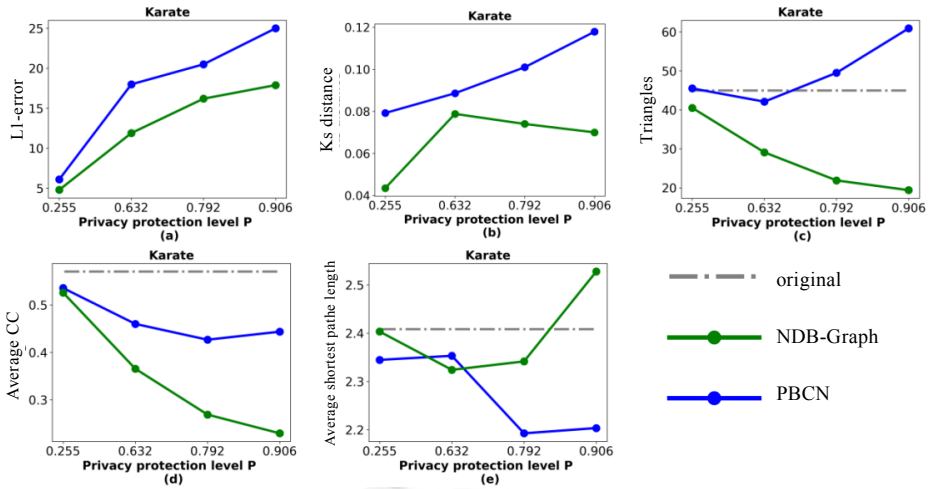

图 9 Karate 数据集上 NDB-Graph 和 PBCN 方法的对比

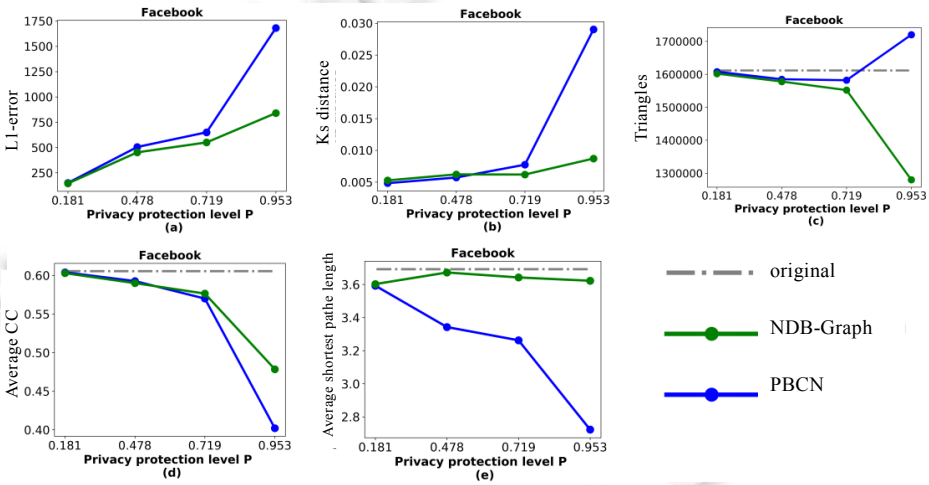

图 10 Facebook 数据集上 NDB-Graph 和 PBCN 方法的对比

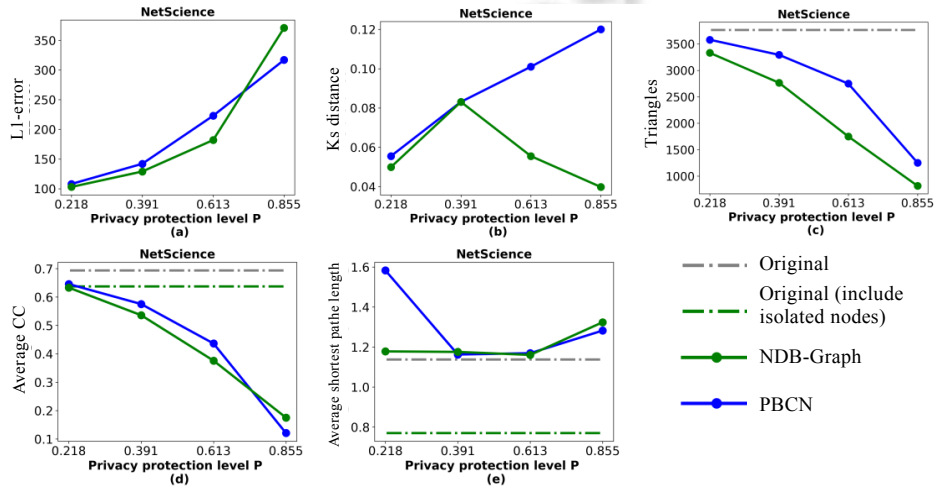

图 11 NetScience 数据集上 NDB-Graph 和 PBCN 方法的对比

© 中国科学院软件研究所 http://www.jos.org.cn

4.3 图聚类实验结果与分析

为了进一步证明 NDB_{Graph} 的实用性, 本文对 NDB_{Graph} 进行图聚类实验. Mülle 等人[21]提出的 PIG 方法是一种基于聚类差分隐私图数据隐私保护方法, 本文在 Facebook 数据集上重新实现了该方法, 并使用相同的 SCAN 算法参数与本文的方法进行比较. 为了评价两种方法的性能, 实验中采用 $F1$-score 和调整兰德系数(ARI)作为评价指标. 每组参数下进行 10 次实验后, 取平均值. 两种方法的最终聚类性能见表 5, 其中, $P(n)Q(m)$表示对应于第 4.1 节和第 4.2 节中介绍的 Facebook 数据集的第 n 组参数 p 和 m 组参数 q.

表 5 SCAN 聚类算法下与 PIG 算法效果对比

	NDB-Graph			PIG[19]		
Parameters	$F1$	ARI	s	ε	$F1$	ARI
$P(1)Q(1)$	0.458	0.451	0.01	5.293	0.785	0.726
$P(1)Q(2)$	0.893	0.887	0.02	4.595	0.741	0.602
$P(1)Q(3)$	0.899	0.865	0.03	4.185	0.701	0.543
$P(1)Q(4)$	0.892	0.892	0.04	3.892	0.689	0.517
$P(2)Q(1)$	0.284	0.174	0.05	3.664	0.667	0.499
$P(2)Q(2)$	0.986	0.994	0.06	3.476	0.669	0.487
$P(2)Q(3)$	0.752	0.709	0.07	3.317	0.656	0.485
$P(2)Q(4)$	0.993	0.985	0.08	3.178	0.430	0.469
$P(3)Q(1)$	0.249	0.107	0.09	3.055	0.237	0.455
$P(3)Q(2)$	0.992	0.987	–	–	–	–
$P(3)Q(3)$	0.618	0.597	–	–	–	–
$P(3)Q(4)$	**1.000**	**1.000**	–	–	–	–

可以看出, 本文的方法在最佳条件下 $F1$ 或 ARI 指数都可以高达 0.99 甚至是 1. 虽然 PIG 和本文的方法之间的安全参数没有比较标准, 但可以看到, 本文的方法对聚类性能的控制更敏感. 因此本文认为, 图数据的负表示在图聚类中也具有很高的效用. 具体来说, 本文统计各个类别的 TP, FP, FN, TN, 加和构成新的 TP, FP, FN, TN, 然后计算 $Precision$ 和 $Recall$, 根据下列公式得到全局指标 micro-$F1$:

$$Precision=\frac{TP}{TP+FP} \tag{17}$$

$$Recall=\frac{TP}{TP+FN} \tag{18}$$

$$F1\text{-}Score=\frac{2\times(Precision\times Recall)}{Precision+Recall} \tag{19}$$

ARI 指数反映了两种聚类结果的重叠程度, 用 a_{ij} 和 b_{ij} 表示两种聚类下的重叠实例数, n_{ij} 表示总重叠实例数量, ARI 指数可以由以下公式(20)计算, 其中, i 代表实际的分类数, j 代表这里的聚类类别数, n 代表实例总数:

$$ARI=\frac{\sum_{ij}\binom{n_{ij}}{2}-\left[\sum_i\binom{a_i}{2}\sum_j\binom{b_j}{2}\right]\Big/\binom{n}{2}}{\left[\sum_i\binom{a_i}{2}\sum_j\binom{b_j}{2}\right]\Big/2-\left[\sum_i\binom{a_i}{2}\sum_j\binom{b_j}{2}\right]\Big/\binom{n}{2}} \tag{20}$$

4.4 基于NDB的GNN推荐系统性能评估

随着对图结构数据的研究与应用变得广泛, 随之带来的是用户关系隐私暴露的风险. 另外, 在应用了图神经网络的推荐系统中, 一方面, 用户享受着来自服务器的推荐服务; 另一方面, 也将面临着交互数据、喜好偏好隐私泄露的风险. 针对该隐私泄露的风险, 本文基于负数据库技术提出了一种自定义隐私需求联邦推荐系统框架，实验具体如下

本节将基于 LightGCN 网络模型, 并在以下两个数据集上分别验证所提出方法的性能以及结果分析.

- Lastfm 1K 音乐推荐数据集: 由 lastfm 发布的用户收听纪录数据集, 包含 1 892 个用户、4 489 个项目(音乐)以及 42 135 个交互;

© 中国科学院软件研究所 http://www.jos.org.cn

- Gowalla 数据集: 基于地理位置的推荐系统数据集, 包含 29 858 个用户、40 981 个项目以及 1 027 370 个交互.

本文使用 LightGCN 模型, 它是对神经图协同过滤(NGCF)模型的简化. 在 NGCF 的基础上, 删除了变换矩阵和激活函数, 同时只聚合相连的邻居, 不整合目标节点本身, 并将单独的 GCN 层改为加权求和. 实验结果表明, 该简化后的模型仍具有较高的精度.

对于两个数据集中的所有用户, 实验中的一些参数设置见表 6.

表 6 神经网络部分参数设置

EPOCH	Layers	Embedding dim	学习率	Top-K	BPR batch
1 000	3	64	0.001	20	2 048

实验中, 本文假设用户对隐私保护级别的需求呈正态分布, 即: 对隐私保护级别有一般需求的人群占大多数人, 以及少数人对隐私保护级别有较高或较低的需求. 因此, 实验中每个用户对应的参数 sec_{need} 都是由一个符合正态分布的随机函数生成的. 实验代码基于 PyTorch 实现, 总隐私预算的每次变换重复 10 组实验. 最终结果取平均值. 本文设置负数据库生成参数 K=3, r=15 和 p_1=0.85, p_2=0.1, p_3=0.05. 对于每个数据集, 对应的 $q_2,\ldots,q_{L-1}$ 参数设置情况以及相应的 P_{diff} 值见表 7 和表 8.

表 7 Lastfm 1K 数据集上的 $q_2,\ldots,q_{12}$ 参数设置

Group	$q_2,\ldots,q_{12}$	$q_2,\ldots,q_{12}$ 对应的 P_{diff}	q_1+q_{13}	(q_1+q_{13})对应的 P_{diff}
1	0.02	0.14773	0.78	0.87113
2	0.03	0.20634	0.67	0.85308
3	0.04	0.25742	0.56	0.82915
4	0.05	0.30230	0.45	0.79592
5	0.06	0.34210	0.34	0.74662

表 8 Gowalla 数据集上的 $q_2,\ldots,q_{18}$ 参数设置

Group	$q_2,\ldots,q_{18}$	$q_2,\ldots,q_{18}$ 对应的 P_{diff}	q_1+q_{19}	(q_1+q_{19})对应的 P_{diff}
1	0.008	0.09201	0.848	0.91483
2	0.010	0.11243	0.810	0.91119
3	0.020	0.20213	0.620	0.88705
4	0.030	0.27536	0.430	0.84488
5	0.040	0.33628	0.240	0.75247

根据数据集不同的 item 数, Lastfm 1K 数据集对应的 L=13, Gowalla 数据集对应的 L=19. 实验中使用的主要评价指标是精度和归一化折损累计增益(NDCG), *Precision@K* 是衡量用户在推荐的 K 个项目中将选择的项目得分的度量. 计算方法如公式(21)所示:

$$Precision@K=\frac{|R^K(u)\cap T(u)|}{K} \tag{21}$$

其中,$T(u)$表示正确标注的项目集, $RK(u)$表示前 K 个推荐项目集. NDGC 以 DCG/IDCG (ideal DCG, 又称最理想化的 DCG)为代表. 列表中的项目顺序很重要, 不同位置的贡献是不同的. 一般来说, 排名靠前的项目影响较大, 排名靠后的项目影响较小. DCG 会增加排名第一的项目的影响力, 而排名靠后的项目则削弱其影响力. 具体地, 公式(22)所示为 *NDCG@K* 的计算方法:

$$NDCG@K=\frac{1}{|U|}\sum_{u\in U}\frac{\sum_{k=1}^{K}\frac{I(R_k^K(u)\in T(u))}{\log(k+1)}}{\sum_{k=1}^{K}\frac{1}{\log(k+1)}} \tag{22}$$

为了验证所提方法的有效性, 表 9 列出了与基于 EdgeRand 方法的比较结果论文[46]中关于边缘差分隐私的研究. 该算法天然地保留了邻接矩阵的稀疏结构, 本文将其作为用户的输入扰动算法. 其中, 本文将参数 s 设置为 0.00001, 0.0001, ..., 0.1, 其对应的隐私预算ε见表 9. Lastfm 1K 数据集上的最佳结果略高于提出的方法, 但相应的隐私预算ε为 12.21, 这意味着添加的噪声较少, 并且 Gowalla 数据集上的最终结果比提出的方法稍

© 中国科学院软件研究所 http://www.jos.org.cn

差. 可以看出: 随着噪声水平的增加, EdgeRand 方法的精度显著下降; 相比之下, 所提出的方法的精度相对稳定, 保持了原始未保护模型一半以上的精度.

表 9 基于 LightGCN 的实验结果

Dataset		Lastfm 1K		Gowalla	
Result		*Precision*@20	*NDCG*@20	*Precision*@20	*NDCG*@20
Ours	Group 1	0.075 1	0.208 5	0.055 8	0.154 3
	Group 2	0.075 0	0.207 5	**0.055 9**	**0.154 5**
	Group 3	0.072 9	0.201 4	0.055 6	0.153 9
	Group 4	0.065 2	0.177 6	0.052 7	0.145 8
	Group 5	0.040 9	0.105 6	0.035 9	0.095 7
EdgeRand[42]	ε=12.206	**0.075 3**	**0.209 1**	0.055 6	0.154 1
	ε=9.903	0.074 2	0.205 4	0.052 9	0.147 1
	ε=7.600	0.067 4	0.186 3	0.034 1	0.087 6
	ε=5.293	0.031 8	0.079 3	0.021 8	0.049 7
	ε=2.944	0.003 0	0.006 4	0.015 0	0.033 1
Original		0.075 2	0.209 6	0.055 9	0.154 7

4.5 安全性实验结果分析

本节在 Karate 和 Facebook 数据集上, 通过计算公式(9)中的 P_{equal} 对安全性进行实验. 表 10 展示了在两个数据集上计算 P_{equal} 值的实验结果, 实验结果表明: 在 Lastfm 1K 数据集上, 当$-\log_2 P_{equal}$ 值小于 2 000 时(前 3 组), 推荐系统精度、NDCG 值下降不明显, 并且前两组中精度下降几乎都在 0.3%以内. 但是即使如此, $-\log_2 P_{equal}$ 值也有 300 以上, 意味着攻击者在这些参数下猜出正确原始数据的概率仍然非常小. 在 Gowalla 数据集上, 前 3 组中, 精度、NDCG 值也几乎接近原始模型, 相应的$-\log_2 P_{equal}$ 值在 400 以上. 因此可以认为, 以上几组实验中, 具有一定安全性的情况下, 提出的模型也仍然具有较高的精度.

表 10 实验中 Lastfm 1K 和 Gowalla 数据集上的$-\log_2 P_{equal}$ 值

Group	Lastfm 1K 数据集上的$-\log_2 P_{equal}$ 值	Gowalla 数据集上的$-\log_2 P_{equal}$ 值
Group 1	325.905	429.173
Group 2	655.063	422.477
Group 3	1 731.799	6 983.488
Group 4	7 103.884	108 438.827
Group 5	29 452.149	894 138.158

同样的, 表 11 展示了在两个数据集上计算 P_{equal} 值的实验结果. 为了表示方便, P_{equal} 值的计算结果用 $-\log_2 P_{equal}$ 值表示(越大, 代表数据安全性越高). 虽然$-\log_2 P_{equal}$ 值在一些准确度较高的参数设置下相对较小, 比如在 Karate 数据集上最小值为 1.49, 这也意味着攻击者在这些参数下猜出正确原始数据的概率小于 50%, 因此也导致了较高的准确度. 在某些参数设置下, 实验中的$-\log_2 P_{equal}$ 值很高, 比如 Facebook 数据集的最大值是 65 013, Karate 数据集的最大值是 193, 因此对应的发布数据相对原始数据的误差也较大.

表 11 Karate 与 Facebook 数据集上的$-\log_2 P_{equal}$ 值

Parameters	Karate 数据集上的$-\log_2 P_{equal}$ 值	Facebook 数据集上的$-\log_2 P_{equal}$ 值
$P(1)Q(1)$	70.183	2 230.058
$P(1)Q(2)$	41.202	211.002
$P(1)Q(3)$	167.298	297.441
$P(1)Q(4)$	25.441	227.187
$P(2)Q(1)$	30.940	21 230.430
$P(2)Q(2)$	92.678	19.564
$P(2)Q(3)$	67.761	2 448.600
$P(2)Q(4)$	4.436	19.893
$P(3)Q(1)$	41.367	65 013.300
$P(3)Q(2)$	193.633	4.014
$P(3)Q(3)$	37.644	13 684.400
$P(3)Q(4)$	1.490	3.856

© 中国科学院软件研究所 http://www.jos.org.cn

5 总 结

在本文中, 本文提出了一种负表示图数据的发布方案. 首先, 通过 *QK*-hidden 算法将边序列二进制编码下的图转换为 NDB; 其次, 本文提出了一种从 NDB 获取扰动图 NDB-Graph 的方法以及一些特征属性的计算方法. 本文还对 NDB-Graph 的聚类性能进行了实验. 此外, 还提出了基于 NDB 的 GNN 推荐系统的框架模型. 具体而言, 用户根据不同用户对自身数据的隐私需求, 生成相应隐私保护级别的 NDB_s 数据并发送给服务器; 然后, 服务器接收来自不同用户的 NDB_s 数据, 聚合总扰动图矩后计算阵; 最后, 用户单独下载对应的模型参数用于本地推荐. 最终的实验结果表明, 负表示图数据方法能够很好地保持基本特征属性并具有一定的安全性. 所提出的基于 NDB 技术的图神经网络推荐系统的结果也表明, 本文的方法可以具有良好的准确性. 未来, 本文将继续扩展这项工作, 并尝试将其应用到其他可行的应用场景中.

References:

[1] Zhao G, Wang QG, Yao F, *et al*. A survey of large-scale graph neural network systems. Ruan Jian Xue Bao/Journal of Software, 2022, 33(1): 150−170 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6311.htm [doi: 10.13328/j.cnki.jos.006311]

[2] Abawajy JH, Ninggal MIH, Herawan T. Privacy preserving social network data publication. IEEE Communications Surveys and Tutorials, 2016, 18(3): 1974−1997. [doi: 10.1109/COMST.2016.2533668]

[3] Li XY, Zhang CH, Jung T, *et al*. Graph-based privacy preserving data publication. In: Proc. of the 35th Annual IEEE Int'l Conf. on Computer Communications (INFOCOM 2016). IEEE, 2016. 1−9. [doi: 10.1109/INFOCOM.2016.7524584]

[4] Dwork C. The differential privacy frontier. In: Proc. of the Theory of Cryptography Conf. Springer, 2009. 496−502.

[5] Hay M, Li C, Miklau G, *et al*. Accurate estimation of the degree distribution of private networks. In: Proc. of the 9th IEEE Int'l Conf. on Data Mining. IEEE, 2009. 169−178. [doi: 10.1109/ICDM.2009.11]

[6] Esponda F. Everything that is not important: Negative databases. IEEE Computational Intelligence Magazine, 2008, 3(2): 60−63. [doi: 10.1109/MCI.2008.919079]

[7] Esponda F, Forrest S, Helman P. Enhancing privacy through negative representations of data. Technical Report, Albuquerque: New Mexico University, 2004. 72−84.

[8] Dasgupta D, Saha S. Password security through negative filtering. In: Proc. of the 2010 Int'l Conf. on Emerging Security Technologies. IEEE, 2010. 83−89. [doi: 10.1109/EST.2010.37]

[9] Zhao DD, Luo WJ, Liu R, *et al*. Negative iris recognition. IEEE Trans. on Dependable and Secure Computing, 2015, 15(1): 112−125. [doi: 10.1109/TDSC.2015.2507133]

[10] Zhao DD, Luo WJ. Real-Valued negative databases. In: Proc. of the Artificial Life Conf. Cambridge: MIT, 2013. 884−890.

[11] Zhao DD, Hu XY, Xiong SW, *et al*. *k*-means clustering and kNN classification based on negative databases. Applied Soft Computing, 2021, 110(1): 107732−107747.

[12] Singh A, Bansal D, Sofat S. Privacy preserving techniques in social networks data publishing-a review. Int'l Journal of Computer Applications, 2014, 87(15): 9−14.

[13] Liu K, Terzi E. Towards identity anonymization on graphs. In: Proc. of the 2008 ACM SIGMOD Int'l Conf. on Management of Data. ACM, 2008. 93−106. [doi: 10.1145/1376616.1376629]

[14] Cheng J, Fu AWC, Liu J. K-Isomorphism: Privacy preserving network publication against structural attacks. In: Proc. of the 2010 ACM SIGMOD Int'l Conf. on Management of Data. ACM, 2010. 459−470. [doi: 10.1145/1807167.1807218]

[15] Ren XM, Jiang DX. A personalized (α,β,l,k)-anonymity model of social network for protecting privacy. Wireless Communications and Mobile Computing, 2022. 141−157.

[16] Campan A, Truta TM. Data and structural *k*-anonymity in social networks. In: Proc. of the Int'l Workshop on Privacy, Security.and Trust in KDD. Springer, 2008. 33−54.

[17] Tassa T, Cohen DJ. Anonymization of centralized and distributed social networks by sequential clustering. IEEE Trans. on Knowledge and Data Engineering, 2011, 25(2): 311−324. [doi: 10.1109/TKDE.2011.232]

© 中国科学院软件研究所 http://www.jos.org.cn

[18] Jiang HW, Zhan QH, Liu WJ, *et al*. Clustering-anonymity approach for privacy preservation of graph data-publishing. Ruan Jian Xue Bao/Journal of Software, 2017, 28(9): 2323−2333 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5178.htm [doi: 10.13328/ j.cnki.jos.005178]

[19] Zhou YH, Zhang B, Yang YG, *et al*. Cluster-based social network privacy protection method. Computer Science, 2019, 46(10): 154−160 (in Chinese with English abstract). https://www.jsjkx.com/CN/abstract/article18582.shtml [doi: 10.11896/jsjkx.180901749]

[20] Ahmed F, Liu AX, Jin R. Publishing social network graph eigenspectrum with privacy guarantees. IEEE Trans. on Network Science and Engineering, 2019, 7(2): 892−906. [doi: 10.1109/TNSE.2019.2901716]

[21] Mülle Y, Clifton C, Böhm K. Privacy-integrated graph clustering through differential privacy. In: Proc. of the EDBT/ICDT Workshops. 2015. 247−254.

[22] Liu P, Xu YX, Jiang Q, *et al*. Local differential privacy for social network publishing. Neurocomputing, 2020, 391: 273−279. [doi: 10.1016/j.neucom.2018.11.104]

[23] Chen R, Fung BCM, YU PS, *et al*. Correlated network data publication via differential privacy. The VLDB Journal, 2014, 23: 653−676. [doi: 10.1007/s00778-013-0344-8]

[24] Ma XB, Yang JY, Guan SY. Differentially private social graph publishing for community detection. In: Proc. of the Int'l Conf. on Security and Privacy in Communication Systems. Springer, 2020. 208−214. [doi: 10.1007/978-3-030-63095-9_11]

[25] Huang HP, Zhang DJ, Xiao F, *et al*. Privacy-preserving approach PBCN in social network with differential privacy. IEEE Trans. on Network and Service Management, 2020, 17(2): 931−945. [doi: 10.1109/TNSM.2020.2982555]

[26] Hay M, Miklau G, Jensen D, *et al*. Anonymizing social networks. 2007. 180. https://doi.org/10.1201/9781420091502-c15

[27] Min X, Li W, Yang JZ, *et al*. Self-supervised graph neural network with pre-training generative learning for recommendation systems. Scientific Reports, 2022, 12(1): 15882−15896. [doi: 10.1038/s41598-022-19528-3]

[28] Zhu ZG, Li WY, Jiang P, *et al*. Survey of graph neural networks in session recommender systems. Computer Engineering and Applications, 2023, 59(5): 55−69 (in Chinese with English abstract). http://cea.ceaj.org/CN/10.3778/j.issn.1002-8331.2207-0397 [doi: 10.3778/j.issn.1002-8331.2207-0397]

[29] Huang WM, Liu BS, Tang H. Privacy protection for recommendation system: A survey. Journal of Physics: Conf. Series. 2019, 012087−012096. [doi: 10.1088/1742-6596/1325/1/012087]

[30] Tan ZW, Zhang LF. Survey on privacy preserving techniques for machine learning. Ruan Jian Xue Bao/Journal of Software, 2020, 31(7): 2127−2156 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6052.htm [doi: 10.13328/j.cnki.jos.006052]

[31] Qiu YQ, Huang CY, Wang JZ, *et al*. A privacy-preserving subgraphlevel federated graph neural network via differential privacy. In: Proc. of the 15th Int'l Conf. on Knowledge Science, Engineering and Management. Springer, 2022. 165−177. [doi: 10.1007/978-3-031-10989-8_14]

[32] Liu ZW, Yang LW, Fan ZW, *et al*. Federated social recommendation with graph neural network. ACM Trans. on Intelligent Systems and Technology (TIST), 2022, 13(4): 1−24. [doi: 10.1145/3501815]

[33] Liu YQ, Fang SZ, Wang LY, *et al*. Neural graph collaborative filtering for privacy preservation based on federated transfer learning. The Electronic Library, ahead-of-print. 2022. 729−742. [doi: 10.1108/EL-06-2022-0141]

[34] Shin H, Kim S, Shin J, *et al*. 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]

[35] Fang L, Du BQ, Wu C. Differentially private recommender system with variational autoencoders. Knowledge-based Systems, 2022, 250: 109044−109058. [doi: org/10.1016/j.knosys.2022.109044]

[36] Gao C, Huang C, Lin DS, *et al*. Dplcf: Differentially private local collaborative filtering. In: Proc. of the 43rd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2020. 961−970. [doi: 10.1145/3397271.3401053]

[37] Wu CH, Wu FZ, Lyu LJ, *et al*. A federated graph neural network framework for privacy-preserving personalization. Nature Communications, 2022, 13(1): 3091−4001. [doi: 10.1038/s41467-022-30714-9]

[38] Jia HX, Moore C, Strain D. Generating hard satisfiable formulas by hiding solutions deceptively. Journal of Artificial Intelligence Research, 2007, 28: 107−118. [doi: 10.1613/jair.2039]

© 中国科学院软件研究所 http://www.jos.org.cn

[39] Zhao DD, Luo WJ, Liu R, *et al*. Experimental analyses of the *k*-hidden algorithm. Engineering Applications of Artificial Intelligence, 2017, 62: 331–340. [doi: 10.1016/j.engappai.2016.05.010]

[40] Zhao DD, Hu XY, Xiong SW, *et al*. A fine-grained privacy-preserving *k*-means clustering algorithm upon negative databases. In: Proc. of the 2019 IEEE Symp. Series on Computational Intelligence (SSCI). IEEE, 2019. 1945–1951.

[41] Lam SKT, Frankowski D, Riedl J. Do you trust your recommendations? An exploration of security and privacy issues in recommender systems. In: Proc. of the Int'l Conf. on Emerging Trends in Information and Communication Security. Springer, 2006. 14–29.

[42] Tai CH, Yu PS, Yang DN. Privacy-preserving social network publication against friendship attacks. In: Proc. of the 17th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. ACM, 2011. 1262–1270. [doi: 10.1145/2020408. 2020599]

[43] Zhou B, Pei J. Preserving privacy in social networks against neighborhood attacks. In: Proc. of the 24th IEEE Int'l Conf. on Data Engineering. IEEE, 2008. 506–515. [doi: 10.1109/ICDE.2008.4497459]

[44] Zachary WW. An information flow model for conflict and fission in small groups. Journal of Anthropological Research, 1977, 33(4): 452–473. [doi: doi.org/10.1086/jar.33.4.3629752]

[45] Leskovec J, Krevl A. SNAP datasets: Stanford large network dataset collection. 2014. http://snap.stanford.edu/data

[46] Wu F, Long YH, Zhang C, *et al*. Linkteller: Recovering private edges from graph neural networks via influence analysis. In: Proc. of the 2022 IEEE Symp. on Security and Privacy (SP). IEEE, 2022. 2005–2024. [doi: 10.1109/SP46214.2022.9833806]

附中文参考文献:

[1] 赵港, 王千阁, 姚烽, 等. 大规模图神经网络系统综述. 软件学报, 2022, 33(1): 150–170. http://www.jos.org.cn/1000-9825/6311.htm [doi: 10.13328/j.cnki.jos.006311]

[18] 姜火文, 占清华, 刘文娟, 等. 图数据发布隐私保护的聚类匿名方法. 软件学报, 2017, 28(9): 2323–2333. http://www.jos.org.cn/1000-9825/5178.htm [doi: 10.13328/j.cnki.jos.005178]

[19] 周艺华, 张冰, 杨宇光, 等. 基于聚类的社交网络隐私保护方法. 计算机科学, 2019, 46(10): 154–160. https://www.jsjkx.com/CN/abstract/article_18582.shtml [doi: 10.11896/ jsjkx. 180901749]

[28] 朱志国, 李伟玥, 姜盼, 等. 图神经网络会话推荐系统综述. 计算机工程与应用, 2023, 59(5): 55–69. http://cea.ceaj.org/CN/10.3778/j.issn.1002-8331.2207-0397 [doi: 10.3778/j.issn.1002-8331.2207-0397]

[30] 谭作文, 张连福. 机器学习隐私保护研究综述. 软件学报, 2020, 31(7): 2127–2156. http://www.jos.org.cn/1000-9825/6052.htm [doi: 10.13328/j.cnki.jos.006052]

赵冬冬(1989－), 男, 博士, 副教授, CCF 专业会员, 主要研究领域为隐私保护, 虹膜识别, 数据挖掘.

彭思芸(1998－), 女, 硕士生, 主要研究领域为机器学习, 隐私保护.

徐虎(1998－), 男, 硕士生, 主要研究领域为机器学习, 隐私保护.

周俊伟(1986－), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为算机视觉, 安全保护.

© 中国科学院软件研究所 http://www.jos.org.cn