0254-6124/2024/44(6)-1155-11 *Chin. J. Space Sci.* 空间科学学报

ZHOU Taichun, GUO Guohang, XIAO Zhigang, LI Hu. Anomaly Detection of Satellite Telemetry Data Based on Latent Space Interpolation Autoencoder (in Chinese). *Chinese Journal of Space Science*, 2024, **44**(6): 1155–1165. CSTR:32142.14.cjss.2023–0147. DOI:10.11728/cjss2024.06.2023–0147

基于隐空间插值自编码器的卫星遥测参数异常检测*

周台春[1,2] 郭国航[1,2] 肖志刚[1] 李 虎[1]

1(中国科学院国家空间科学中心 北京 100190)

2(中国科学院大学 北京 100049)

摘 要 卫星遥测参数是地面运管系统评估卫星在轨运行正常状态的关键指标, 遥测参数异常检测对于保障卫星安全可靠运行和任务顺利执行至关重要. 针对现有卫星遥测异常检测算法对参数特征提取存在区分度缺乏、有效异常决策信息提取不充分等问题, 本文提出一种基于隐空间插值优化的异常检测方法, 将隐空间优化约束后的自编码器的表示学习能力与核密度估计方法的密度估计能力相结合, 有效地进行异常检测. 采用量子科学卫星的真实遥测参数数据和公开数据集进行验证, 其结果表明所提方法在真实遥测参数上比最优对比方法的 Auc 值和 F1 值分别提升了 5.6% 和 5.8%. 与其他异常检测算法相比, 该方法有较强的正常和异常样本辨别能力, 有效解决了特征缺乏区分性以及决策信息提取不充分的问题, 同时具有良好的噪声抗干扰性和有效性.

关键词 科学卫星, 遥测参数, 自编码器, 异常检测, 隐空间优化

中图分类号 V557.3

Anomaly Detection of Satellite Telemetry Data Based on Latent Space Interpolation Autoencoder

ZHOU Taichun[1,2] GUO Guohang[1,2] XIAO Zhigang[1] LI Hu[1]

1(*National Space Science Center*, *Chinese Academy of Sciences*, *Beijing* 100190)

2(*University of Chinese Academy of Sciences*, *Beijing* 100049)

Abstract Satellite telemetry parameters are the critical indicators for the ground operation and management system to assess the normal state of satellite operation in orbit, and anomaly detection of telemetry parameters is essential to guarantee the safe and reliable operation of satellites and the smooth execution of tasks. In response to the existing satellite telemetry anomaly detection algorithms for parameter feature extraction there is a lack of differentiation, effective anomaly decision-making information is not sufficiently extracted and other problems, this paper proposes an anomaly detection method

* 中国科学院战略性先导科技专项科学卫星任务运控技术项目资助 (XDA15040100)

2023-12-13 收到原稿, 2024-01-19 收到修定稿

©The Author(s) 2024. This is an open access article under the CC-BY 4.0 License (https://creativecommons.org/licenses/by/4.0/)

based on the optimization of latent space interpolation, the latent space optimization constraints after the self-coder's representation learning ability and the density estimation ability of the Kernel Density Estimation (KDE) method are combined to effectively carry out the anomaly detection. Real telemetry parameter data from quantum science satellites and public datasets are used for validation, and the results show that the proposed method improves the Auc and F1 values over the optimal comparison method by 5.6% and 5.8%, respectively, on real telemetry parameters. Compared with other anomaly detection algorithms, the proposed method has strong ability to discriminate normal and abnormal samples, effectively solves the problems of lack of differentiation of features and insufficient extraction of decision information, and has good noise immunity and effectiveness.

Key words Scientific satellite, Telemetry parameters, Autoencoder, Anomaly detection, Latent space optimization

0 引言

受极端空间环境因素、系统设计问题以及设备质量因素等[1,2]影响, 卫星在其整个生命周期内可能会发生或多或少的异常或故障, 严重情况下甚至会导致卫星运行任务失败. 卫星遥测参数数据与卫星的性能状态息息相关, 其反映的是卫星设备或系统的功能及性能的变化, 是地面工作者监视和评估卫星是否正常运行的关键指标. 通过对卫星遥测参数进行异常检测, 可以有效识别卫星在轨期间所发生的异常或意外事件, 并采取必要的干预措施进行调整, 这对确保卫星安全稳定、可靠运行具有重要意义, 也是目前航天领域中一个备受关注的热门研究课题[3].

在航天器领域, 借助文献 [4] 中的定义, 卫星的异常通常是指卫星在运行期间出现与预期不符或未经计划的事件 (行为). 异常检测是指识别数据集或系统中超出正常范围的数据或行为的过程[5]. 卫星遥测参数异常检测是进行卫星故障诊断和实时健康监控的重要手段, 其方法主要分为 4 类: 基于预设阈值、基于专家系统、基于物理模型以及基于数据驱动的异常检测. 其中基于预设阈值的异常检测, 需要人工判别, 费时费力, 而且某些卫星遥测参数没有超过阈值的异常无法检测; 基于专家系统的异常检测, 只能基于专家经验建立已知异常的检测规则, 无法处理未定义的异常现象; 基于物理模型的检测, 建模过程复杂, 而且无法对大型复杂系统的所有部件都建模. 而基于数据驱动的异常检测, 通过收集的海量历史遥测参数数据, 利用统计、数据挖掘和机器学习等领域的方法应用在卫星遥测参数数据的建模过程中, 这种方法不需要完全依赖完整的专家经验和领域知识, 可拓展能力强, 可根据遥测参数数据对检测模型动态更新.

近年来, 随着任务密度的提高以及待监测参数量的提升, 数据驱动的异常检测吸引了越来越多的学者进行研究. Wang 等[6] 针对皮纳卫星的高维遥测数据设计了一种离线和在线两阶段异常检测算法, 在线阶段进行遥测参数的单点异常检测, 而离线阶段则针对遥测数据的集体异常进行检测. Zhang 等[7] 针对多变量遥测参数提出了一种基于最大信息系数的层次聚类方法, 通过挖掘遥测参数之间的相关性, 将遥测参数分成子空间并利用长短期记忆网络 (LSTM) 对其进行单参数异常检测. Shi 等[8] 提出了一种基于特殊点序列分割的卫星遥测时间序列聚类方法, 该方法在具有突出形状特征的时间序列数据集的检测效果较好. Pan 等[9] 提出一种卫星电力系统异常检测主成分分析 (KPCA) 算法, 该算法利用 KPCA 对数据集空间的特征矩阵进行结构分析进而实现异常检测, 其有效性在风云卫星电源分系统数据集上得到了验证. 李海玲等[10] 利用异常检测规则知识和故障案例, 设计了一种关键字提取的关联引擎方法, 充分体现了关联案例在航天器异常检测的实际作用.

随着深度学习的发展, 已有研究提出了一种两阶段的异常检测方法: 利用深度网络得到样本的特征表示, 根据获取的样本特征构建模型计算异常分数以实现异常检测. 例如 Li 等[11] 提出一种两阶段异常检测框架, 利用自监督学习和生成单类分类器进行异常检测; Liu 等[12] 提出一种将印象提取器 (IE-Net) 与专家网络 (Expert-Net) 相结合的无监督异常检测方法. 此外, Deng 等[13] 提出局部切割空间分析策略和 Erfani 等[14] 采用深度信念网络 (DBN) 与一类支持向量机 (OC-SVM) 的组合来代替降维, 企图将异常样本和正

常样本区分开. 虽然两阶段的异常检测算法是一种处理具有复杂相关关系数据异常问题的有效方法, 但直接通过对特征空间的密度估计来有效识别卫星遥测参数中的异常还面临着以下问题.

(1) 对于具有复杂相关关系的卫星遥测数据异常检测, 仅仅只使用深度网络进行降维和特征学习, 而缺乏适当的约束, 会使得获取的遥测参数数据间的依赖关系不够准确, 这将导致映射到特征空间之后正常样本和异常相交叠, 两者无法有效区分.

(2) 异常分数的构建只考虑代理任务中的单维度特征, 导致异常决策信息提取不够充分, 不能反映卫星的真实状态, 且易受噪声数据干扰, 从而导致异常检测性能不理想.

基于上述原因, 本文将自编码器 (AE) 获取的隐空间进行插值优化, 并利用核密度估计模型进行异常分数构建, 提出一种新的用于卫星遥测参数异常检测的算法. 与标准自编码器不同, 本文提出的异常检测算法由两支自编码器构成并在隐空间进行插值约束, 同时约束后的自编码器不仅考虑了重构误差, 还对数据在隐空间中的分布进行优化, 使得数据的密度分布和关键特征在低维表示中得以保留, 密度值高的更有可能是正常对象, 密度值低的则被认为是异常对象, 从而为正常数据构建一个更有区别的轮廓.

本文提出的隐空间优化卫星遥测参数异常检测方法, 通过对自编码器压缩后的隐空间进行插值, 并且引入带有正则化的评价网络进行约束, 以获取更加紧凑的特征分布, 从而提高对正常数据构建准确轮廓的能力, 增强正常和异常样本之间的区分度. 采用非参数的 KDE 方法来构建异常分数, 同时综合考虑代理任务中多个维度不同层次的重构特征, 将隐空间优化约束后的自编码器的表示学习能力与 KDE 的密度估计能力相结合, 能够有效检测出遥测参数中的异常数据. 利用隐空间插值自编码器与核密度估计模型相结合实现卫星遥测参数异常检测, 实验结果表明, 本方法的有效性和噪声抗干扰性较好, 并在量子科学实验卫星的遥测数据上表现出了较好的性能.

1 问题定义

1.1 遥测参数形式化描述

为了便于对卫星遥测参数的异常检测进行建模分析, 给出其形式化描述如下.

定义 1 设某科学卫星的遥测参数向量集合为 $\boldsymbol{P}$, 采集的遥测参数为

$$\boldsymbol{P}=(p_1,p_2,\cdots,p_i,\cdots,p_k). \tag{1}$$

其中, p_i表示第i个参数的取值情况, k为参数个数.

定义 2 以星上时为基准, 设某时刻t收集到的遥测数据向量

$$\boldsymbol{V}^{(t)}=(t,\boldsymbol{P}^{(t)})=(t,p_1^{(t)},p_2^{(t)},\cdots,p_i^{(t)},\cdots,p_k^{(t)}). \tag{2}$$

其中, $\boldsymbol{P}^{(t)}$表示t时刻采集的k维遥测参数向量.

定义 3 设某设备或分系统的健康状态为δ, $\delta\in\{0,1\}$, 此处, $\delta=0$表示正常, $\delta=1$表示异常. 在某一时刻t, 某设备或分系统的状态向量记为$\boldsymbol{Q}^{(t)}$, 有

$$\boldsymbol{Q}^{(t)}=(t,\delta^{(t)}). \tag{3}$$

其中, $\delta^{(t)}$表示对应时刻设备的健康状态.

1.2 基于遥测参数数据的异常检测问题模型

由于受到空间环境和星上状态动态变化等因素的影响, 新的异常类型会不断出现, 无法穷尽所有类型的异常. 此外, 异常样本通常数量较少, 而正常状态下的数据样本较容易获取. 因此, 问题可以归约为半监督异常检测: 通过使用正常状态下的遥测参数数据构建模型, 然后使用该模型来检测异常. 设给定的遥测样本数据训练集为$X_{\text{train}}=\{(x_i,\delta^{(i)})\}_{i=1}^{N}$, 其中, $x_i\in R^k$. R^k表示k维实数空间, 则该遥测参数异常检测问题的数学建模为

$$G:R^k\to\delta. \tag{4}$$

其中, $G(\cdot)$表示一个映射函数. 在训练样本集X_{train}, 对于每组遥测参数样本$(x_i,\delta^{(i)})$, $\delta^{(i)}=0$. 首先, 利用特征提取函数$\varphi(\cdot)$学习训练样本的特征, 即从原始遥测数据空间R^k映射到特征空间R^h的过程, 有

$$\varphi:R^k\to R^h. \tag{5}$$

根据样本特征来构建异常分数函数$F(\cdot)$, 即

$$F:R^h\to R. \tag{6}$$

其中, R为实数空间. 由异常分数构建判别函数$T(\cdot)$, 进行异常决策, 有

$$T:R\to\delta^{(i)}. \tag{7}$$

其中, $\delta^{(i)}\in\{0,1\}$.

综上所述, 对于给定的训练样本集X_{train}, 学到的异常检测模型为

$$G(x_i) = T \circ F \circ \varphi(x_i) = \delta^{(i)}. \tag{8}$$

其中, $x_i \in X_{\text{train}}$, 对于每一个样本x_i, 最终都有一个$\delta^{(i)}$与之对应, $\delta^{(i)} = 0$表示正常, $\delta^{(i)} = 1$表示异常.

2 研究方法

基于以上问题定义, 设计卫星遥测参数异常检测的整体框架如图 1 所示, 主要包括遥测数据的准备阶段和基于隐空间插值自编码器的异常检测方法两个部分.

2.1 数据准备

在数据的准备过程中, 将原始遥测数据进行采集和解析, 按时间拼接成数据集. 对拼接的卫星各个遥测参数数据进行数据预处理, 其中包括对离群点数据采用3σ原则进行剔除, 对于缺失值数据使用多项式插值方法进行填补, 以及将状态量参数进行 one-hot 编码, 对连续量参数使用Z–Score标准化处理, 标准化处理公式为

$$y' = \frac{y - \mu}{\sigma}. \tag{9}$$

其中, y为某一遥测参数的具体数值, μ为该参数的样本平均数, σ为该参数的样本标准差. 最后将经过预处理后的数据进行划分, 分为训练集数据和测试集数据, 以进行遥测参数的异常检测.

2.2 算法模型

提出的卫星遥测参数异常检测模型 (Latent Space Interpolation Autoencoder Anomaly Detection, LSIA-AD) 主要分为特征学习和异常检测构建. 特征学习阶段主要是利用隐空间插值自编码器网络进行训练, 异常检测阶段先进行异常分数计算, 然后进行异常判别.

2.2.1 隐空间插值自编码器网络结构

自编码器作为一种无监督学习的深度学习方法, 能够学习到隐空间数据的关键特征并用较少的信息进行表示, 而且可以从该表示中重构出原始数据[15]. 针对具有复杂关系的卫星遥测参数, 在自编码器过程中对隐空间进行插值优化处理, 通过隐空间进行插值的方法来生成新样本, 从而使得数据分布更加紧凑密集, 提取的遥测参数特征信息更丰富有效. 最终构建的异常检测模型如图 2 所示. 插值过程的两组隐变量z_1, z_2是由两组随机样本x_1, x_2经编码器g_θ编码所得, 然后通过凸优化插值得到混合隐变量z_1', 再由解码器h_φ对z_1'进行解码, 因此插值的计算公式如下:

$$x_\alpha = h_\varphi(\alpha z_1 + (1-\alpha) z_2), \tag{10}$$

$$z_1 = g_\theta(x_1), \tag{11}$$

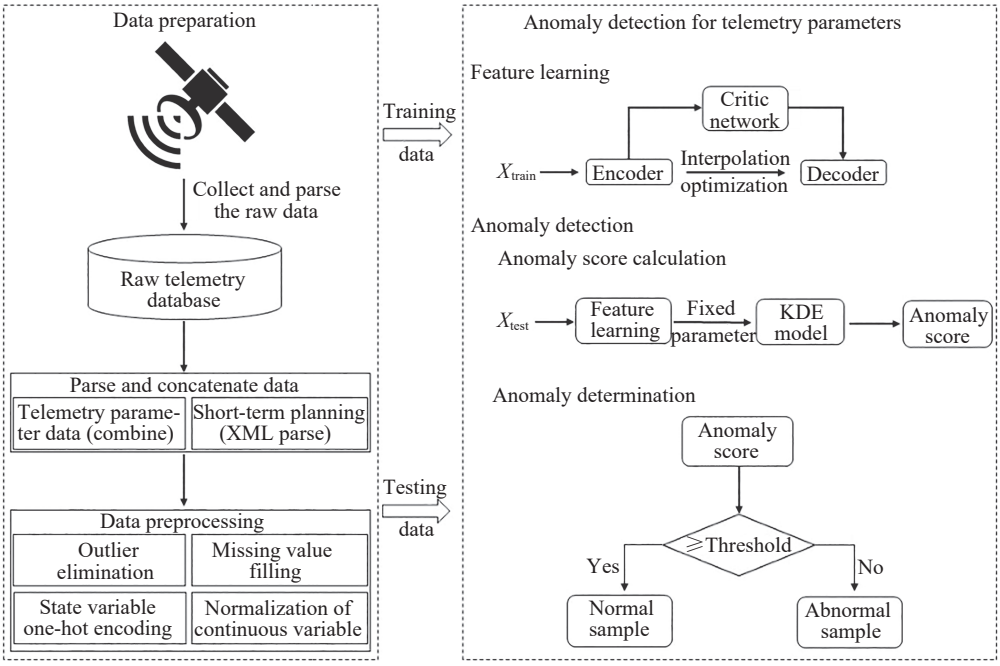


图 1 算法框架

Fig. 1 Framework of the algorithm

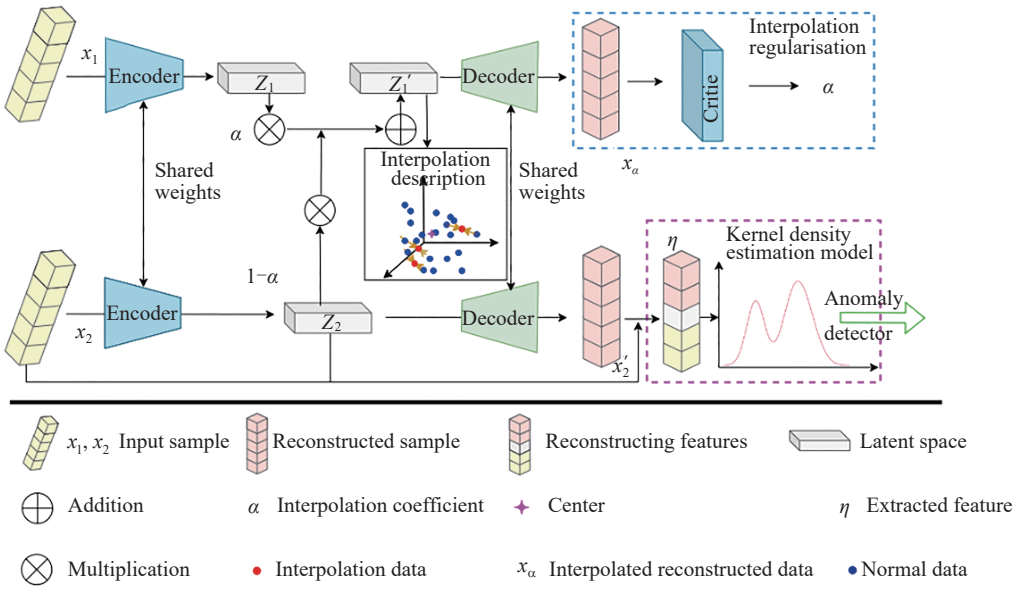


图 2 隐空间插值自编码器网络结构

Fig. 2 Framework of autoencoder for latent space optimization

$$z_2 = g_\theta(x_2). \tag{12}$$

其中, $\alpha \in [0,1]$, 当参数α从 0 到 1 不断变化取值时, 会生成一系列用于插值的数据点, 这将为后续分类器的训练提供基础.

如图 2 所示, 隐空间优化的自编码器框架由两部分构成, 即上支部分和下支部分. 上支部分是一个评价模块, 由评价自编码器和评价网络组成, 该模块的作用是用来评估插值的质量, 以及预测插值约束系数 α ($\alpha \in [0,1]$), 其计算公式如下:

$$\alpha = s_\xi(h_\varphi(\alpha z_1 + (1-\alpha)z_2)). \tag{13}$$

其中, $s_\xi(\cdot)$表示评价网络. 下支部分由自编码器和 KDE 组成, 该部分作用是根据学习到的特征表示进行异常判别. 在模型的训练阶段这两个部分是相互对抗作用的. 具体流程是, 对原始遥测训练数据空间 R^k中的任意两组输入样本, 经过编码器$g_\theta(\cdot)$映射到隐空间z, 然后进行插值优化, 上支评价模块对插值后重构的数据进行评估, 此时下支进行特征学习. 而在测试阶段, 主要工作的是下支部分的模型, 此时将训练阶段的自编码器和 KDE 模型参数固定, 对测试的遥测数据样本进行异常判别.

因此, 评价网络对插值效果的好坏评判是至关重要的, 其思想类似于对抗神经网络[16], 评估网络具有与编码器等效的网络结构. 评估网络$s_\xi(\cdot)$是将插值经过解码生成的重构数据x_α作为输入, 目的是由插值的效果来反向预测参数α. 由于式 (10) 中隐空间变量 z_1和z_2在插值过程具有对称性, 因此可以将参数α的范围设置为$[0, 0.5]$. 此时, AE 的作用是通过在其损失函数中加入正则化项来参与训练, 达到“欺骗”评价网络$s_\xi(\cdot)$的目的, 使$s_\xi(\cdot)$认为参数α的值是 0, 则$s_\xi(\cdot)$网络的损失优化函数为

$$\begin{aligned} l_\xi =& \|s_\xi(x_\alpha) - \alpha\|^2 + \\ & \|s_\xi(x + (1-\gamma)g_\theta(h_\varphi(x)))\|^2, \end{aligned} \tag{14}$$

$$x_\alpha = h_\varphi(\alpha g_\theta(h_\varphi(x_1)) + (1-\alpha)g_\theta(h_\varphi(x_2))). \tag{15}$$

其中, γ为超参数. 式 (14) 中第一项是训练$s_\xi(\cdot)$由插值数据来预测α, 第二项则作为正则化器进行约束. 该正则化器的作用是, 当非插值数据作为$s_\xi(\cdot)$的输入时, 企图使评价网络的输出为 0; 通过在数据空间中对原始样本x和重构样本$g_\theta(h_\varphi(x))$之间进行插值, 即便自编码器重构的数据效果较差, 也能保证评论网络中存在真实数据的输入. 因此, 经过插值的自编码器引入正则化项优化后的损失函数为

$$l_{g,h} = \|x - g_\theta(h_\varphi(x))\|^2 + \varepsilon\|s_\xi(x_\alpha)\|^2. \tag{16}$$

式中, ε为一个超参数, 其目的是调节正则化器的权重. 此处, 该正则化项的作用是不论α的取值如何, 都能有效地使$s_\xi(\cdot)$的输出为 0, 从而达到“欺骗”$s_\xi(\cdot)$的目的, 使其将输入的插值数据视为是未经插值过的(即$\alpha = 0$).

2.2.2 异常检测阶段

研究不仅对自编码器的隐空间进行插值优化, 还考虑了其不同角度的重构误差特征, 以用于遥测参数的异常检测. 由于卫星遥测参数间的复杂相关性, 因此提出的 LSIA-AD 方法选取的重建误差特征η是多维的, 其中包含了将欧氏距离d_{euc}和余弦相似度$d_{\cos}$作为重建特征. 计算公式如下:

$$\eta = [z, d_{\text{euc}}, d_{\cos}], \tag{17}$$

$$d_{\text{euc}} = \sqrt{\sum_{i=1}^{k}(x_{1i} - x'_{1i})^2}, \tag{18}$$

$$d_{\cos} = \frac{\sum_{i=1}^{k}(x_{ji} \times x'_{ji})}{\sqrt{\sum_{i=1}^{k}(x_{ji})^2} \times \sqrt{\sum_{i=1}^{k}(x'_{ji})^2}}. \tag{19}$$

其中, η表示重建误差特征, x_{ji}表示第j个样本的第i维度数据, x'_{ji}表示第j个样本的第i维度重构数据, z表示隐变量. 将获取的低维数据特征信息和重构误差作为输入来构建异常函数, 并计算样本数据的异常分数.

利用核密度估计模型作为异常分数的计算, 使用学习到的低表示对输入数据的概率密度分布进行建模. 将预处理后的遥测数据由式 (17) 获得的低维表示为$\eta_1, \eta_2, \cdots, \eta_n$, 通过 KDE 模型计算其概率密度分布. 根据核密度估计模型的定义, 将插值优化后的自编码器提取的特征信息作为输入, 则重构特征$\eta_1, \eta_2, \cdots, \eta_n$的密度分布函数为

$$f_h(\eta) = \frac{1}{n}\sum_{i=1}^{n} K(\eta - \eta_i) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{\eta - \eta_i}{h}\right). \tag{20}$$

其中, K为非负的核函数, 本文在实验过程中使用的是高斯核函数, h为带宽. 最后, 对于每一个遥测参数的测试样本, 都有一个异常判别函数, 用于判断该样本是否异常.

3 实验验证与结果分析

3.1 实验准备

实验采集的数据为量子科学实验卫星于 2017 年 1 月 7 日至 2019 年 2 月 28 日在轨运行期间的遥测参数数据. 数据总计包含 205522 条记录, 涵盖了量子卫星在地星之间存在的隐形传态、密钥分发和纠缠分发这三种实验类型. 将收集到的遥测参数数据进行解析后获得的部分数据见表 1.

按照时间顺序从获取的科学实验计划中提取出在 2017 年 1 月 7 日至 2019 年 2 月 28 日期间该卫星所开展的实验类型数据, 并对其进行编码处理. 最终提取到的实验类型数据及编码结果见表 2.

根据时间对应关系, 将解析后的遥测数据与该卫星的科学实验计划数据进行拼接合并, 最终合并好的数据见表 3.

根据解析出的数据表 3, 利用专家经验选取了与卫星姿控系统工作状态相关的 23 个参数, 其中包含电源电压、载荷工作模式标识 (整型数据) 和卫星是否处于 SAA 区标识 (0/1 状态量) 等属性, 将其进行数据预处理后整合为 Micius 数据集.

为进一步验证算法的泛化性能和异常检测的性能, 本文还选取了 OODS 异常检测数据库中广泛应用于异常检测领域的公开卫星基准数据集 Satellite, 该数据集与量子科学实验卫星设备遥测参数性质类似. 实验数据集的相关信息列于表 4.

表 1 解析出的原始遥测数据

Table 1 Parsed raw telemetry data

Time (UT)		P0x1356WW68	P0x1060WW9	P0x1060WW7	···	P0x1060WW123
7 Jan. 2017	01:03	10.45200062	–51.59808349	–0.23409990	···	–1.05939996
	01:07	10.45200062	–46.31356811	–0.20004299	···	–0.47624000
	01:10	1.24800002	–32.03112792	–0.00009200	···	0.34847998
⋮	⋮	⋮	⋮	⋮	⋮	⋮
28 Feb. 2019	01:39	0.312000006	32.17395019	–0.137237995	···	1.004840016
	01:40	0.156000003	30.81161499	–0.075713999	···	0.819400012
	01:41	0.312000006	29.042785644	–0.04183999	···	0.641279995

表 2 对实验类型数据的提取和编码

Table 2 Extraction and coding of experimental type data

Start time (UT)		End time (UT)		Experiment type	Label experiment
7 Jan. 2017	00:11	7 Jan. 2017	00:44	星地量子纠缠分发	3
	01:44	7 Jan. 2017	03:09	地星量子隐形传态	1
	23:45	8 Jan. 2017	01:10	星地量子密钥分发	2
⋮	⋮	⋮	⋮	⋮	⋮
23 Jan. 2019	01:34	23 Jan. 2019	02:59	星地量子纠缠分发	3
	23:36	24 Jan. 2019	01:01	星地量子纠缠分发	3
23 Jan. 2019	23:16	25 Jan. 2019	00:41	星地量子密钥分发	2
⋮	⋮	⋮	⋮	⋮	⋮

表 3 拼接后的数据

Table 3 Data after concatenation

Time (UT)		P0x1356WW68	P0x1060WW9	···	P0x1060WW123	Labels
7 Jan. 2017	01:03	10.45200062	–51.59808349	···	–1.05939996	3
	01:07	10.45200062	–46.31356811	···	–0.47624000	1
	01:10	1.24800002	–32.03112792	···	0.34847998	2
⋮	⋮	⋮	⋮	⋮	⋮	
28 Feb. 2019	01:39	0.312000006	32.17395019	···	1.004840016	3
	01:40	0.156000003	30.81161499	···	0.819400012	1
	01:41	0.312000006	29.042785644	···	0.641279995	2

表 4 实验数据集的基本信息

Table 4 Basic information of the experimental dataset

数据集	维度	样本数	样本异常比例
Satellite*	36	6535	0.316
Micius	23	205522	0.034

3.2 对比算法

将本文提出的方法 (LSIA-AD) 与以下经典的深度学习的方法进行比较, 具体如下.

OC-SVM[17]: 通过构建超平面将正常样本与异常样本分离, 用于识别异常数据点.

IFOREST[18]: 通过随机选择特征和切分值构建孤立树, 利用路径长度来识别异常样本.

LOF[19]: 通过比较每个数据点与其邻域之间的密度计算异常因子, 用于识别局部异常点.

Deep SVDD[20]: 通过使用神经网络将正常样本映射到一个球形区域内的潜在空间, 并利用该球形区域来描述正常数据的分布. 计算新样本与该球形区域之间的距离来判断其是否为异常点.

DAGMM[21]: 通过使用自编码器学习数据的低维表示, 并基于该表示计算每个样本的重构误差. 使用高斯混合模型对重构误差进行建模, 将正常样本与异常样本分开.

GAOD[22]: 通过利用随机仿射变换的方法来训练分类器实现检测一般数据 (图像和表格) 的异常.

3.3 算法有效性评估

异常检测实际上可看作是一种分类任务, 分类的目标只有两类, 一般正例作为异常样本, 负例作为正

*http://odds.cs.stonybrook.edu/satellite-dataset/

常样本. 选取 Auc 值和 F1 分数作为度量不同模型的异常检测性能指标, 采用文献 [21,22] 中的实验设置方式. 使用完全干净的数据进行训练, 并对实验重复运行 20 次以获取 Auc 值和 F1 的均值和标准差, 以避免偶然因素的干扰; 每次实验, 随机选择 50% 的正常样本数据构成训练集, 将余下的 50% 正常样本数据和所有异常样本数据作为测试集; 根据表 4 中给出的样本数据异常比例τ, 确定异常分数阈值来识别异常. 各个对比算法采用网格搜索法得到最优超参数. 表 5 给出了 LSIA-AD 与所有对比算法在 Micius 和 Satellite 数据集上的检测效果.

由表 5 可知, LSIA-AD 在数据集 Micius 和 Satellite 上的 Auc 和 F1 值高于其余的对比算法, 这表明 LSIA-AD 性能较好. 在数据集 Micius 上, LSIA-AD 比最优秀的对比算法 LOF 在 Auc 值和 F1 值上分别提高了 5.6% 和 5.8%. 在 Satellite 数据集上, LSIA- AD 比最优秀的对比算法 DAGMM 在 Auc 值上提高了 1.5%, 比最优秀的对比算法 LOF 在 F1 值上提高了 4.8%. 总体而言, LSIA-AD 相比 DAGMM 的优势在于以下几个方面: 首先, 通过在隐空间上进行插值约束, 使得其获取的数据分布更为紧凑, 该方法在保持原始遥测数据分布的特征表示学习方面比 DAGMM 的自编码器表现更优. 其次, LSIA-AD 采用 KDE 模型来计算异常分数, KDE 是一种非参数估计, 允许根据数据拟合函数形式而无需先验假设条件. 此外, 对于 DSVDD, 尽管该方法的硬边界设计非常强大, 但由于在模型训练过程中缺乏一些泛化性能的约束, 可能会导致出现过拟合现象. 这种过拟合现象表现为在训练过程中将所有样本映射为特征空间中的某一个点, 从而使得异常检测效果较差.

为了更进一步验证 LSIA-AD 的优越性, 本文从统计学的角度, 采用威尔克森 (Wilcoxon) 符号秩检验对 LSIA-AD 与所有对比算法在这两个数据集上的运行结果进行统计学检验评估. 具体而言, 需要逐一比较 LSIA-AD 与每个对比方法, 对于 LSIA-AD 与每一个对比方法的评估给出如下零假设H_0和备择假设H_1, 即

$$H_0 : A \approx B,\ H_1 : A > B.$$

其中, A 表示 LSIA-AD 在 Micius 或 Satellite 数据集上的运行结果, B 表示相对应数据集上某种对比算法的运行结果. 计算每对方法评估的 p 值, 并在显著性水平$\alpha = 0.01$的条件下进行检验假设. 关于 Micius 和 Satellite 上 Auc 和 F1 的 p 值检验结果见表 6.

由表 6 可知, 在$\alpha = 0.01$情况下的显著性水平时, 只有 DAGMM 算法与 LSIA-AD 算法的配对 Wilcoxon 检验在 Satellite 数据集上 Auc 的 p 值为 0.0111, 略大于α, 而 LSIA-AD 对其他对比算法的 Auc 和 F1 值的威尔克森符号秩检验 p 值均小于α, 综合来看, 该检验的结果是显著的. 因此, 零假设 H_0被拒绝, 而选择接受备择假设H_1, 这说明本文 LSIA-AD 算法的异常检测性能较好.

此外, 在实际应用过程中, 描述数据运行平台情况以及不同算法的资源消耗是非常重要的. 这可以帮助研究人员更好地理解各种算法的性能和效率, 从而为决策提供有力的支持. 因此, 本文针对不同深度学习算法在 Micius 和 Satellite 两个数据集上的资源消耗情况进行实验分析, 具体结果列于表 7.

所有方法的运行配置为: NVIDIA GeForce GTX 1050, python-3.7, torch-gpu-1.10. 由表7 可知, DSVDD

表 5 各实验数据集上 LSIA-AD 与对比算法的平均 Auc 和 F1 (%)

Table 5 Average Auc and F1 score (%) for LSIA-AD and baseline algorithms on experimental datasets

Method	Micius		Satellite	
	Auc ± Std	F1 ± Std	Auc ± Std	F1 ± Std
OC-SVM	76.3 ± 0.6	56.7 ± 0.8	75.8 ± 0.5	67.8 ± 0.4
IFOREST	66.0 ± 2.2	38.4 ± 5.5	80.1 ± 1.3	68.8 ± 0.7
LOF	83.8 ± 0.6	62.5 ± 1.0	84.4 ± 0.8	74.4 ± 0.7
DSVDD	50.7 ± 7.3	33.8 ± 9.1	72.5 ± 4.4	65.1 ± 3.8
DAGMM	63.7 ± 6.8	30.2 ± 8.4	86.1 ± 2.1	69.4 ± 3.5
GAOD	64.9 ± 5.5	49.7 ± 6.7	76.8 ± 3.6	67.3 ± 3.3
LSIA-AD	**89.4** ± 1.1	**68.3** ± 2.9	**87.6** ± 0.8	**79.2** ± 0.9

注 加粗的数值表示最优结果.

在两个数据集上的运行时间和内存消耗是最少的, 其次是本文提出的算法 LSIA-AD, 而 DAGMM 算法的运行时间和内存消耗是最多的. 这是因为, 不同深度学习算法的网络结构和参数数量不同, 从而导致计算量和内存占用不同. 同时, 不同算法的实现方式和优化方法不同, 对硬件资源的利用效率也不同. 具体而言, DSVDD 利用深度神经网络进行特征提取, 将数据映射到高维特征空间中, 这种方法可以减少数据预处理的时间, 同时通过最小化到球心的距离来找到一个紧致的球形边界, 避免了使用所有数据点进行模型训练, 从而降低了运行时间和内存消耗. DAGMM 算法运行时间长、消耗内存多主要是因为其包含深度自编码器的训练过程和高斯混合模型的参数估计过程, 这些过程需要大量的计算资源和时间来完成. 本文提出 LSIA-AD 虽然由两支自编码器构成, 但是由于参数共享, 在一定程度上减少了内存消耗和训练时间. 综上所述, 从内存消耗和运行时间角度分析, 本文提出的算法相对于对比算法来说具有更好的性能, 这为本文提出的算法的实际应用提供了一定的参考.

3.4 噪声抗干扰性检验

在实际应用中, 训练集数据可能包含有噪声或者在构建训练数据集时存在异常样本数据被误标记成正常样本数据的情况. 为检验算法对噪声的抗干扰性, 本文在训练模型时采用添加污染数据后的 Micius 训练集进行实验, 以观察性能指标的变化. 在本实验中, 研究了 LSIA-AD 面对训练数据受到不同程度污染影响时的表现. 每次实验中, 随机选择 50% 的正常样本数据以及不同污染比例 c (c=[0, 1%, 2%, 3%, 4%, 5%, 8%, 10%]) 的异常样本数据进行混合, 构成训练集用于模型训练, 以检验算法的噪声抗干扰性效果, 实验结果如图 3 所示.

由图 3 可知, 在 Micius 数据的训练集上添加不同比例的异常数据, LSIA-AD 和全部对比方法运行 20 次实验后的平均 Auc 和 F1 值的变化情况. 随着添加污染数据的比例c从 1% 逐渐增加到 10%, 所有算法的 Auc 和 F1 值基本都存在波动的现象, 然而各个方法对添加污染数据的波动程度有所差异. 观察到在添加 1%～5% 比例的污染数据后, 本文提出的 LSIA-AD 方法依然可以保持较好的检测效果, 而添加污染数据比例大于 5% 后, 模型性能开始下降. 但是总体而言, 本文提出的方法在所有加入污染比例的数据上的 Auc 和 F1 值效果都比对比方法好, 这表明 LSIA-AD 在 Micius 数据集中添加不同污染比例的数据下表现出了良好的噪声抗干扰性. 此外, 针对所选取的 DSVDD, DAGMM 和 GAOD 三种深度学习的对比算法, 随着训练样本中添加的污染数据比例不断增加, 这三种方法的 Auc 和 F1 值虽然在变小, 但波动不大. 而传统方法 OC-SVM, IFOREST 和 LOF 在污染数据比例逐渐增加的情况下, 这三种方法的 Auc 和 F1 值变化较大, 不太稳定. 尤其是 OC-SVM

表 6 关于 Auc 和 F1 的 Wilcoxon 检验 p 值

Table 6 Auc values and Wilcoxon rank-sum test p-values for F1

	Dataset	DAGMM	IFOREST	DSVDD	OC-SVM	LOF	GAOD
Auc	Micius	3.39×10^{-8}	3.38×10^{-8}	3.38×10^{-8}	3.40×10^{-8}	3.40×10^{-8}	3.39×10^{-8}
	Satellite	1.11×10^{-2}	3.38×10^{-8}	3.39×10^{-8}	3.40×10^{-8}	4.59×10^{-8}	3.39×10^{-8}
F1	Micius	3.40×10^{-8}	3.41×10^{-8}	3.39×10^{-8}	3.40×10^{-8}	6.00×10^{-7}	7.56×10^{-9}
	Satellite	3.33×10^{-3}	3.38×10^{-8}	3.39×10^{-8}	3.37×10^{-8}	3.95×10^{-8}	3.36×10^{-8}

表 7 不同深度学习算法的资源消耗情况

Table 7 Resource consumption of different deep learning algorithms

		DSVDD	DAGMM	GAOD	LSIA-AD
Micius	运行时间/s	400.0	4114.0	483	1245.0
	消耗内存/MByte	566.1	1872.5	1995.3	1748.8
Satellite	运行时间/s	294.0	1340.0	467.0	308.0
	消耗内存/MByte	485.7	1726.6	1964.3	1593.0

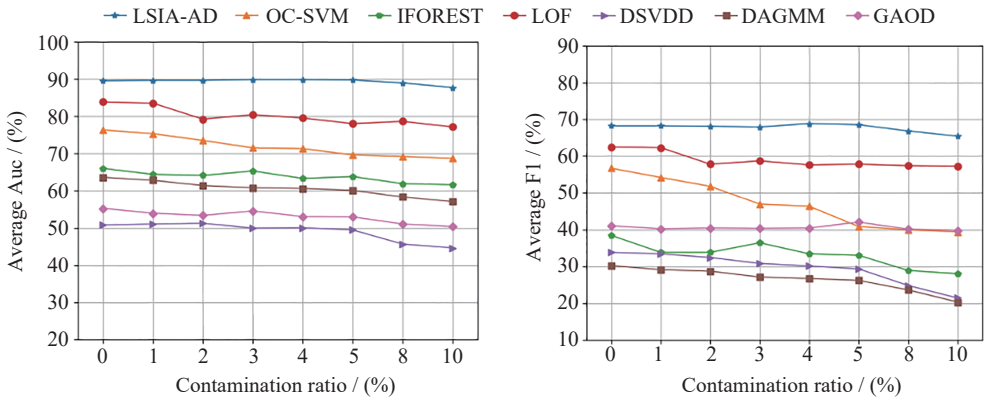


图 3 在 Micius 数据集上添加不同污染比例数据的异常检测结果

Fig. 3 Anomaly detection results on contaminated training data on the Micius dataset

算法, 在不改变该算法的参数设置情况下, OC-SVM 对数据集中掺杂污染比例的数据更为敏感. 由此可见, 在进行异常检测时, 尽量使用污染率低或干净的数据来训练模型, 这样可以获得更好的检测效果.

3.5 消融实验

为了验证 LSIA- AD 算法中隐空间插值优化的有效性, 进行了消融实验.

图 4 描述了隐空间插值优化对本文提出算法的影响, 给出了隐空间插值优化在数据集 Micius 和 Satellite 上 20 次实验的 F1 和 Auc 的均值和标准差. 从图 4 可以看出, 在两个数据集上, 当深度自编码器的网络结构相同时, 采用隐空间插值算法 (LSIA-AD) 相较于不采用隐空间插值算法 (AE-KDE), 其 AUC 和 F1 值均有所提升, 这表明进行插值约束可以更好地优化隐空间, 使得获取的数据分布更紧凑, 从而提高了 LSIA-AD 模型的异常检测效果. 这进一步证实了在训练过程中, 经过插值约束后的隐空间能够充分提取出更关键的特征信息, 并且在本文提出的框架中自编码器经过插值优化后与评价网络相互提升了性能. 通过引入正则化的评价网络, 深度自编码器能够避免落入局部最优值, 从而实现更好的压缩效果. 而经过隐空间插值约束后的自编码器, 为评价网络输入了更有意义的特征表示, 从而实现了噪声抗干扰性的密度估计.

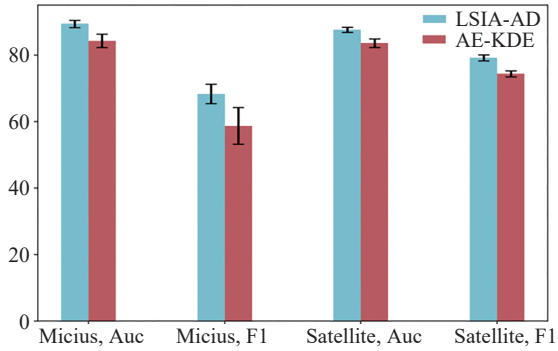


图 4 隐空间插值对 LSIA-AD 算法 Auc 和 F1 的影响

Fig. 4 Impact of latent space interpolation on Auc and F1 of LSIA-AD algorithm

4 结论

针对卫星遥测异常检测中存在的对参数特征提取缺乏区分度、有效异常决策信息提取不充分等问题, 提出了一种基于数据驱动的隐空间插值自编码器异常检测方法 (LSIA-AD), 该方法由评价网络模块和 AE 与 KDE 模块两部分构成. LSIA-AD 通过将样本映射到隐空间中进行插值, 获取了更丰富的特征信息和更加紧凑的特征分布, 然后将隐变量表征和重建特征作为 KDE 模型的输入进行评分来量化数据的异常情况, 从而提高了异常检测的准确性. 在实验研究中, 本文提出的方法在量子科学卫星真实遥测参数数据集和异常检测领域公开数据集上表现出了良好的性能, 并且在受污染数据上的实验也有较好的表现, 具有较好的噪声抗干扰性, 这些实验结果体现了模型的准确性和有效性, 可以为卫星地面操作人员在卫星健康监测和管理方面提供有力的决策支持.

参考文献

[1] LIU Enyu. Research on Fault diagnosis Method and Health Management of satellite Attitude Control System[D]. Shenyang: Shenyang Ligong University, 2023

(刘恩雨. 卫星姿态控制系统故障诊断方法及健康管理研究[D]. 沈阳: 沈阳理工大学, 2023)

[2] LI HU, GUO Guohang, HU Tai, *et al*. Ensemble learning for state recognition of payload from telemetry data[J]. *Journal of National University of Defense Technology*, 2021, **43**(6): 33-40 (李虎, 郭国航, 胡钛, 等. 遥测参数数据载荷状态判别集成学习方法 [J]. 国防科技大学学报, 2021, **43**(6): 33-40)

[3] JIANG H X, ZHANG K, WANG J Y, *et al*. Anomaly detection and identification in satellite telemetry data based on pseudo-period[J]. *Applied Sciences*, 2020, **10**(1): 103

[4] YAIRI T, TAKEISHI N, ODA T, *et al*. A data-driven health monitoring method for satellite housekeeping data based on probabilistic clustering and dimensionality reduction[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2017, **53**(3): 1384-1401

[5] LIU Chao. Spacecraft Anomaly Detection Method based on the Deep Learning Research[D]. Harbin: Harbin Industrial University, 2022 (刘超. 基于深度学习的航天器异常检测方法研究 [D]. 哈尔滨: 哈尔滨工业大学, 2022)

[6] WANG Chan, WANG Huiquan, JIN Zhonghe. Pico-satellite telemetry anomaly detection through clustering[J]. *Journal of Harbin Institute of Technology*, 2018, **50**(4): 110-116 (王婵, 王慧泉, 金仲和. 皮纳卫星遥测数据异常检测聚类分析方法 [J]. 哈尔滨工业大学学报, 2018, **50**(4): 110-116)

[7] ZHANG L W, YU J S, TANG D Y, *et al*. Anomaly detection for spacecraft using hierarchical agglomerative clustering based on maximal information coefficient[C]//2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA). Kristiansand, Norway: IEEE, 2020: 1848-1853

[8] SHI X T, PANG J Y, LIU D T, *et al*. Satellite telemetry time series clustering with improved key points series segmentation[C]//2017 Prognostics and System Health Management Conference (PHM-Harbin). Harbin, China: IEEE, 2017: 1-7

[9] PAN D W, LIU D T, ZHOU J, *et al*. Anomaly detection for satellite power subsystem with associated rules based on kernel principal component analysis[J]. *Microelectronics Reliability*, 2015, **55**(9/10): 2082-2086

[10] LI Hailing, DONG Guangran, DANG Qi. A spacecraft anomaly detection case correlation method[J]. *Telecommunication Engineering*, 2018, **58**(7): 843-847 (李海玲, 董广然, 党琦. 一种航天器异常检测案例关联方法 [J]. 电讯技术, 2018, **58**(7): 843-847)

[11] LI C L, SOHN K, YOON J, *et al*. CutPaste: Self-supervised learning for anomaly detection and localization [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 9659-9669

[12] LIU Y F, ZHUANG C Q, LU F. Unsupervised two-stage anomaly detection[OL]. arXiv preprint arXiv: 2103.11671, 2021

[13] DENG T Q, YE D S, MA R, *et al*. Low-rank local tangent space embedding for subspace clustering[J]. *Information Sciences*, 2020, **508**: 1-21

[14] ERFANI S M, RAJASEGARAR S, KARUNASEKERA S, *et al*. High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning[J]. *Pattern Recognition*, 2016, **58**: 121-134

[15] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. *Nature*, 1986, **323**(6088): 533-536

[16] BERTHELOT D, RAFFEL C, ROY A, *et al*. Understanding and improving interpolation in autoencoders via an adversarial regularizer[OL]. arXiv preprint arXiv 1807.07543, 2018

[17] SCHÖLKOPF B, WILLIAMSON R, SMOLA A, *et al*. Support vector method for novelty detection[C]//Proceedings of the 12th International Conference on Neural Information Processing Systems. Denver, CO: MIT Press, 1999: 582-588

[18] LIU F T, TING K M, ZHOU Z H. Isolation forest[C]//2008 Eighth IEEE International Conference on Data Mining. Pisa, Italy: IEEE, 2008: 413-422

[19] BREUNIG M M, KRIEGEL H P, NG R T, *et al*. LOF: Identifying density-based local outliers[J]. *ACM SIGMOD Record*, 2000, **29**(2): 93-104

[20] RUFF L, VANDERMEULEN R, GÖRNITZ N, *et al*. Deep one-class classification[C]//Proceedings of the 35th International Conference on Machine Learning (PMLR). Stockholm, Sweden, 2018: 4393-4402

[21] ZONG B, SONG Q, MIN M R, *et al*. Deep autoencoding gaussian mixture model for unsupervised anomaly detection[C]//6th International Conference on Learning Representations (ICLR). Vancouver, BC, Canada, 2018**: 1-19**

[22] BERGMAN L, HOSHEN Y. Classification-based anomaly detection for general data[OL]. arXiv preprint arXiv: 2005.02359, 2020

作者简介



周台春 男, 1999 年 2 月生于湖北黄冈, 中国科学院国家空间科学中心计算机应用技术专业硕士研究生, 主要研究方向为大数据处理分析技术、遥测参数异常检测和凸函数不等式及其应用等.
E-mail: taichunzhou@163.com



肖志刚 (通信作者) 男, 1976 年 9 月生于河南周口, 中国科学院国家空间科学中心副研究员, 硕士生导师, 主要研究方向为卫星智能运控技术、大数据处理分析技术和卫星载荷健康管理技术等.
E-mail: xiaozhigang@nssc.ac.cn