

基于对称算法的多源异构能源数据快速脱敏方法

田 钺¹ 吴 漾¹ 董若烟¹ 孔庆波¹ 方继宇¹

TIAN Yue WU Yang DONG Ruoyan KONG Qingbo FANG Jiyu

摘 要

针对多源异构能源数据因构成差异与噪声干扰,导致脱敏处理时间复杂度高的问题,文章提出了一种基于对称算法的多源异构能源数据快速脱敏方法。利用 $L_{2,1}$ 范数量化数据初始化特性敏感性,构建考虑数据相似性的脱敏目标函数,约束变形参数分布以平滑数据差异、过滤噪声;通过对称算法迭代目标函数时调换仿射变化对与乘法逆元,增加有限域多项式复杂度实现脱敏。测试结果表明,该方法严格雪崩准则距离稳定在较低水平,脱敏效果可靠,且时间复杂度稳定在 0.6 ms 以内。

关键词

对称算法;多源异构能源数据;快速脱敏; $L_{2,1}$ 范数;变形参数分布

doi: 10.3969/j.issn.1672-9528.2025.11.041

0 引言

为保护数据的安全性和隐私性,需要对数据中的敏感部分进行处理,使其在不影响数据使用和分析的前提下,无法直接识别出原始的敏感信息^[1]。其中,数据脱敏是较为常见的技术手段之一。其通过对数据中的敏感信息进行基于脱敏规则的变形,实现对敏感隐私数据的可靠保护^[2]。

在相关研究中,高博等人^[3]聚焦于石油勘探数据脱敏与安全保护问题,针对石油勘探数据中敏感信息的特点,提出了一种基于 BERT-CRF 模型与数据加密相结合的方法。利用 BERT-CRF 模型对石油勘探文本数据进行命名实体识别,精准定位数据中的敏感实体。采用特定的数据加密算法对识别出的敏感实体进行加密处理,实现了数据的脱敏。该方法能够有效保护石油勘探数据中的敏感信息,在保证数据可用性的同时,提高了数据的安全性。但是敏感实体识别过程的时间开销较大,导致石油勘探数据脱敏的整体时间复杂度较高。窦亚珍^[4]致力于面向深度学习的动态数据脱敏技术研究,提出了一种动态数据脱敏框架。该框架根据深度学习模型的训练需求和数据特点,实时调整脱敏策略。该方法能够不影响深度学习模型性能的前提下,有效降低数据泄漏风险。但是动态数据脱敏框架需要大量数据作为训练,在实际应用场景中存在时间复杂度相对较高的局限。李思慧等人^[5]围绕基于差分隐私的数据脱敏技术展开深入探讨,分析了差分隐私的原理和实现方

法,并将其应用于数据脱敏过程中。通过在数据发布/共享时添加适量的噪声,在保证数据可用性的同时,有效保护了数据中个体的隐私信息。该方法能够在不同数据场景下实现良好的脱敏效果,但是噪声添加的合理性以及数量对于数据脱敏的时间复杂度影响较大。蔡彬彬^[6]提出了一种基于差分隐私的多维数据脱敏方法。针对多维数据的特点,设计了噪声添加机制,在保证数据多维结构信息不被破坏的前提下,实现对敏感数据的脱敏处理。但是数据多维结构信息不被破坏的约束限制了脱敏的效率。

相关研究在多源异构能源数据脱敏中,或因模型复杂导致时间开销大,或因约束条件限制效率,存在明显短板。本文聚焦上述问题,将对称算法与 $L_{2,1}$ 范数结合,构建考虑数据相似性的目标函数,通过调换仿射变化对与乘法逆元强化混淆效果,开展基于对称算法的多源异构能源数据快速脱敏方法研究。

1 多源异构能源数据快速脱敏方法设计

1.1 多源异构能源数据脱敏目标函数构建

对于多源异构能源数据而言,其自身在构成上的差异化特征,以及噪声和异常值含量都会影响对其进行脱敏处理的效果^[7]。本文结合对称算法对其进行脱敏处理时,设置了以 $L_{2,1}$ 范数度量为基础的损失函数,利用 $L_{2,1}$ 范数对多源异构能源数据初始化特性的敏感性量化,确保数据脱敏的效果^[8]。

对称算法对多源异构能源数据进行脱敏的目标函数可以表示为:

$$F = \min \sum_{r=1}^k (\alpha_r)^\gamma \|W - H_r H_r^T\|_{2,1} \quad (1)$$

对应的约束条件为:

1. 贵州电网有限责任公司信息中心 贵州贵阳 550081

[基金项目] 中国南方电网有限责任公司科技项目“能源大数据中心海量实时架构与多源异构数据集成融合技术研究”(GZKJXM20232513)

$$\text{s.t.} \begin{cases} H_r \geq 0 \\ \forall r, \\ \|\alpha\|_1 = 1 \\ \alpha \geq 1 \end{cases} \quad (2)$$

式中： F 表示对称算法对多源异构能源数据进行脱敏的目标函数； H_r 表示聚类数，即多源异构能源数据中每一行最大值对应位置的指示样本； r 表示多源异构能源数据中的元素数量； α_r 表示权重向量； α 表示对称算法对多源异构能源数据进行脱敏的变形参数，通过令 $\|\alpha\|_1$ 的值为1，约束 α 的解为非平凡值； W 表示用于度量多源异构能源数据中第*i*个和第*j*个样本之间相似性的重构矩阵； γ 表示超参数，利用其控制 α 项的分布，避免其解出现极端稀疏的情况，确保大部分 α 处于无限接近0的区间范围内。

按照上述方式，利用随机生成非负矩阵完成多源异构能源数据脱敏目标函数构建。该过程不仅能覆盖数据多源异构特性，还可借助聚类划分矩阵过滤噪声与异常值，大幅减少脱敏迭代次数，为后续利用对称算法迭代目标函数奠定基础。

1.2 对称算法迭代目标函数

结合多源异构能源数据脱敏目标函数，引入对称算法对其进行迭代，借助对称算法的非线性组件保障多源异构能源数据脱敏过程的混淆效果^[9]。

以有限域的数学方法为基础，本文考虑到利用有限域 $GF(2^8)$ 多项式表达式迭代目标函数的输出结果仅为9项，在复杂程度方面的表现相对较低，会降低多源异构能源数据的脱敏效果^[10]。针对此，从提升对称算法代数式项数的角度出发，将原始的仿射变化对和乘法逆元进行调换。其中，在有限域 $GF(2^8)$ 上对多源异构能源数据进行放射变化的处理方式可以表示为：

$$G_\alpha = \alpha(x)a(x) \quad (3)$$

式中： G_α 表示放射变化输出； $\alpha(x)$ 表示多源异构能源数据脱敏变形参数的解； $a(x)$ 表示变形参数解的仿射变化对。在构建多源异构能源数据脱敏目标函数时，约束 α 的分布范围为无限接近0的区间内。因此，为避免放射变化离散程度过高导致的 $a(x)$ 分布与 α 分布重叠较低问题，设置 $a(x)$ 的分布范围也为无限接近0的区间内。

为保障有限域 $GF(2^8)$ 的背景下，多源异构能源数据的字节元素均能匹配对应的乘法逆元，设置元素 $p(x)$ ，且满足 $p(x) \cdot q(x)$ 对有限域 $GF(2^8)$ 多项式表达式中的不可约多项式取模结果为单位元。结合这一设置，多源异构能源数据脱敏过程中乘法逆元的判定方式可以表示为：

$$p(x)q(x) \bmod 1 \notin GF(2^8) \quad (4)$$

对于满足式(4)的 $q(x)$ ，输出其为脱敏过程中乘法逆元，并对输出的多源异构能源数据元素进行计算，即

$$G_\alpha' = (G_\alpha, \alpha(x))^{-1}。$$

根据多源异构能源数据逆元的选取结果，在有限域 $GF(2^8)$ 上对其进行加法运算时，设置仿射变化对作为平衡参量，以此确保加法运算执行的对称性。其计算方式可以表示为：

$$\begin{cases} y = (G_\alpha, \alpha(x))^{-1} + a(x) \\ x = G_\alpha^{-1}(y + a(x))^{-1} \end{cases} \quad (5)$$

式中： y 和 x 分别表示多源异构能源数据非线性表达变化的正向输出和逆向输出。

借助式(5)得到多源异构能源数据非线性表达变化的正向输出和逆向输出后，利用二者的对称性完成对称算法对目标函数的迭代。此迭代过程通过增加有限域上多项式表达式复杂度，强化数据混淆效果，在保障脱敏输出防御能力的同时，因前期借助 $L_{2,1}$ 范数量化特性、构建目标函数过滤噪声，减少了冗余计算，最终实现多源异构能源数据的快速脱敏，保障数据安全。

2 测试分析

2.1 测试准备

在测试阶段，以KingbaseES数据库管理系统中的关系型数据库信息作为测试数据。作为一个大型通用跨平台系统，KingbaseES负载大量能源业务的在线事务处理，以及密集数据仓库的管理和查询响应应用。在具体的数据构成方面，其主要包括风机、光伏板和储能电池等发电设备的运行参数信息，同时以每分钟一次的频率更新。其中，测试数据集的数据情况如表1所示。

表1 测试数据集

字段名	类型	说明
id	NUMBER(8)	主键
valtype	NUMBER(2)	数据采集的积分周期
tm	NUMBER(4)	数据时间
datetime	date	通讯采集上的原始数据
rawdata	NUMBER	1有效2无效
rawtype	NUMBER(2)	经过小数点处理后的数据
repdata	NUMBER	1原始值2人工修改3人工补录4数据
reptype	NUMBER(1)	电量类型
tmeng	NUMBER	导入5主备替换
engtype	date	根据报表值计算出的一次侧积分间隔值
valid	NUMBER(1)	电量值状态
rowtype	NUMBER(1)	是否校验
bak	NUMBER(64)	数据来源

数据库能源数据信息的记录值以纵向存储形式存在，每列对应一条记录值，更新一次产生的数据为2万个。以上述的数据构成为基础，测试阶段选取连续30 min内KingbaseES

数据库管理系统中更新的数据信息作为测试的数据集，分析设计技术对于多源异构能源数据的脱敏效果。

2.2 测试结果与分析

分别设置面向深度学习的动态数据脱敏技术，以及基于差分隐私的数据脱敏技术作为测试的对照组，对比分析了不同技术在数据脱敏方面的性能。在具体的分析阶段，针对数据脱敏的可靠性，选取严格雪崩准则距离作为评价指标；针对数据脱敏的效率，选取时间复杂度作为评价指标。

首先，对比分析不同技术下的严格雪崩准则距离，利用其评价数据脱敏技术的防御能力，测试结果如图 1 所示。

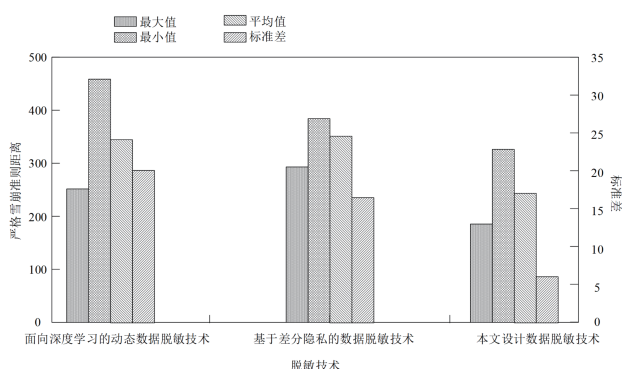


图 1 严格雪崩准则距离对比图

结合图 1 所示的信息可以看出，在 3 种不同数据脱敏技术下，本文设计技术下严格雪崩准则距离明显低于对照组。从数据安全的角度分析，较小的严格雪崩准则距离表明数据脱敏技术对测试数据库能源数据信息变化具有较好的扩散效果，能够实现能够更有效地隐藏明文与密文之间的关系的目的。结合严格雪崩准则距离越小，数据脱敏技术的抗攻击性越强这一逻辑关系可以判断，本文设计的技术具有良好的抗攻击能力。这是因为本文设计就是通过引入对称算法对多源异构能源数据脱敏目标进行迭代，避免了由于复杂程度较低引起的抗攻击性较低问题，进而保障了对数据的脱敏效果。

其次，对比分析了不同技术的时间复杂度，对随机生成 1 000 次大小为 128 bit 测试数据的时间进行统计，其结果如图 2 所示。

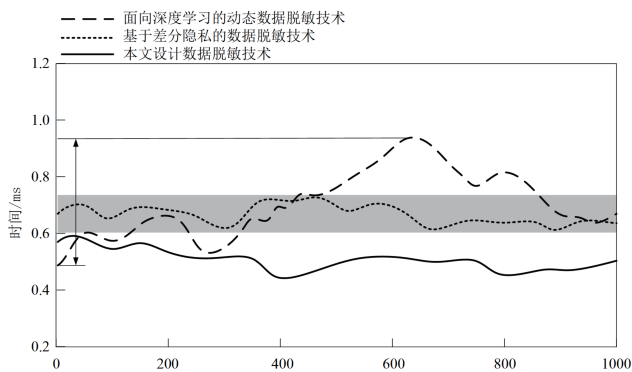


图 2 时间复杂度对比图

结合图 2 所示的测试结果可以看出，在 3 种不同数据脱敏技术下，本文设计技术对应的时间复杂度最低，基本稳定在 0.6 ms 以内，且并未出现明显的波动，表明设计技术可以实现对测试多源异构能源数据的快速脱敏。这是因为本文设计技术利用 $L_{2,1}$ 范数对多源异构能源数据初始化特性的敏感性进行量化，避免了由于数据自身结构和构成引起的数据脱敏构成计算量较大，脱敏速度较慢的问题，进而实现了对多源异构能源数据的快速脱敏。

3 结语

对于多源异构能源数据而言，其在构成上表现出的差异，以及数据中包含的噪声都在一定程度上限制了数据脱敏的效率。针对此，本文开展了基于对称算法的多源异构能源数据快速脱敏方法研究。通过引入 $L_{2,1}$ 范数，实现了对多源异构能源数据初始化特性的敏感性量化，进而平滑多源异构能源数据的结构差异；结合数据之间相似性构建了脱敏目标函数，并对脱敏阶段的变形参数分布进行约束，实现了对数据中噪声的有效过滤。利用对称算法对目标函数进行迭代时，通过调换仿射变化对和乘法逆元，增加了有限域上多项式表达式的复杂度，保障了数据的脱敏效果。该方法严格雪崩准则距离较低、时间复杂度较低，脱敏效果可靠。然而，应对超大规模数据集时，该方法的脱敏适配性仍存在局限。未来研究可以探索跨领域多类型异构数据脱敏的通用适配方案，推动该方法在更广泛场景落地应用，为能源数据安全治理提供更完善技术支撑。

参考文献：

- [1] 陈新, 覃月凤, 周学林, 等. 多维标度分析中个体差异尺度模型的向量序列加速算法 [J]. 计算数学, 2025, 47(2): 255-284.
- [2] 苏林茂, 宋学贵, 向一, 等. 大数据与人工智能融合下敏感数据识别与脱敏技术在网络监管中的应用 [J]. 中国科技资源导刊, 2025, 57(2): 27-39.
- [3] 高博, 常莉, 张金波, 等. 基于 BERT-CRF 和数据加密的石油勘探数据脱敏与安全保护研究 [J]. 自动化与仪器仪表, 2025(3): 20-24.
- [4] 窦亚珍. 面向深度学习的动态数据脱敏技术研究 [J]. 电脑编程技巧与维护, 2025(2): 125-128.
- [5] 李思慧, 戴明超, 蔡伍洲. 基于差分隐私的数据脱敏技术研究 [J]. 网络安全与数据治理, 2025, 44(2): 39-43.
- [6] 蔡彬彬. 基于差分隐私的多维数据脱敏方法研究 [J]. 信息记录材料, 2024, 25(5): 160-162.
- [7] 许华琳. 面向云计算的安全架构设计: 对称与非对称加密算法的应用 [J]. 中国宽带, 2025, 21(2): 109-111.

多自由度机械臂三维空间多障碍物避障与路径规划算法研究

李星宇¹

LI Xingyu

摘要

现有方法在规划机械臂避障路径时缺乏对障碍物运动趋势的预测,无法提前规避潜在碰撞风险,导致规划性能下降。因此,文章提出了一种多自由度机械臂在三维空间中对多障碍物避障的路径规划算法,建立机械臂坐标系的齐次变换矩阵,定位末端位姿,实现机械臂正运动学建模;运动过程中根据机动性能、制动距离、碰撞条件等多重综合约束,制定优先级约束实现避障;对满足约束条件的代价函数进行评估,建立代价函数;完成初步代价筛选同时对动态障碍物避障,运用 D* Lite 算法筛选避障最优路径。实验表明,机械臂可以提前预测环境中存在的障碍物,保证避障距离。路径曲率较小,规划路径更为稳定、舒适,证明了该方法的有效性。

关键词

多自由度机械臂;多障碍物;三维空间;路径规划算法

doi: 10.3969/j.issn.1672-9528.2025.11.042

0 引言

随着工业对人机协作、复杂环境适应性的需求激增,服务机器人及空间探索技术的快速发展,多自由度机械臂已成为高维复杂空间中任务执行的核心载体。从柔性制造车间的人机协作装配,到航天器舱外的在轨服务,再到手术室内的微创机器人手术,机械臂需在三维非结构化环境中实现自主避障与动态路径规划。

蔡泽平等^[1]采用激光扫描的方式获得障碍物的参数,基于表面点的障碍物感知和斥力生成避障算法,提出一种安全避障策略。但该算法采集的数据过于庞大,运算复杂难以

求出抖动时的准确路径。黄帅等人^[2]设计了无人驾驶矿卡局部避障路径规划算法。引入 Frenet 坐标系简化障碍物和参考路径信息,采用边界转换策略将避障路径规划中的非凸约束转换为凸约束,构建二次规划模型;求解优化函数,实现了避障路径规划。该方法未考虑到地面风险因素对路径规划的影响,设计路径非最优路径,算法避障精度不高。因此本文提出一种多自由度机械臂三维空间多障碍物避障与路径规划算法,实现低成本的运行轨迹自动规划。

1 多自由度机械臂多障碍物避障与路径规划设计

1.1 机械臂运动建模

多自由度机械臂依靠关节变量的变化来驱动其末端执行器实现各种运动。在建模过程中,通常采用 DH (de-

1. 江苏师范大学机电工程学院 江苏徐州 221116

[8] 刘烨,李帆,郑泽晨.基于上下文编码器网络的油气曲线图像脱敏方法研究[J].计算机应用与软件,2024,41(2):216-221.

[9] 袁梦真,许潇,张彦豪.基于人工智能的非结构化数据脱敏方法研究[J].网络安全与数据治理,2023,42(S1):184-190.

[10] 滑中豪,李宁,王啸,等.GNSS 观测数据的数据脱敏算法及质量验证[J].导航定位学报,2023,11(5):164-171.

【作者简介】

田钺(1996—),男,贵州施秉人,本科,工程师,研究方向:电力行业数字化技术、数据运营。

吴漾(1984—),男,贵州贵阳人,硕士,工程师,研究方向:数据资产管理、电力行业数字化、数据标准。

董若烟(1995—),女,贵州贵阳人,硕士研究生,工程师,研究方向:大数据、人工智能。

孔庆波(1982—),男,贵州贵阳人,本科,工程师,研究方向:大数据、人工智能。

方继宇(1978—),男,贵州贵阳人,本科,高级工程师,研究方向:大数据、人工智能。

(收稿日期:2025-06-13 修回日期:2025-11-05)