

基于多维岭化马氏距离算法的无人机通信安全

王一帆¹ 陈玲玲^{1*} 陈一凡¹

WANG Yifan CHEN Lingling CHEN Yifan

摘要

随着无人机互联网 (internet of drones, IoD) 在智慧城市等场景中的深度应用, 多源感知数据的准确性正面临严峻挑战。为此, 文章提出一种以最小化真实感知地面各部分占比结果与所得感知地面各部分占比结果之间均方误差的优化目标。为实现这一目标, 设计了一种多维岭化马氏距离 (multi-dimensional ridge mahalanobis distance, MDRMD) 算法, 该算法采用逐维建模策略, 分别计算各个无人机在不同特征维度上的马氏距离, 并通过聚合生成综合异常指标, 同时引入岭化方差, 以增强模型在高维、小样本条件下的稳定性。仿真结果表明, 该算法在无人机准确率控制方面显著优于传统马氏距离与 K-means 等常用方法, 具备良好的可扩展性与高效的部署效率。

关键词

无人机互联网; 异常数据识别; 岭化马氏距离; 多维特征融合

doi: 10.3969/j.issn.1672-9528.2026.02.033

0 引言

传统地面通信系统由于网络容量和覆盖范围的限制, 难以全面满足用户日益多样化的无线接入需求^[1]。在此背景下, 无人机通信技术凭借其高机动性、低延迟数据传输、灵活部署等优势迅速发展, 成为重要的补充解决方案^[2-4]。IoD 通过

无人机与地面控制站 (ground control station, GCS) 的协同工作, 为移动用户 (mobile user, MU) 提供包括交通监控以及灾害救援在内的多种服务。该系统利用无人机群采集图像、视频等感知数据并传输至 GCS 进行整合分析, 这种空对地通信模式在复杂环境中展现出更强的适应性和更广的覆盖范围, 显著提升了通信系统的整体性能。

然而, 该技术在实际应用中仍面临数据安全挑战: 一方

1. 吉林化工大学信息与控制工程学院 吉林吉林 132022

[8]MA J, GAO W, WEI Z Y, et al. Detect rumors using time series of social context information on microblogging websites [C]//CIKM'15: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. New York: ACM, 2015: 1751-1754.

[9]VU D T, JUNG J J. Rumor detection by propagation embedding based on graph convolutional network[J]. International journal of computational intelligence systems, 2021, 14(1): 1053-1065.

[10]VU D-T, JUNG J J. Detecting emerging rumors by embedding propagation graphs [C]//Information Retrieval Technology: 15th Asia Information Retrieval Societies Conference, AIRS 2019, Hong Kong, China. New York: ACM, 2020: 173-184.

[11]CHEN J, LIU W Q, HUANG Z Z, et al. Universal deep GNNs: rethinking residual connection in GNNs from a path

decomposition perspective for preventing the over-smoothing [EB/OL]. (2022-05-30)[2025-12-30]. <https://doi.org/10.48550/arXiv.2205.15127>.

[12]HUANG Q, ZHOU C, WU J, et al. Deep spatial-temporal structure learning for rumor detection on Twitter [J]. Neural computing and applications, 2023, 35(18): 12995-13005.

【作者简介】

王梦园 (1999—), 女, 湖北随州人, 硕士研究生, 研究方向: 自然语言处理。

罗卓卿 (1999—), 男, 湖南娄底人, 硕士研究生, 工程师, 研究方向: 自然语言处理。

杨广浩 (1998—), 男, 湖北襄阳人, 硕士研究生, 工程师, 研究方向: 自然语言处理。

(收稿日期: 2025-07-13 修回日期: 2026-01-26)

面，无人机可能因硬件故障或被恶意控制而产生异常感知数据；另一方面，攻击者可能通过劫持无人机获取异常数据，影响系统决策。

为应对上述问题，现有研究大多聚焦于感知数据的聚合优化。其中，信任评估机制被引入以筛选高可信度的无人机节点，从而降低异常节点对融合质量的负面影响^[5-7]。然而，即使在引入信任机制的前提下，异常值仍难以彻底剔除，因此，异常检测技术作为必要的补充手段愈发受到重视，主流研究主要涵盖三类方法：一是基于聚类的检测策略可对数据进行分簇以发现异常样本，但该方法在大规模数据集下计算开销较大^[8]；二是基于统计的检测方式实现简单，但过度依赖数据分布的先验假设，难以适应实际环境中的动态变化^[9]；相比之下，基于距离的检测方法无需预设簇结构或依赖分布假设，兼具高计算效率与良好适应性，因而在无人机感知数据异常检测中表现出更高实用价值，该类方法通过评估数据点之间的相对距离来识别异常，不仅规避了聚类方法的计算瓶颈，也克服了统计方法的分布依赖性，已成为保障IoD系统感知可靠性的关键路径。

针对感知数据异常识别问题，本文提出的MDRMD算法旨在提升IoD系统中感知数据的鲁棒性和异常识别能力。该算法在传统马氏距离的基础上进行了三项关键改进：首先，采用逐维分析方法，分别计算感知向量在各特征维度上的马氏距离，并通过加权策略融合为综合异常指标，显著提升了对局部异常特征的检测灵敏度；其次，针对小样本或高维场景下方差估计不稳定的问题，创新性地引入岭回归正则化技术，对每个维度的方差进行稳健性修正，有效避免了因方差估计偏差导致的误判；最后，算法设计了基于动态分位数的自适应阈值机制，能够根据数据分布特征自动调整异常判定标准，从而实现对异常样本的精准识别与剔除。这一系列改进使算法在保持计算效率的同时，大幅提升了异常检测的准确性和环境适应性。

1 系统模型

在本文中，构建了一个IoD感知与通信模型，该模型由一个GCS和 N 架分布式无人机节点组成。系统运行过程中，各无人机首先独立完成环境感知任务，并将感知数据通过无线链路传输至GCS。GCS在接收到全部感知数据后，首先对数据进行异常剔除处理，随后对保留的正常数据进行聚合，最终生成该区域的综合感知结果。本文所研究的IoD感知与通信模型如图1所示。

为评估地面控制站最终得到的区域感知结果是否准确，本文采用加权平均感知结果 X_G 与真实感知状态 X_T 之间的误差作为评估指标。

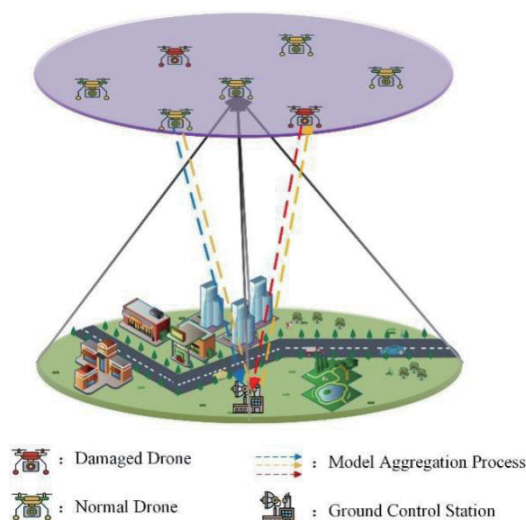


图1 IoD感知与通信模型

首先，设每架无人机 u_i 的感知数据为 x_{u_i} ，其马氏距离为 $d_{u_i}^2$ 。计算地面控制站的加权感知结果 X_G 用公式表示为：

$$X_G = \sum_{i=1}^n w_i x_{u_i} \quad (1)$$

其中，权重 w_i 定义为：

$$w_i = \frac{1/d_{u_i}^2}{\sum_{j=1}^n 1/d_{u_j}^2} \quad (2)$$

该权重表明，马氏距离越小的无人机所占权重越大，马氏距离大的感知数据反而在此设定中占比更低。为定量描述聚合结果与真实感知数据之间的差异，本文进一步构建基于最小均方误差的感知异常函数为：

$$\min_{w_i} \|X_G - X_T\|^2 \quad (3)$$

通过合理分配权重 w_i ，使得最终聚合结果 X_G 尽可能接近真实的环境状态 X_T ，从而提高系统感知准确性。

2 多维岭化马氏距离算法

在改进后的马氏距离方法中，无人机 u_i 第 p 个特征维度均值 $\bar{x}_{u_i}^p$ 的计算为：

$$\bar{x}_{u_i}^p = \sum_{i=1}^N x_{u_i}^p \quad (4)$$

式中： N 表示参与感知任务的无人机数量； $x_{u_i}^p$ 表示第 i 架无人机在第 p 个特征维度上的感知结果。接着，计算第 p 个特征维度的方差，计算公式为：

$$S^2 = \frac{1}{N} \sum_{i=1}^N (x_{u_i}^p - \bar{x}_{u_i}^p)^2 \quad (5)$$

为避免在样本量不足或特征维度较高时方差估计不稳定性问题，本文引入岭回归的正则化思想，构造具有数值稳定性的岭化方差估计量 $S^p(\lambda^p)$ ，用公式表示为：

$$S^p(\lambda^p)^2 = S^2 + \lambda^p \quad (6)$$

式中： λ^p 为岭参数。

在此基础上，无人机 u_i 在第 p 个特征维度上的马氏距离

$d_{u_i}^p(\bar{x}_{u_i}^p)$ 计算为:

$$d_{u_i}^p(\bar{x}_{u_i}^p) = \frac{|x_{u_i}^p - \bar{x}_{u_i}^p|}{S^p(\lambda^p)} \quad (7)$$

随后, 将各特征维度上的马氏距离进行平均, 得到无人机 u_i 的整体马氏距离 $d_{u_i}(x_{u_i})$:

$$d_{u_i}(x_{u_i}) = \frac{1}{p} \sum_{y=1}^p d_{u_i}^p \quad (8)$$

为了识别异常值, 本文引入基于分位数的自适应阈值判别机制, 设定 95% 分位点作为异常识别阈值 τ :

$$\tau = \text{quantile}(d_{u_i}, 0.95) \quad (9)$$

当 $d_{u_i}(x_{u_i}) > \tau$ 时, 判定无人机 u_i 的感知数据为异常值并予以剔除。剔除异常感知数据后, 保留剩余的 n 份正常数据用于后续的聚合计算与可信融合分析。

3 仿真分析

在本节中, 展示了基于改进马氏距离的无人机感知数据异常检测方法的实验结果。为验证所提方法的有效性, 基于 Python 进行仿真设计, 对比分析了改进马氏距离方法与 LOF 方法、K-means 聚类方法以及原始马氏距离方法在不同异常比例下的识别性能。

图 2 展示了 4 种异常检测方法的准确率随异常比例变化的对比结果, 所有方法均在低异常比例 (<25%) 下保持较为稳定的准确率, 但随着异常占比的增高 (25%~50%), 准确率均呈现下降的趋势, 这表明现有算法在高异常比例环境下的检测能力普遍受限。

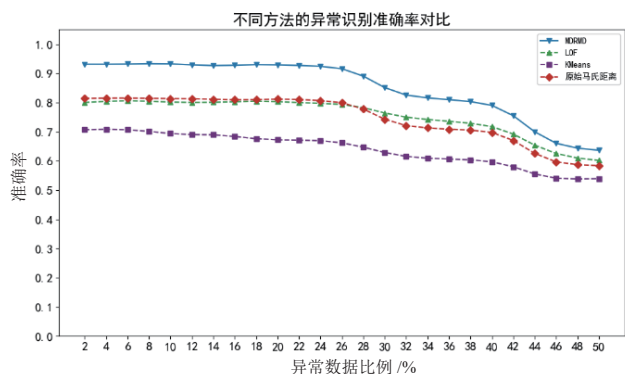


图2 不同异常占比下各算法异常数据识别准确率对比

其中, 改进马氏距离方法始终保持最优性能, 这得益于其采用的岭回归方差修正技术和特征维度独立处理机制, 有效解决了传统方法在协方差矩阵估计中的数值不稳定问题。原始马氏距离方法在低异常比例 (<20%) 时, 异常检测准确率较高, 但随着异常比例增加, 由于未引入正则化处理导致协方差矩阵退化, 其准确率快速下降。LOF 方法由于基于局部密度检测的特性, 在中等异常比例 (25%~50%) 区间表现

略优于原始马氏距离。K-means 方法准确率始终处于最低水平, 表明其固定聚类中心的特性导致对异常数据识别能力有限。整体而言, 改进马氏距离方法通过创新的方差估计方式, 显著提升了异常检测的准确性和鲁棒性。

图 3 展示了 4 种异常检测方法的精确率随异常比例变化的对比结果。整体来看, 所有方法在低异常比例 (2%~25%) 区间均呈现缓慢下降趋势, 但在 25% 异常比例处出现明显的性能回升拐点, 这种特殊现象源于当异常比例达到 25% 这一临界值时, 异常样本的分布特征开始趋于稳定, 反而为检测算法提供了更明显的判别依据。其中, 改进马氏距离方法在整个测试范围内始终保持最高精确率, 这得益于其特征维度独立处理机制避免了高维数据下的协方差矩阵退化。其余 3 种方法在异常比例超过 26% 后, 精确率反而出现回升趋势并高于低异常占比时的水平, 其中, 原始马氏距离方法因协方差矩阵逐渐适应异常分布, 精确率得以提升; LOF 方法由于局部密度特征在高异常比例下更易区分, 精确率得以提升; K-means 方法则因异常簇规模扩大而得以提升。整体而言, 改进马氏距离方法表现出较高的异常数据检测精确率。

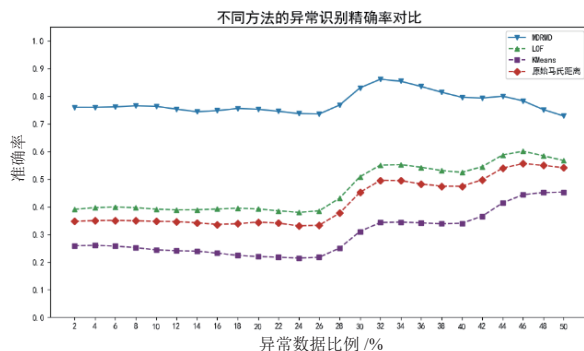


图3 不同异常占比下各算法异常数据识别精确率对比

图 4 展示了 4 种异常检测方法的 F_1 值随异常比例变化的对比结果。在 0%~25% 异常占比区间, 改进马氏距离保持 0.78 的稳定 F_1 值, 显著优于 LOF 方法、K-means 与原始马氏距离。当异常占比超过 25% 后, 改进马氏距离性能开始下降, 最终所有方法的 F_1 值收敛至 0.3~0.4 区间。这种变化趋势主要源于在低异常比例时, 改进马氏距离通过优化的方差处理, 能有效捕捉稀疏异常; 而 LOF 受限于固定邻域参数, K-means 和原始马氏距离则因算法固有局限导致性能较低。随着异常比例增加, 数据分布特性发生质变, 异常样本的聚集效应使各类方法都难以维持原有判别能力。实验结果表明, 改进的马氏距离算法, 在异常数据检测中展现的综合性能优于原始马氏距离算法以及传统的异常数据检测算法。

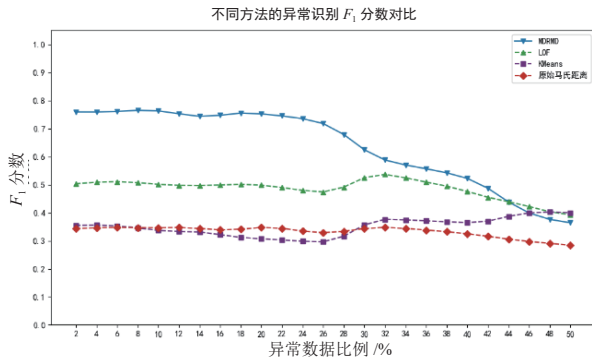


图 4 不同异常占比下各算法异常数据识别 F_1 值对比

图 5 展示了基于改进马氏距离的区域感知误差分析结果。图中实线描绘了最小均方误差随无人机数量增加的变化规律，当无人机数量从 0 增至 76 架时，误差值从初始的 4.2×10^{-5} 迅速下降至 0.3×10^{-5} 以下，呈现出指数衰减趋势，虚线标注的 76 架无人机收敛点表明系统达到最优规模，此时误差稳定在 3×10^{-6} 量级并趋于平稳。该仿真结果验证了改进马氏距离算法在区域感知应用中的有效性。

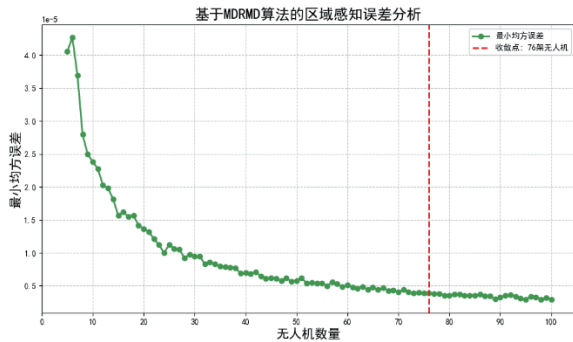


图 5 最小均方误差随不同无人机数量的变化

4 总结

本文针对 IoD 中多源感知数据的异常值干扰问题，提出了一种基于 MDRMD 的异常检测算法。该算法以最小化所得感知结果与真实感知结果之间的均方误差为优化目标，通过逐维建模计算各无人机特征维度的马氏距离并聚合，结合岭化方差增强高维小样本条件下的稳定性，设计自适应阈值方法有效去除异常样本，以此来优化群体感知性能。仿真实验表明，所提方法在检测准确率上较传统马氏距离算法提升 10%，具有良好的可扩展性和部署效率，为复杂 IoD 环境下的协同感知任务提供了可靠的技术支撑。

参考文献:

[1]TAN Y W, LIU J J, WANG J D. How to protect key drones in unmanned aerial vehicle networks? an SDN-based topology deception scheme[J]. IEEE transactions on vehicular technol-

ogy, 2022, 71(12): 13320-13331.

[2]YUAN X P, HU Y L, SCHMEINK A. Joint design of UAV trajectory and directional antenna orientation in UAV-enabled wireless power transfer networks[J]. IEEE journal on selected areas in communications, 2021, 39(10): 3081-3096.

[3]CHAI S Q, LAU V K N. Multi-UAV trajectory and power optimization for cached UAV wireless networks with energy and content recharging-demand driven deep learning approach[J]. IEEE journal on selected areas in communications, 2021, 39(10): 3208-3224.

[4]MESSAOUDI K, OUBBATI O S, RACHEDI A, et al. A survey of UAV-based data collection: challenges, solutions and future perspectives[J]. Journal of network and computer applications, 2023, 216: 103670.

[5]ZHENG J, XU J P, DU H Y, et al. Trust management of tiny federated learning in internet of unmanned aerial vehicles[J]. IEEE internet of things journal, 2024, 11(12): 21046-21060.

[6]KANG X, WU Y D. A trust-based pollution attack prevention scheme in peer-to-peer streaming networks[J]. Computer networks, 2014, 72:62-73.

[7]YE J W, KANG X, LIANG Y C, et al. A trust-centric privacy-preserving blockchain for dynamic spectrum management in IoT networks[J]. IEEE internet of things journal, 2022, 9(15): 13263-13278.

[8]HEIDARI J, DANESHPOUR N, ZANGENEH A. A novel K-means and K-medoids algorithms for clustering non-spherical-shape clusters non-sensitive to outliers[J]. Pattern recognition, 2024, 155: 110639.

[9]LEE H, YUN S. Strategies for imputing missing values and removing outliers in the dataset for machine learning-based construction cost prediction[J]. Buildings, 2024, 14(4): 933.

【作者简介】

王一帆 (1999—)，女，陕西渭南人，硕士研究生，研究方向：无人机通信安全。

陈玲玲 (1980—)，通信作者 (email:cll807900@163.com)，女，吉林长春人，博士，教授，研究方向：认知无线电。

陈一凡 (2001—)，男，安徽亳州人，硕士研究生，研究方向：认知车联网频谱共享。

(收稿日期：2025-07-25 修回日期：2026-02-02)