

面向多方数据融合分析的隐私计算技术综述

刘圣龙¹, 黄秀丽², 江伊雯¹, 姜嘉伟¹, 田月池^{3,4,5}, 周泽峻^{3,4,5}, 牛犇^{3,4}

(1. 国家电网有限公司大数据中心, 北京 100052;

2. 国网智能电网研究院有限公司电力网络安全防护与监测技术实验室, 江苏 南京 210000;

3. 中国科学院信息工程研究所, 北京 100085;

4. 网络空间安全防御重点实验室, 北京 100085; 5. 中国科学院大学网络空间安全学院, 北京 100049)

摘要: 数据时代, 泛在共享环境下个人隐私信息面临的威胁十分广泛, 如App频繁超范围采集个人信息、大数据杀熟等, 跨系统交换的多方隐私计算需求迫切。因此, 聚焦面向泛在共享环境下跨系统交换的多方隐私计算需求, 以多方数据融合中隐私数据的安全共享和受控传播应用需求为切入点, 从多方隐私计算、多方隐私信息共享控制、多方数据协同安全计算等方面梳理了现有相关工作, 进行综述。分析了泛在共享环境下个人隐私信息保护的背景和研究现状。综述并对比分析了近年来多方隐私计算、多方隐私信息共享控制、多方数据协同安全计算等方面国内外最新研究成果, 针对多方隐私计算, 介绍了全生命周期隐私保护、隐私信息流转控制、敏感数据安全交换等技术; 针对多方隐私信息共享控制, 介绍了本地控制、延伸控制和脱敏控制三类技术; 针对多方数据协同安全计算, 介绍了学术界、产业界常用的技术。最后, 对多方隐私计算面临的挑战和发展方向进行展望, 传统的基于匿名、加扰、访问控制等技术的隐私脱敏方案、基于密码学的方案、基于联邦学习的方案等仍具有局限性, 而隐私计算理论给出的面向全生命周期保护的计算框架和信息系统框架, 需结合不同应用场景, 践行隐私信息全生命周期保护。

关键词: 隐私计算; 隐私信息共享控制; 协同安全计算

中图分类号: TP393

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2024078

Review of privacy computing techniques for multi-party data fusion analysis

LIU Shenglong¹, HUANG Xiuli², JIANG Yiwen¹, JIANG Jiawei¹,
TIAN Yuechi^{3,4,5}, ZHOU Zejun^{3,4,5}, NIU Ben^{3,4}

1. Big Data Center of State Grid Corporation of China, Beijing 100052, China

2. State Grid Laboratory of Power Cyber-Security Protection and Monitoring Technology,
State Grid Smart Grid Research Institute Co., LTD, Nanjing 210000, China

3. Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100085, China

收稿日期: 2024-04-30; 修回日期: 2024-10-15

通信作者: 牛犇, niuben@iie.ac.cn

基金项目: 国家电网有限公司总部管理科技项目 (5108-202218280A-2-393-XG)

Foundation Item: Science and Technology Project of State Grid Corporation of China (5108-202218280A-2-393-XG)

引用格式: 刘圣龙, 黄秀丽, 江伊雯, 等. 面向多方数据融合分析的隐私计算技术综述[J]. 网络与信息安全学报, 2024, 10(6): 24-36.

Citation Format: LIU S L, HUANG X L, JIANG Y W, et al. Review of privacy computing techniques for multi-party data fusion analysis[J]. Chinese Journal of Network and Information Security, 2024, 10(6): 24-36.

4. Key Laboratory of Cyberspace Security Defense, Beijing 100085, China

5. School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract: In the data era, threats to personal privacy information in ubiquitous sharing environments are widespread, such as apps frequently collecting personal information beyond scope, and big data-enabled price discrimination against frequent customers. The need for multi-party privacy computing for cross-system exchanges is urgent. This work focused on the needs of multi-party privacy computing for cross-system exchanges in ubiquitous sharing environments, taking the security sharing and controlled dissemination of private data in multi-party data fusion applications as the starting point, and provided reviews of existing relevant work from the perspectives of multi-party privacy computing, multi-party privacy information sharing control, and multi-party data collaborative secure computing. First, the background and research status of personal privacy information protection in a ubiquitous sharing environment were analyzed. Then, the latest domestic and foreign research results in recent years regarding multi-party privacy computing, multi-party privacy information sharing control, and multi-party data collaborative security computing were reviewed and comparatively analyzed. Regarding multi-party privacy computing, technologies such as full lifecycle privacy protection, privacy information flow control, and secure exchange of sensitive data were introduced. In terms of multi-party privacy information sharing control, localized control, extended control, and anonymization control techniques were discussed. In the aspect of multi-party data collaborative secure computing, commonly used techniques in both academia and industry were discussed. Finally, the challenges and development directions of multi-party privacy computing were prospected. There were still limitations for anonymity, scrambling, or access control-based traditional privacy desensitization measures, cryptography-based measures, and federated learning-based measures, while privacy computing theory provided a computational and information system framework for full-lifecycle protection, which needed to be combined with different application scenarios to implement full-lifecycle privacy information protection.

Keywords: privacy computing, privacy information sharing control, data collaborative secure computing

0 引言

数据时代，移动通信、网络技术和信息技术快速发展，以及智能设备的快速升级，促进了万物互联与智慧互通，社交、导航、网购、物流、支付、直播、共享经济、在线医疗、ChatGPT等服务模式不断涌现迭代，促使隐私信息频繁跨终端、跨系统、跨生态圈，甚至跨国境交互已成常态，并呈现海量多模趋势。然而，由此导致的App频繁超范围采集个人信息，后台信息服务系统中的隐私数据越权使用、大数据杀熟、个人画像结果滥用、个人信息过度留存等问题与日俱增，全球隐私泄露事件层出不穷，严重威胁个人隐私、企业秘密，甚至国家利益。世界各国高度重视，纷纷颁布法律法规，如欧盟颁布了《通用数据保护条例》（general data protection regulation, GDPR），我国先后颁布了《中华人民共和国数据安全法》和《中华人民共和国个人信息保护法》，说明个人信息保护与数据安全是两个不同问题，对应的关键

技术也存在差异。

泛在共享环境下个人隐私信息面临的威胁十分广泛，以图1为例，用户每日在健康App中完成运动打卡，通过可穿戴设备实时监测和采集其血压、血糖、心率等机体指标，并通过该App上传到后台服务器，后台服务器聚合多条用户数据，并共享给其他信息系统。根据这种隐私信息的传播共享特征，可以归纳出两种典型的隐私信息共享模式：一是纵向有损受控传播，如后台服务器对用户发布的信息进行脱敏处理（如隐去身份、模糊位置等）或转发（给当前App的其他用户，或其他相关App用户）；二是横向无损安全共享，如后台服务器将聚合后的多条用户数据与医疗、金融、科研机构等其他信息系统共享，以便为用户提供个性化医疗咨询、保险定价、新药研发等。在上述操作过程中，用户的ID、机体指标、位置轨迹、运动偏好等诸多隐私信息会在整个流程涉及的信息系统中传播，严重威胁用户隐私。

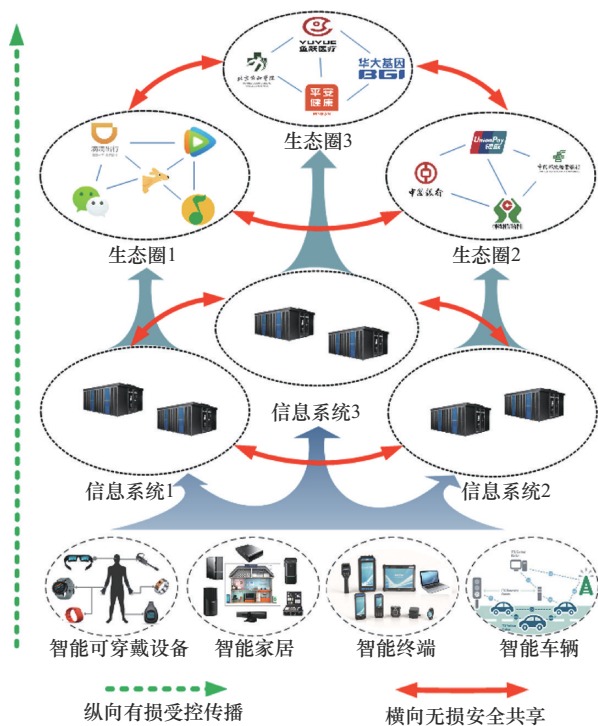


图1 泛在共享环境下信息服务架构

Figure 1 Structure of information service architecture in ubiquitous sharing environment

本文聚焦面向泛在共享环境下跨系统交换的多方隐私计算需求,以多方数据融合中隐私数据的安全共享和受控传播应用为切入,主要围绕多方隐私计算、多方隐私信息共享控制、多方数据协同安全计算3个方面的内容开展调研工作。多方隐私计算提供了实现数据隐私保护的技术手段;多方隐私信息共享控制提供了管理和控制数据共享的策略和机制;多方数据协同安全计算综合运用这些技术和机制,在保护数据隐私的前提下,实现多方的协同计算目标。

1 多方隐私计算

1.1 全生命周期隐私保护

在隐私保护的模型与机理方面,隐私演算(privacy calculus)^[1]是从经济学角度看待隐私保护问题,通过分析公开个人信息的风险和收益,决定是否公开。李风华等^[2]于2015年在国内外首次提出了“隐私计算”,涵盖了基本定义、框架、隐私计算应遵循的原则、算法设计准则、隐私保护效果评估等内容,从计算的角度给出了全生命周期权益控制的体系框架,并于2021年出版国

内外首部隐私计算学术专著《隐私计算理论与技术》^[3],进一步丰富了隐私计算理论的内涵和外延。隐私计算框架在隐私信息全生命周期的各个环节中建立应用场景、保护需求与计算模型之间的映射关系,包括隐私信息抽取、场景描述、隐私控制、隐私操作、隐私效果评估5个步骤,如图2所示。基于场景描述和保护需求,适应性地选择相应环节的计算方法实现相应的计算功能。

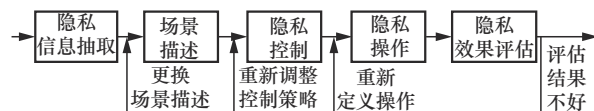


图2 隐私计算框架

Figure 2 Privacy computing framework

在全流程隐私保护方面,微软研究了服务全流程隐私保护,发布了“*Protecting Data and Privacy in the Cloud*”白皮书,定义了具体的用户隐私保护方法和流程。HP实验室、爱丁堡龙比亚大学使用加密机制将策略与数据关联,对策略进行属性编码和匿名化处理以防被恶意利用,为全生命周期内隐私数据访问控制提供支持。Koo等^[4]将大数据生命周期分为收集、存储、分析、利用、销毁5个阶段,根据识别出的威胁和安全问题定义了大数据生命周期的安全分类法。Anciaux等^[5]将生命周期策略绑定到上下文数据从而使数据逐步退化从而保护隐私。Ambrose^[6]认为在线信息的使用寿命变短,随时间推移而变化,因此可以利用分类法提出评估信息生命周期中用户行使“遗忘权”的代价,即是否值得行使“遗忘权”。Lin等^[7]提出了一种新的云服务选择方法,基于云服务覆盖用户数据整个生命周期的隐私保护能力来评估云服务,通过计算所有生命周期阶段的综合保护能力,对所有云服务进行排序并推荐给用户。Kang等^[8]提出了基于隐私信息生命周期的隐私保护技术的分类,将隐私保护技术分为3种:操作技术、通用技术和策略/管理技术。张文静等^[9]用真实位置和扰动位置之间的互信息度量单一时刻位置的隐私性,用真实位置和多个扰动位置之间的互信息度量攻击者掌握多个扰动位置时的隐私泄露。俞艺涵等^[10]提出了一种基于信息熵与BP(back propagation)神经网络的隐私数据度量与分级模型,使用信息熵对数

数据集隐私量进行分层计算,并利用BP神经网络,在不预设度量权值的情况下,实现对隐私数据的度量与分级。俞艺涵等^[11]提出了基于多层模糊综合评估的隐私保护效果评估方法,从隐私风险来源、隐私类别和隐私保护实施3个方面建立评估指标,通过划分最小评估单元,多层次、分区块地迭代使用模糊综合评估法进行评估。

1.2 隐私信息流转控制

流转控制策略设置主要依赖两大类解决方案帮助用户提高流转控制策略设置的效率,分别为基于访问控制模型和基于策略推荐的解决方案。

在基于访问控制模型的解决方案中,访问控制技术和模型能够提高用户设置流转控制策略的效率,并为用户的设置过程提供理论指导和安全性证明。在自主访问控制和强制访问控制的基础上,20世纪90年代,基于角色的访问控制(role-based access control, RBAC)^[12-13]被提出;在RBAC研究的基础上,Covington等^[14]提出了基于上下文的访问控制(context-based access control, CBAC);随着科技的发展和进步,新的场景和需求不断涌现,新的访问控制技术和模型也不断产生,例如,面向复杂信息系统中的细粒度访问控制和大规模用户动态扩展问题,产生了基于属性的访问控制(attribute based access control, ABAC)^[15];在这些访问控制技术的基础上,结合具体的应用场景和用户需求,出现了基于信任、基于任务和基于行为等一系列的新型访问控制技术及其管理模型。李风华等^[16-17]提出了面向网络空间的访问控制(cyberspace-oriented

access control, CoAC),通过对访问控制过程涉及的各种要素(包括访问请求实体、广义时态、接入点、资源、访问设备、网络、网络交互图以及资源传播链等)进行综合考虑,CoAC可对泛在网络中的信息及数据实施细粒度、多层次、灵活可变的访问控制,解决了动态开放的跨域泛在网络中的细粒度访问控制问题。

在基于策略推荐的解决方案中,策略推荐系统通过分析大量信息,为用户提供决策依据,减轻用户的决策负担。可基于实现原理的不同,策略推荐系统进行分类。最先出现的是基于矩阵分解的推荐系统,主要使用线性运算,从交互数据中学习用户的偏好等信息,如基于矩阵分解的不依赖辅助信息的推荐系统^[18-19]和基于矩阵分解的依赖辅助信息的推荐系统^[20-22];随着计算机数据处理能力的提升和待处理数据量的增加,出现了基于深度学习的推荐系统,采用多层神经网络对数据进行分析以实现更为精确的推荐,如基于深度学习的不依赖辅助信息的推荐系统^[23-25]和基于深度学习的依赖辅助信息的推荐系统^[26-28];除此之外,随着人工智能技术的不断发展,一些新的推荐技术也在不断被提出,如基于记忆网络的推荐系统^[29]等。策略推荐系统对比分析如表1所示。在信息流转前,可以使用推荐系统,结合分享的内容、隐私需求和人际关系等因素,生成信息流转控制策略,从而避免信息的不必要传播,减少信息的传播面和降低信息泄露和留存的风险。

1.3 敏感数据安全交换

在基于区块链的解决方案中,Wang等^[30]提

表1 策略推荐系统对比分析
Table 1 Comparison of strategy recommendation systems

文献	主要技术	辅助信息	流转控制策略设置	用户决策辅助效果
[18-19]	矩阵分解	无	简单易实现,初始设置快,但灵活性和精度有限	通过较简单的模型提供基本推荐,用户负担小,适合初次决策
[20-22]	矩阵分解+辅助信息	用户和物品的属性信息	通过辅助信息优化策略设置,调整灵活,但依赖于辅助信息的质量,设置复杂	通过辅助信息自动化推荐,减少了用户筛选无关信息的时间
[23-25]	深度神经网络	无	能处理复杂数据流转需求,但设置复杂,需大量数据支持	深度学习处理复杂决策场景,提供多样化且准确的推荐,减少用户自己发现新选项的负担
[26-28]	深度神经网络+辅助信息	用户和物品的属性信息	结合多种辅助信息,设置和调整策略更智能化,但设置和调整复杂,需大量计算资源	结合多种信息源,自动化程度高,最大程度减轻用户筛选负担
[29]	记忆网络	可包括用户历史行为等辅助信息	可记忆长时间跨度信息,策略设置高效,但模型复杂,设置时间较长	提供记忆和推理能力的推荐,减少用户重复选择和频繁决策的负担

出了基于区块链的安全、可审计的智能电网私有数据共享方案,该方案设计了基于区块链的新型框架用于无信任的私有数据计算和数据使用跟踪,其中智能合约指定细粒度的数据使用策略,同时使用 DPaaS 模型保护数据隐私,该模型只共享数据处理结果,不共享原始数据,但是,DPaaS 模型是基于区块链实现的分布式云,它的操作由区块链上的矿工执行,而矿工不方便管理,可能会执行恶意行为。Li 等^[31]提出了一种基于区块链的隐私保护和奖励性的物联网私有数据共享方案,该方案基于莫雷诺技术和可否认的环签名实现用户的匿名性和抗诬陷,数据拥有者匿名上传数据,数据使用者匿名获取数据,且能在匿名情况下否认对他的行为的诬陷,但该方案是针对数据交易的数据共享,缺少对用户的身份认证,匿名性会导致恶意行为无法被追溯和定责,不适用于敏感数据共享。Jin 等^[32]提出了一种基于区块链的安全、隐私保护和可互操作的医疗数据共享方案,该方案将链下个人数据和非个人身份信息存储在区块链中,该系统能够成功地跟踪利益相关者之间的个人信息移动,使追踪资料泄露来源更容易,但是区块链是一个透明化的系统,将数据存储在可能会带来数据泄露问题。Lath 等^[33]提出了一种智能交通系统中用于安全通信的数据交换和存储框架,使用区块链技术来防止数据被篡改、拦截、遗漏,在多种攻击下被证明具备稳健性。Lu 等^[34]提出了一种基于区块链和联邦学习的工业物联网数据共享方案,该方案利用联邦学习构建数据模型并共享数据模型而不是原始数据,从而将数据共享问题转化为机器学习问题,然后提出了新的区块链赋能的协作架构,在分布式的多方之间共享数据,同时与差分隐私技术结合,以减少数据泄露的风险。

在基于零知识证明、不经意传输、秘密分享等技术的解决方案中,Lim 等^[35]基于密钥同态加密的不经意伪随机函数和零知识证明提出了一种隐私保护的数据集集成和共享方案,该方案在数据集共享和集成过程中,不会泄露任何个人的身份信息,它可在不向第三方透露任何敏感资料的情况下,核实隐私保护数据集的正确性,它也不需要为了共享和集成不同的数据集而在参与者之

间共享密钥,这使得密钥管理更加简单,但是该方案缺少访问授权机制,任何合法用户都能获取相关数据集,数据拥有者无法对自己的数据进行访问控制。Huang^[36]提出了一种可定责和高效的工业物联网数据共享方案,该方案利用不经意传输协议和零知识证明技术对用户隐私进行保护,利用公钥同态加密、置换和随机化技术提高问责有效性,利用零知识证明防止数据提供者的恶意行为,但是该方案同样缺少访问授权机制。Mohanty 等^[37]提出了一种基于时间锁合约的隐私保护的社会车联网数据共享方案,该方案利用支付通道网络减少交易确认时间和链上开销,以增加交易的吞吐量,同时,引入匿名惩罚机制对上传错误数据的恶意用户进行匿名惩罚,但是该方案是针对数据交易的数据共享,缺少对用户的身份认证,不适用于敏感数据共享。Shi 等^[38]提出了一种物联网环境下的三重信任架构,确保从物联网边缘设备收集的数据不能被修改,以及分布式身份管理机制,增强个人隐私、安全和数字身份的控制。Gupta 等^[39]提出了一种基于秘密共享技术的系统,其名为 OBSCURE,使用秘密共享技术进行聚合查询,并具有通信高效和信息论安全的特点,可以防止网络和服务器了解用户的数据库查询内容或查询结果。Xu 等^[40]基于动态可搜索加密技术提出了一种云辅助医疗系统的可搜索和隐私保护的数据共享方案,该方案采用隐私保护的平等测试协议实现了关键字范围搜索和多关键字的搜索。

当前,隐私计算理论与关键技术^[3]给出了泛在共享环境下一种全生命周期保护的解决方案,其他相关研究更多停留在节点间的信息流转与交换,还有待进一步拓展与融合,对复杂应用场景形成更为个性化的补充与支撑。

2 多方隐私信息共享控制

2.1 隐私信息本地控制

联邦学习是利用本地算力对本地数据进行部分模型训练的一种分布式模型训练架构,其本质是原始数据不出域、算力分布利用。根据参与方之间数据分布的不同,可以把联邦学习分为 3 类:横向联邦学习、纵向联邦学习和联邦迁移学习。横向联邦学习的本质是样本的联合,适用于参与

者间业务相同但客户不同,即特征重叠多,用户重叠少时的场景,如不同地区的银行间,它们的业务相似(特征相似),但用户不同(样本不同)。纵向联邦学习的本质是特征的联合,适用于用户重叠多,特征重叠少的场景,如同一地区的商超和银行。而当参与者间特征和样本重叠都很少时可以考虑使用联邦迁移学习,如不同地区的银行和商超间的联合。联邦学习系统通常由一个负责聚合的中央服务器和多个参与训练的客户端组成。McMahan等^[41-43]提出了联邦学习的基本框架,其核心理念是数据不动模型动,数据可用不可见。在该框架下,各数据中心通过共享本地梯度以完成学习任务,能够在不泄露本地训练样本的前提下实现高精度的全局模型训练。最近的改进集中在克服联邦学习中的统计挑战^[44-45]和提高安全性^[46-47]方面。也有部分研究工作旨在使联邦学习更加个性化^[48-50]。

2.2 隐私信息共享的延伸控制

当前延伸控制研究主要聚焦于两个方面:基于起源的访问控制和策略粘贴技术。其中,基于起源的访问控制将起源数据作为决策依据,保护起源敏感资源;策略粘贴采用密码学技术将访问控制策略绑定到数据中,确保策略随数据流转。

在基于起源的访问控制方面,Nguren等^[51]做出了一系列开创性工作,给出了形式化策略规约语言和模型、动态权责分离方案和实施框架。Sun等^[52]设计了一个具有分层架构的源位置感知访问控制框架,能够自动感知信息的来源,实现了将信息的传播路径与访问控制策略结合。Bates等^[53]在Linux系统中创建了一个可信的来源感知执行环境,并基于此实现了一个文件传输应用程序还原出文件的传播路径,阻止来自敏感节点的文件传输任务。

在策略粘贴技术方面,Karjoth等^[54]提出了一种称为“黏性政策”(sticky policy)的跨系统访问控制方法,将访问控制策略绑定到原始数据上,从而解决跨系统传播问题。此后,此类方案在隐私保护的跨云服务中流行,Pearson等^[55]提出了一种实用的跨云隐私保护框架,通过将隐私数据与安全策略绑定,实现隐私数据的跨云控制。随后,Gnesi等^[56]提出了多对象个人数据的

隐私保护,Such等^[57-58]也在多人照片中进行隐私保护处理,并还考虑了其他涉及多人隐私数据中的策略冲突消解方法。Humbert等^[59-60]相继提出了针对亲人之间染色体数据和多人之间位置数据的隐私保护方案。Olteanu等^[61]提出了在线交互隐私确认系统发布隐私图片时,如果当事人在线,则当事人可以决定是否允许该图片被共享,如果当事人不在线,则隐私图片会被隐匿处理。文献[62-63]使用加密机制将策略与数据相关联,并对策略进行属性编码和匿名化处理以防被恶意利用,为全生命周期内的数据访问控制提供支持。但这些工作未限定使用哪些要素制定控制决策,未考虑审计信息,不能做到对数据非授权操作的溯源。为此,李凤华等^[64-65]提出了一种跨社交网络的图片隐私延伸控制方案和溯源取证方案,利用基于传播链的访问控制模型限制后续用户的操作权限,并将隐私策略嵌入图片文件,通过图片加密算法保护图片隐私信息和隐私策略的机密性、完整性,确保用户隐私策略被正确执行;同时,溯源取证方案使用嵌套签名对信息溯源记录的安全性进行保障,基于溯源记录能够还原出信息流转时的传播路径。

2.3 隐私脱敏算法及其控制

在基于匿名的隐私脱敏方面:基于匿名的隐私脱敏主要包含了 k -匿名^[66]、 l -多样性^[67]和 t -接近性^[68]等匿名化技术,其中 l -多样性和 t -接近性是 k -匿名的衍生技术,对于指定敏感属性,通过对其种类的多样化,以及确保其在各个匿名的等价类分组中的分布与全体记录分布保持邻近关系来提高匿名效果。基于匿名的脱敏算法应用对比分析如表2所示,类似地,周长利等^[69]设计 p -敏感性,保证关键属性至少有 p 个不同的值。泛化的基本思想是将轨迹上所有的采样点都泛化为对应的匿名区域,使攻击者无法获得准确位置。Wang等^[70]提出一种实现轨迹 k -匿名隐私保护的位置重组机制,该机制引入概率相似性和地理相似性来合成满足基本轨迹和采样轨迹隐私保护要求的伪轨迹。为进一步提高轨迹数据效用,Tojiboev等^[71]提出一种在矢量网格环境中添加噪声轨迹的匿名轨迹数据库发布模型,该方案不仅保证了匿名数据库的数据利用率,而且具有较低的

表2 基于匿名的脱敏算法应用对比分析

Table 2 Comparison of application of anonymous-based desensitization algorithms

文献	基于算法	场景	思想
[69]	p -敏感性	位置隐私	确保关键属性至少有 p 个不同的值
[70]	k -匿名	轨迹隐私	使用概率相似性和地理相似性合成伪轨迹,使每个轨迹满足 k -匿名
[71]	添加噪声轨迹	轨迹隐私	在矢量网格环境中添加噪声轨迹
[72]	个性化轨迹隐私保护	轨迹隐私	重构轨迹以提高数据的语义一致性和形状相似性
[73]	三维网格划分	轨迹隐私	划分三维网格,并对位置进行交换或删除
[74]	保留原始轨迹语义	轨迹隐私	时空扰动自适应地保留轨迹语义
[75]	动态群组隐私保护	数据融合	云服务器周期性收集数据,并计算聚合函数

时间复杂度。抑制的基本思想是在轨迹数据正式发布之前,剔除或删除现有轨迹中用户高频率访问位置或一些敏感位置。Dai等^[72]提出一种个性化轨迹隐私保护方法,通过合理替换敏感停止点后的轨迹重构,尽可能地提高轨迹数据的语义一致性和形状相似性,从而较好地平衡隐私和数据可用性。Chen等^[73]提出一种基于三维网格划分的轨迹隐私保护方法,在每个三维单元内进行轨迹间的位置交换或删除不满足条件的位置,而非整个轨迹,有效减少了轨迹匿名过程中的信息丢失。为进一步保留原始轨迹的语义,Naghizade等^[74]提出一种时空扰动自适应地保留轨迹语义的算法,尽可能保留原始轨迹的同时保护了中间敏感的停止点。Shen等^[75]提出了一种面向雾计算环境下动态群组的隐私保护和抗合谋的数据融合方案,在该协议中,云服务器可以周期性地收集原始数据并在其上计算任意的聚合函数,实现了终端设备的动态加入和退出。

在基于差分的隐私脱敏方面:Dwork^[76]提出了一种不依赖于攻击者背景知识的差分隐私技术。对于差分隐私的研究目前主要集中在两个方向:集中式数据模型和本地化数据模型。其中,本地化差分隐私由Duchi等^[77]提出,目前在苹果

和Google的数据收集算法中都使用了本地化差分隐私技术。苹果的iOS 10系统以及MacOS中部署了差分隐私算法^[78],用于收集经过保护的用户输入数据来分析用户习惯。Google在客户端Chrome中部署了RAPPOR算法^[79],用于收集隐私保护的用户隐私数据。基于扰动的差分隐私保护方法计算复杂度较低,但提高了在大数据挖掘和机器学习领域的应用效率。基于差分的脱敏算法应用对比分析如表3所示,Shokri等^[80]最先利用差分隐私机制设计了一种隐私保护的分布式学习方法,在这一方法中隐私损失是根据模型的参数计算得到,但模型参数数目庞大,因此会导致较大的隐私损失。在此基础上,Abadi等^[81]引入更高效的基于差分隐私的梯度下降算法和一种隐私损失记录方式,对隐私损失进行自动化计算。Wei等^[82]提出一种基于差分隐私概念的联邦学习框架,在模型聚合前向客户端的参数中加入满足差分隐私的噪声。考虑到单纯应用安全多方计算易受隐私推理攻击,单纯应用差分隐私在各客户端数据量相对较小的情况下会导致精度较低。因此,Truex等^[83]提出一种替代方法,将差分隐私与安全多方计算相结合,以此平衡隐私性与准确性。

表3 基于差分的脱敏算法应用对比分析

Table 3 Comparison of application of differential-based desensitization algorithms

文献	基于算法	场景	思想
[80]	差分隐私分布式学习	分布式学习	利用差分隐私机制设计隐私保护的分布式学习方法,通过模型参数计算隐私损失
[81]	差分隐私梯度下降算法	深度学习	基于差分隐私的梯度下降算法,并使用隐私损失记录方式对隐私损失进行自动化计算
[82]	差分隐私联邦学习	联邦学习	基于差分隐私的联邦学习框架,在模型聚合前向客户端的参数中加入噪声
[83]	差分隐私、安全多方计算	联邦学习	将差分隐私与安全多方计算相结合

基于信息论的脱敏算法应用对比分析如表4所示。在位置隐私保护场景中，Shaham等^[84]结合基于用户连续位置变化的新背景信息，提出一种新的度量标准即过渡熵用于研究位置隐私保护，然后提出了两种算法来改善产生虚假位置算法的过渡熵。Niu等^[85]提出一种融合信息缓存与信息熵的隐私度量方法，并基于该度量方法设计了信息缓存的虚假位置选择算法。朱素霞等^[86]提出基于相对熵和 K -means的形状相似差分隐私轨迹保护机制，利用相对熵计算真实位置对隐私预算的敏感程度，设计位置敏感的隐私级别实时计算算法。在位置服务中所涉及的隐私信息不但包括用户的真实位置，用户的查询内容常常也会泄露一些敏感的信息。Wang等^[87]提出两种算法MEE和MER，通过为查询兴趣设计复杂的分区策略，利用动态编程技术分别实现最大期望熵和最小最大比率。在数据发布场景中，Wu等^[88]通过刻画有相关性数据对隐私的威胁，提出一种新的差分隐私保护概念，当数据独立时，差分隐私模型是该模型的一个特殊情况；当数据之间有相关性时，采用群体隐私保护方案实现依赖情形下的差分隐私，从而实现新的隐私模型。彭长根等^[89]将隐私保护系统描述成一种通信模型，该模型利用信息熵描述隐私信息在数据信宿中的平均隐私信息量，通过隐私信息的不确定性直接反映隐私信息泄露的风险。

当前，对隐私信息的共享控制主要聚焦在域内，通过设置访问控制策略、粘贴特殊策略等技术，并结合不同场景和需求下的脱敏算法，实现隐私信息的细粒度共享控制。对适用于频繁跨域情况下需多次脱敏应用场景的解决方案还有待进一步研究。

3 多方数据协同安全计算

多方安全计算是指多方之间利用密码技术进行联合计算，本质是原始数据不出域、结果安全交换，支撑计算结果安全共享。在具体应用方面，Zhu和Wang等^[90-91]将安全两方余弦相似度计算协议引入基于位置服务中，构造了高效隐私保护的空間范围查询算法，能够在保证空間查询信息和待查询坐标数据安全的前提下实现空間范围查询，并进一步将提出的算法应用在社交网络和在線打车系統中。在机器学习领域，Zhu等^[92-93]将上述安全两方余弦相似度计算协议应用到线性SVM和神经网络的数据分类过程中，构造了系列数据安全分类方案，能够在保证模型数据安全和用户待预测数据敏感性的前提下进行高效数据分类。Mohassel等^[94]构造了一个经典的两方计算协议，数据所有者将隐私数据分发给两个服务器，服务器通过协同计算对用户数据进行模型训练。在该协议中，每个服务器仅能得到用户的部分数据，因此，每个用户的敏感信息能够得到有效的保护。Wang等^[95]提出了一种基于安全多方向量比较算法的协作式模型学习方案，以此来解决医疗数据共享的隐私问题，提高在线医疗诊断服务的准确性。通过该方案，医疗中心可以在云的帮助下安全地学习全局诊断模型和本地诊断模型，并且每个医疗中心的敏感医疗数据都得到了很好的保护。Kanagavelu等^[96]提出了基于安全多方计算的模型聚合框架，该框架以点对点的方式进行模型聚合，同时通过建立分段机制来减少模型聚合过程中的通信开销，提高了框架的可扩展性。

在同态加密方面，支持对密文直接进行运

表4 基于信息论的脱敏算法应用对比分析

Table 4 Comparison of application of information theory-based desensitization algorithms

文献	基于算法	场景	思想
[84]	过渡熵	位置隐私	结合用户连续位置变化的新背景信息，提出过渡熵指标
[85]	信息缓存、信息熵	位置隐私	融合信息缓存与信息熵，设计基于信息缓存的虚假位置选择算法
[86]	相对熵、 K -means、差分隐私	位置隐私	通过相对熵计算真实位置对隐私预算的敏感程度，并设计位置敏感的隐私级别实时计算算法
[87]	MEE、MER	查询隐私	通过设计复杂的分区策略，利用动态编程技术实现最大期望熵和最小、最大比率
[88]	差分隐私模型	数据发布	通过刻画有相关性数据对隐私的威胁，提出一种新的差分隐私保护模型
[89]	通信模型	数据发布	利用信息熵反映隐私信息泄露的风险

算, 所得到的结果仍是密文, 且与对应明文运算后再加密的结果相同。利用上述同态特性, 用户能够将数据计算从明文域转化到密文域中以保护敏感数据安全。目前出现的同态加密算法可以被分为 3 种类型: 部分同态加密、全同态加密和对称同态加密。在具体应用方面, 同态加密支持在密文域中直接对数据进行处理, 在将数据外包至服务器的互联网服务中得到了广泛应用。在基于位置服务中, Yi 等^[97]基于 Paillier 同态加密提出了面向移动用户的 k 近邻范围查询方案, 能够保证查询过程中用户相关位置数据的安全性。Mu 等^[98]将 Paillier 加密算法和 ElGamal 算法进行了结合, 提出了任意形状的多边形范围查询方案, 支持社交网络中好友位置的模糊查询。Zhu 等^[99]提出了一种隐私保护的隐私中心位置探测方案, 服务器可计算一个用户集合的位置中心而不获取各用户的精确位置。为提升密文空间范围查询的效率, Khoshgozaran 等^[100]利用四叉树数据结构将空间进行了划分, 并通过希尔伯特填充曲线对服务空间进行降维处理, 极大地提升了密文空间查询效率。Thomas 等^[101]利用点和多边形顶点之间的角度关系, 提出了一种新型的多边形范围查询策略, 并借助加法同态实现了安全的多边形范围查询。在外包学习中, Gilad-Bachrach 等^[102]设计了可以在加密数据上运行的神经网络 Cryptonets, 其主要原理是用户上传数据之前先对数据进行同态加密, 然后对模型进行多项式逼近处理, 使之能够在加密后的数据上进行模型推断。Nandakumar 等^[103]提出了更为高效的全同态加密网络, 能够完成模型在加密数据上的训练任务。但该方案仅能应用于结构较为简单的神经网络, 无法应用于真实场景数百层的复杂网络。崔建京等^[104]总结了同态加密在各种机器学习算法中的应用研究进展, 主要包括神经网络、 k 近邻、决策树和随机森林 4 种机器学习算法, 并分析了其正确性、安全性和执行效率。王启正^[105]利用神经网络对同态加密后的隐私图像进行特征提取, 且保证在多项式时间内完成计算。

尽管安全多方计算和同态加密在数据安全性方面具有优势, 但作为传统的数据安全技术, 无法解决数据跨域交换场景下的隐私保护问题, 且

它们在性能上仍面临重大瓶颈。安全多方计算的主要问题是高昂的通信开销, 参与方之间需要频繁地交换大量数据, 在网络延迟高和带宽有限的情况下问题尤其显著。大量的加解密操作及布尔电路或算术电路的评估也会使计算成本急剧上升, 在大部分需要多轮交互的场景下, 计算时间和通信成本还会进一步增加。同态加密不仅涉及大量的模运算和多项式运算, 而且需要存储的密文量比明文大得多, 具有极高的计算和存储成本, 尤其是在涉及多次同态运算的情况下, 代价呈指数级增长, 使计算开销和存储需求急剧增加。这些性能瓶颈严重限制了其在大规模数据处理和复杂计算场景中的应用。

4 结束语

传统隐私脱敏方案通常采用匿名、加扰、访问控制等技术实现隐私保护, 但更多侧重特定环节或单一信息系统, 在复杂交换环境下的安全性无法得到有效保障, 且脱敏后的隐私信息无法完全恢复, 势必造成可用性的降低; 安全多方计算、同态加密等基于密码学的解决方案计算开销大、交互轮次多, 在跨域和大规模数据运算场景并不适合; 联邦学习应用场景虽然可实现数据不出域, 但是计算程序和模型需要部署到各个数据所有方, 会暴露模型的隐私, 也不适合模型需要保护的场景。

隐私计算理论已给出面向全生命周期保护的计算框架和信息系统框架, 但后续还需结合不同应用场景, 基于隐私计算理论体系, 践行隐私信息全生命周期保护。

参考文献:

- [1] DINEV T, HART P. An extended privacy calculus model for E-commerce transactions[J]. Information Systems Research, 2006, 17(1): 61-80.
- [2] 李风华, 李晖, 贾焰, 等. 隐私计算研究范畴及发展趋势[J]. 通信学报, 2016, 37(4): 1-11.
LI F H, LI H, JIA Y, et al. Privacy computing: concept, connotation and its research trend[J]. Journal on Communications, 2016, 37(4): 1-11.
- [3] 李风华, 李晖, 牛犇. 隐私计算理论与技术[M]. 北京: 人民邮电出版社, 2021.
LI F H, LI H, NIU B. Privacy computing theory and technology[M]. Beijing: Posts & Telecom Press, 2021.

- [4] KOO J, KANG G, KIM Y G. Security and privacy in big data life cycle: a survey and open challenges[J]. *Sustainability*, 2020, 12(24): 10571.
- [5] ANCIAUX N L G, HEERDE H J, et al. Implanting life-cycle privacy policies in a context database[M]. Enschede: University of Twente, 2006.
- [6] AMBROSE M L. It's about time: privacy, information life cycles, and the right to be forgotten[J]. *Stan Tech L Rev*, 2012, 16: 369.
- [7] LIN L, LIU T T, HU J, et al. A privacy-aware cloud service selection method toward data life-cycle[C]//Proceedings of the 2014 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS). Piscataway: IEEE Press, 2014: 752-759.
- [8] KANG Y, LEE H, CHUN K, et al. Classification of privacy enhancing technologies on life-cycle of information[C]//Proceedings of the International Conference on Emerging Security Information, Systems, and Technologies (SECUREWARE 2007). Piscataway: IEEE Press, 2007: 66-70.
- [9] 张文静, 刘樵, 朱辉. 基于信息论方法的多等级位置隐私度量与保护[J]. *通信学报*, 2019, 40(12): 51-59.
- ZHANG W J, LIU Q, ZHU H. Evaluation and protection of multi-level location privacy based on an information theoretic approach[J]. *Journal on Communications*, 2019, 40(12): 51-59.
- [10] 俞艺涵, 付钰, 吴晓平. 基于 Shannon 信息熵与 BP 神经网络的隐私数据度量与分级模型[J]. *通信学报*, 2018, 39(12): 10-17.
- YU Y H, FU Y, WU X P. Metric and classification model for privacy data based on Shannon information entropy and BP neural network[J]. *Journal on Communications*, 2018, 39(12): 10-17.
- [11] 俞艺涵, 付钰, 吴晓平. 基于多层模糊综合评估的隐私保护效果评估方法[J]. *网络与信息安全学报*, 2020, 6(6): 121-127.
- YU Y H, FU Y, WU X P. Evaluation method of privacy protection effect based on multi-layer fuzzy comprehensive evaluation[J]. *Chinese Journal of Network and Information Security*, 2020, 6(6): 121-127.
- [12] SANDHU R S. Role-based access control[M]//Advances in Computers. Amsterdam: Elsevier, 1998: 237-286.
- [13] SANDHU R S, COYNE E J, FEINSTEIN H L, et al. Role-based access control models[J]. *Computer*, 1996, 29(2): 38-47.
- [14] COVINGTON M J, LONG W D, SRINIVASAN S, et al. Securing context-aware applications using environment roles[C]//Proceedings of the Sixth ACM Symposium on Access Control Models and Technologies. New York: ACM, 2001: 10-20.
- [15] HU V C, FERRAILOLO D, KUHN R, et al. Guide to attribute based access control (ABAC) definition and considerations (draft) [J]. NIST Special Publication, 2013, 800(162): 1-54.
- [16] LI F H, LI Z F, HAN W L, et al. Cyberspace-oriented access control: a cyberspace characteristics-based model and its policies[J]. *IEEE Internet of Things Journal*, 2019, 6(2): 1471-1483.
- [17] 李风华, 王彦超, 殷丽华, 等. 面向网络空间的访问控制模型[J]. *通信学报*, 2016, 37(5): 9-20.
- LI F H, WANG Y C, YIN L H, et al. Novel cyberspace-oriented access control model[J]. *Journal on Communications*, 2016, 37(5): 9-20.
- [18] KOREN Y, BELL R, VOLINSKY C. Matrix factorization techniques for recommender systems[J]. *Computer*, 2009, 42(8): 30-37.
- [19] HU Y F, KOREN Y, VOLINSKY C. Collaborative filtering for implicit feedback datasets[C]//Proceedings of the 2008 Eighth IEEE International Conference on Data Mining. Piscataway: IEEE Press, 2008: 263-272.
- [20] MA H, ZHOU D Y, LIU C, et al. Recommender systems with social regularization[C]//Proceedings of the Fourth ACM International Conference on Web Search and Data Mining. New York: ACM, 2011: 287-296.
- [21] MA H. An experimental study on implicit social recommendation [C]//Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2013: 73-82.
- [22] PORTEOUS I, ASUNCION A, WELLING M. Bayesian matrix factorization with side information and dirichlet process mixtures[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2010, 24(1): 563-568.
- [23] GUPTA K, RAGHUPRASAD M Y, KUMAR P. A hybrid variational autoencoder for collaborative filtering[EB]. arXiv Preprint arXiv: 1808.01006, 2018.
- [24] CHENG H T, KOC L, HARMSSEN J, et al. Wide & deep learning for recommender systems[C]//Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. New York: ACM, 2016: 7-10.
- [25] XUE H J, DAI X, ZHANG J, et al. Deep matrix factorization models for recommender systems[C]//IJCAI. 2017: 3203-3209.
- [26] KIM D, PARK C, OH J, et al. Convolutional matrix factorization for document context-aware recommendation[C]//Proceedings of the 10th ACM Conference on Recommender Systems. New York: ACM, 2016: 233-240.
- [27] ZHENG L, NOROOZI V, YU P S. Joint deep modeling of users and items using reviews for recommendation[C]//Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. New York: ACM, 2017: 425-434.
- [28] WU S, TANG Y Y, ZHU Y Q, et al. Session-based recommendation with graph neural networks[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33(1): 346-353.
- [29] HUANG H, ZHANG Q, HUANG X, et al. Mention recommendation for twitter with end-to-end memory network[C]//IJCAI. 2017: 1872-1878.
- [30] WANG Y T, SU Z, ZHANG N, et al. SPDS: a secure and auditable private data sharing scheme for smart grid based on blockchain[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(11): 7688-7699.
- [31] LI T, WANG H Q, HE D B, et al. Blockchain-based privacy-preserving and rewarding private data sharing for IoT[J]. *IEEE Internet of Things Journal*, 2022, 9(16): 15138-15149.
- [32] JIN H, XU C, LUO Y, et al. Toward secure, privacy-preserving, and interoperable medical data sharing via blockchain[C]//Proceedings of the 2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS). Piscataway: IEEE Press, 2019: 852-861.
- [33] LATH N, THAPLIYAL K, KANDPAL K, et al. BDESf-ITS: blockchain-based secure data exchange and storage framework for intelligent transportation system[C]//Proceedings of the IEEE INFOCOM 2022 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Piscataway: IEEE Press,

- 2022: 1-6.
- [34] LU Y L, HUANG X H, DAI Y Y, et al. Blockchain and federated learning for privacy-preserved data sharing in industrial IoT[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(6): 4177-4186.
- [35] LIM H W, POH G S, XU J, et al. $\mathsf{PrivateLink}$: privacy-preserving integration and sharing of datasets[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 564-577.
- [36] HUANG C, LIU D X, NI J B, et al. Achieving accountable and efficient data sharing in industrial internet of things[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(2): 1416-1427.
- [37] MOHANTY S K, TRIPATHY S. *SloVChain*: time-lock contract based privacy-preserving data sharing in *SloV*[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(12): 24071-24082.
- [38] SHI P C, WANG H M, YANG S Z, et al. Blockchain-based trusted data sharing among trusted stakeholders in IoT[J]. *Software: Practice and Experience*, 2021, 51(10): 2051-2064.
- [39] GUPTA P, LI Y, MEHROTRA S, et al. *Obscure*: information-theoretically secure, oblivious, and verifiable aggregation queries on secret-shared outsourced data[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2022, 34(2): 843-864.
- [40] XU C, WANG N N, ZHU L H, et al. Achieving searchable and privacy-preserving data sharing for cloud-assisted E-healthcare system[J]. *IEEE Internet of Things Journal*, 2019, 6(5): 8345-8356.
- [41] MCMAHAN H B, MOORE E, RAMAGE D, et al. Communication-efficient learning of deep networks from decentralized data[EB]. *arXiv Preprint arXiv: 1602.05629*, 2016.
- [42] KONEČNÝ J, MCMAHAN H B, RAMAGE D, et al. Federated optimization: distributed machine learning for on-device intelligence [EB]. *arXiv Preprint arXiv: 1610.02527*, 2016.
- [43] KONEČNÝ J, MCMAHAN H B, YU F X, et al. Federated learning: strategies for improving communication efficiency[EB]. *arXiv Preprint arXiv: 1610.05492*, 2016.
- [44] SMITH V, CHIANG C K, SANJABI M, et al. Federated multi-task learning[J]. *Advances in Neural Information Processing Systems*, 2017, 30.
- [45] ZHAO Y, LI M, LAI L Z, et al. Federated learning with non-IID data[EB]. *arXiv Preprint arXiv: 1806.00582*, 2018.
- [46] BONAWITZ K, IVANOV V, KREUTER B, et al. Practical secure aggregation for privacy-preserving machine learning[C]//*Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*. New York: ACM, 2017: 1175-1191.
- [47] GEYER R C, KLEIN T, NABI M. Differentially private federated learning: a client level perspective[EB]. *arXiv Preprint arXiv: 1712.07557*, 2017.
- [48] CHEN F, LUO M, DONG Z H, et al. Federated meta-learning with fast convergence and efficient communication[EB]. *arXiv Preprint arXiv: 1802.07876*, 2018.
- [49] SMITH V, CHIANG C K, SANJABI M, et al. Federated multi-task learning[J]. *Advances in Neural Information Processing Systems*, 2017, 30.
- [50] LIANG P P, LIU T, LIU Z Y, et al. Think locally, act globally: federated learning with local and global representations[EB]. *arXiv Preprint arXiv: 2001.01523*, 2020.
- [51] NGUYEN D, PARK J, SANDHU R. A provenance-based access control model for dynamic separation of duties[C]//*Proceedings of the 2013 Eleventh Annual Conference on Privacy, Security and Trust*. Piscataway: IEEE Press, 2013: 247-256.
- [52] SUN L S, PARK J, NGUYEN D, et al. A provenance-aware access control framework with typed provenance[J]. *IEEE Transactions on Dependable and Secure Computing*, 2016, 13(4): 411-423.
- [53] BATES A, TIAN D, BUTLER K R B, et al. Trustworthy whole-system provenance for the Linux kernel[J]. *Proceedings of the 24th USENIX Security Symposium*, 2015: 319-334.
- [54] KARJOTH G, SCHUNTER M, WADNER M. Platform for enterprise privacy practices: privacy-enabled management of customer data[M]. *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003: 69-84.
- [55] PEARSON S, CASASSA-MONT M. Sticky policies: an approach for managing privacy across multiple parties[J]. *Computer*, 2011, 44(9): 60-68.
- [56] GNESI S, MATTEUCCI I, MOISO C, et al. My data, your data, our data: managing privacy preferences in multiple subjects personal data[M]. *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2014: 154-171.
- [57] SUCH J M, PORTER J, PREIBUSCH S, et al. Photo privacy conflicts in social media: a large-scale empirical study[C]//*Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. New York: ACM, 2017: 3821-3832.
- [58] SUCH J M, CRIADO N. Resolving multi-party privacy conflicts in social media[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(7): 1851-1863.
- [59] HUMBERT M, AYDAY E, HUBAUX J P, et al. Quantifying interdependent risks in genomic privacy[J]. *ACM Transactions on Privacy and Security*, 2017, 20(1): 1-31.
- [60] OLTEANU A M, HUGUENIN K, SHOKRI R, et al. Quantifying interdependent privacy risks with location data[J]. *IEEE Transactions on Mobile Computing*, 2017, 16(3): 829-842.
- [61] OLTEANU A M, HUGUENIN K, DACOSTA I, et al. Consensual and privacy-preserving sharing of multi-subject and interdependent data[C]//*Proceedings 2018 Network and Distributed System Security Symposium*. 2018: 1-16.
- [62] SPYRA G, BUCHANAN W J, EKONOMOU E. Sticky policies approach within cloud computing[J]. *Computers & Security*, 2017, 70: 366-375.
- [63] PEARSON S, MONT M C, KOUNGA G. Enhancing accountability in the cloud via sticky policies[M]//*Communications in Computer and Information Science*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 146-155.
- [64] LI F H, SUN Z, NIU B, et al. An extended control framework for privacy-preserving photo sharing across different social networks [C]//*Proceedings of the 2019 International Conference on Computing, Networking and Communications (ICNC)*. Piscataway: IEEE Press, 2019: 390-394.
- [65] 李凤华, 孙哲, 牛犇, 等. 跨社交网络的隐私图片分享框架[J]. *通信学报*, 2019, 40(7): 1-13.
- LI F H, SUN Z, NIU B, et al. Privacy-preserving photo sharing framework cross different social network[J]. *Journal on Communications*, 2019, 40(7): 1-13.

- [66] SWEENEY L. K-anonymity: a model for protecting privacy[J]. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2002, 10(5): 557-570.
- [67] MACHANAVAJJHALA A, GEHRKE J, KIFER D, et al. L-diversity: privacy beyond k-anonymity[C]//*Proceedings of the 22nd International Conference on Data Engineering (ICDE'06)*. Piscataway: IEEE Press, 2006: 24.
- [68] LI N H, LI T C, VENKATASUBRAMANIAN S. T-closeness: privacy beyond k-anonymity and l-diversity[C]//*Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering*. Piscataway: IEEE Press, 2007: 106-115.
- [69] 周长利, 马春光, 杨松涛. 基于敏感位置多样性的LBS位置隐私保护方法研究[J]. *通信学报*, 2015, 36(4): 125-136.
- ZHOU C L, MA C G, YANG S T. Research of LBS privacy preserving based on sensitive location diversity[J]. *Journal on Communications*, 2015, 36(4): 125-136.
- [70] WANG Y F, LI M Z, LUO S S, et al. LRM: a location recombination mechanism for achieving trajectory: anonymity privacy protection[J]. *IEEE Access*, 2019, 7: 182886-182905.
- [71] TOJIBOEV R, LEE W, LEE C C. Adding noise trajectory for providing privacy in data publishing by vectorization[C]//*Proceedings of the 2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*. Piscataway: IEEE Press, 2020: 432-434.
- [72] DAI Y, SHAO J, WEI C B, et al. Personalized semantic trajectory privacy preservation through trajectory reconstruction[J]. *World Wide Web*, 2018, 21(4): 875-914.
- [73] CHEN C M, LUO Y L, YU Q Y, et al. TPPG: privacy-preserving trajectory data publication based on 3D-Grid partition[J]. *Intelligent Data Analysis*, 2019, 23(3): 503-533.
- [74] NAGHIZADE E, KULIK L, TANIN E, et al. Privacy- and context-aware release of trajectory data[J]. *ACM Transactions on Spatial Algorithms and Systems*, 2020, 6(1): 1-25.
- [75] SHEN X D, ZHU L H, XU C, et al. A privacy-preserving data aggregation scheme for dynamic groups in fog computing[J]. *Information Sciences*, 2020, 514: 118-130.
- [76] DWORK C. Differential privacy: a survey of results[C]//*International Conference on Theory and Applications of Models of Computation*. Berlin, Heidelberg: Springer, 2008: 1-19.
- [77] DUCHI J C, JORDAN M I, WAINWRIGHT M J. Local privacy and statistical minimax rates[C]//*Proceedings of the 2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. Piscataway: IEEE Press, 2013: 429-438.
- [78] TANG J, KOROLOVA A, BAI X L, et al. Privacy loss in apple's implementation of differential privacy on MacOS 10.12[EB]. *arXiv Preprint arXiv: 1709.02753*, 2017.
- [79] ERLINGSSON Ú, PIHUR V, KOROLOVA A. RAPPOR: randomized aggregatable privacy-preserving ordinal response[C]//*Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. New York: ACM, 2014: 1054-1067.
- [80] SHOKRI R, SHMATIKOV V. Privacy-preserving deep learning[C]//*Proceedings of the 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. Piscataway: IEEE Press, 2015: 909-910.
- [81] ABADI M, CHU A, GOODFELLOW I, et al. Deep learning with differential privacy[C]//*Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. New York: ACM, 2016: 308-318.
- [82] WEI K, LI J, DING M, et al. Federated learning with differential privacy: algorithms and performance analysis[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 3454-3469.
- [83] TRUEX S, BARACALDO N, ANWAR A, et al. A hybrid approach to privacy-preserving federated learning[J]. *Informatik Spektrum*, 2019, 42(5): 356-357.
- [84] SHAHAM S, DING M, LIU B, et al. Privacy preservation in location-based services: a novel metric and attack model[J]. *IEEE Transactions on Mobile Computing*, 2021, 20(10): 3006-3019.
- [85] NIU B, LI Q H, ZHU X Y, et al. Enhancing privacy through caching in location-based services[C]//*Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM)*. Piscataway: IEEE Press, 2015: 1017-1025.
- [86] 朱素霞, 刘抒伦, 孙广路. 基于相对熵和K-means的形状相似差分隐私轨迹保护机制[J]. *通信学报*, 2021, 42(2): 113-123.
- ZHU S X, LIU S L, SUN G L. Shape similarity differential privacy trajectory protection mechanism based on relative entropy and K-means[J]. *Journal on Communications*, 2021, 42(2): 113-123.
- [87] WANG J B, LI Y S, YANG D H, et al. Achieving effective k-anonymity for query privacy in location-based services[J]. *IEEE Access*, 2017, 5: 24580-24592.
- [88] WU G Q, XIA X Y, HE Y P. Achieving dalenius' goal of data privacy with practical assumptions[EB]. *arXiv Preprint arXiv: 1703.07474*, 2017.
- [89] 彭长根, 丁红发, 朱义杰, 等. 隐私保护的信息熵模型及其度量方法[J]. *软件学报*, 2016, 27(8): 1891-1903.
- PENG C G, DING H F, ZHU Y J, et al. Information entropy models and privacy metrics methods for privacy protection[J]. *Journal of Software*, 2016, 27(8): 1891-1903.
- [90] ZHU H, WANG F W, LU R X, et al. Efficient and privacy-preserving proximity detection schemes for social applications[J]. *IEEE Internet of Things Journal*, 2018, 5(4): 2947-2957.
- [91] WANG F W, ZHU H, LIU X M, et al. Efficient and privacy-preserving dynamic spatial query scheme for ride-hailing services[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(11): 11084-11097.
- [92] ZHU H, LIU X X, LU R X, et al. Efficient and privacy-preserving online medical prediagnosis framework using nonlinear SVM[J]. *IEEE Journal of Biomedical and Health Informatics*, 2017, 21(3): 838-850.
- [93] ZHU H, LIU X X, LU R X, et al. EPCS: an efficient and privacy-preserving classification service query framework for SVM[J]. *Security and Communication Networks*, 2016, 9(11): 1309-1320.
- [94] MOHASSEL P, ZHANG Y P. SecureML: a system for scalable privacy-preserving machine learning[C]//*Proceedings of the 2017 IEEE Symposium on Security and Privacy (SP)*. Piscataway: IEEE Press, 2017: 19-38.
- [95] WANG F W, ZHU H, LIU X M, et al. Privacy-preserving collaborative model learning scheme for E-healthcare[J]. *IEEE Access*, 2019, 7: 166054-166065.
- [96] KANAGAVELU R, LI Z X, SAMSUDIN J, et al. Two-phase multiparty computation enabled privacy-preserving federated learning[C]//*Proceedings of the 2020 20th IEEE/ACM International Sym-*

sium on Cluster, Cloud and Internet Computing (CCGRID). Piscataway: IEEE Press, 2020: 410-419.

- [97] YI X, PAULET R, BERTINO E, et al. Practical approximate k nearest neighbor queries with location and query privacy[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(6): 1546-1559.
- [98] MU B, BAKIRAS S. Private proximity detection for convex polygons[C]//Proceedings of the Tsinghua Science and Technology. TUP2016: 270-280.
- [99] ZHU X L, LU Y, ZHU X J, et al. A location privacy-preserving protocol based on homomorphic encryption and key agreement[C]//Proceedings of the 2013 International Conference on Information Science and Cloud Computing Companion. Piscataway: IEEE Press, 2013: 54-59.
- [100] KHOSHGOZARAN A, SHIRANI-MEHR H, SHAHABI C. Blind evaluation of location based queries using space transformation to preserve location privacy[J]. GeoInformatica, 2013, 17(4): 599-634.
- [101] THOMAS T. Secure two-party protocols for point inclusion problem[EB]. arXiv Preprint arXiv: 0705.4185, 2007.
- [102] GILAD-BACHRACH R, DOWLIN N, LAINE K, et al. Cryptonets: applying neural networks to encrypted data with high throughput and accuracy[C]//International Conference on Machine Learning. PMLR, 2016: 201-210.
- [103] NANDAKUMAR K, RATHA N, PANKANTI S, et al. Towards deep neural network training on encrypted data[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE Press, 2019: 40-48.
- [104] 崔建京, 龙军, 闵尔学, 等. 同态加密在加密机器学习中的应用研究综述[J]. 计算机科学, 2018, 45(4): 46-52.
CUI J J, LONG J, MIN E X, et al. Survey on application of homomorphic encryption in encrypted machine learning[J]. Computer Science, 2018, 45(4): 46-52.
- [105] 王启正. 隐私保护下的数据处理[D]. 济南: 山东师范大学, 2019.
WANG Q Z. Data processing under privacy protection[D]. Jinan: Shandong Normal University, 2019.

[作者简介]



刘圣龙 (1988-), 男, 辽宁沈阳人, 国家电网有限公司大数据中心高级工程师, 主要研究方向为网络与数据安全。



黄秀丽 (1979-), 女, 河南开封人, 国网智能电网研究院有限公司高级工程师, 主要研究方向为电力数据安全研究。



江伊雯 (1993-), 女, 安徽宣城人, 国家电网有限公司大数据中心高级工程师, 主要研究方向为数据安全。



姜嘉伟 (1994-), 男, 吉林吉林人, 国家电网有限公司大数据中心工程师, 主要研究方向为网络安全。



田月池 (1998-), 女, 河北保定人, 中国科学院信息工程研究所博士生, 主要研究方向为隐私计算、隐私保护效果评估。



周泽峻 (2000-), 男, 河南洛阳人, 中国科学院信息工程研究所博士生, 主要研究方向为隐私计算、数据安全。



牛犇 (1984-), 男, 陕西西安人, 博士, 中国科学院信息工程研究所研究员、博士生导师, 主要研究方向为数据安全、隐私计算。