

DOI: 10.19650/j.cnki.cjsi.J2312173

# 基于对比序列重构的卫星遥测数据异常检测方法\*

李桢煜, 宋宇晨, 彭喜元, 刘大同

(哈尔滨工业大学电子与信息工程学院 哈尔滨 150080)

**摘要:** 基于遥测数据的异常检测是卫星在轨运维管理的关键技术。但现有方法大多仅采用正常样本建立模型, 异常检测结果对判读阈值敏感、虚警率较高。对此本文提出基于对比序列重构的卫星遥测数据异常检测方法, 充分利用有限异常先验增强异常检测中正常、异常样本差异。先基于变分自编码器提取遥测数据时序演化特征, 引入对比学习方法建立对异常、正常数据差异化输出的编码器, 再用大量正常数据进一步训练整个模型实现对正常数据的精准重构形成对异常数据敏感的时序数据重构模型, 再基于核密度估计方法学习异常判读阈值进一步提升异常检出率。在真实卫星遥测数据上验证表明所提方法能有效降低异常检测的虚警率(均低于 0.002)并保持较高的检出率具备良好的实际应用水平。

**关键词:** 卫星; 遥测数据; 异常检测; 对比学习; 变分自编码器

**中图分类号:** TP311 TH701 **文献标识码:** A **国家标准学科分类代码:** 510.4030

## Contrastive time-series reconstruction method for satellite anomaly detection

Li Zhenyu, Song Yuchen, Peng Xiyuan, Liu Datong

(School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150080, China)

**Abstract:** Anomaly detection based on telemetry data is a key technology for the on-orbit operation and maintenance management of satellite. However, most of the existing methods only use normal samples to build models, while the anomaly detection results are sensitive to the detection threshold, resulting in a high false positive rate. To address this problem, this paper proposes an anomaly detection method based on contrastive time-series reconstruction satellite of telemetry data, which makes full use of the prior knowledge of limited abnormal telemetry samples to enhance the differences between normal and abnormal samples. First, variational autoencoders is used to extract the time-series evolutionary characteristics of telemetry data, specifically the contrastive learning method is introduced to establish an encoder with differentiated outputs of abnormal and normal telemetry data, which uses a large amount of normal telemetry data to further train the whole model to achieve precise time-series reconstruction of normal telemetry data and form a time-series data reconstruction model sensitive to abnormal data. Then the anomaly detection threshold of satellite telemetry data is deduced based on the kernel density estimation method to further improve the detection rate of abnormal samples. Experimental verification was conducted using real satellite telemetry data and the results show that the proposed method can effectively use historical abnormal samples to establish an anomaly detection model, effectively reduce the false positive rate (all below 0.002) of anomaly detection and maintain a high detection rate at the same time, keeping a good practical application level.

**Keywords:** satellite; telemetry data; anomaly detection; contrastive learning; VAE

## 0 引 言

卫星是一类被广泛应用于通信、导航、遥感等领域的航天器<sup>[1-3]</sup>。目前,我国航天事业正处于高速发展阶段,

仅在 2023 年上半年,我国航天发射次数就已达到 25 次,已将近百余颗卫星成功发射入轨。常态化的高密度发射已成为目前我国航天领域发展的新趋势。各型号卫星“快速研制、快速入轨、长时运行”也成为我国卫星领域发展的重要目标和未来趋势。

收稿日期:2023-11-21 Received Date: 2023-11-21

\* 基金项目:国家自然科学基金(62201177)项目、黑龙江省自然科学基金优秀青年项目(YQ2023F006)资助

随着卫星应用由“试验型”向“业务型”的快速转化,其在轨运行期间的任务模式的复杂度和多样性也不断提升。与此同时,伴随星群、星座等卫星应用模式的持续更新和变革,卫星的自主性和智能化水平不断提升,卫星自主任务规划、自主控制的能力快速增强。相应地,卫星在轨运行过程和任务模式的不确定性也逐步增强。在此背景下迫切需要提升卫星的状态监测和评估水平,保证卫星的安全稳定运行。

美国国家科学研究委员会(United States National Research Council, NRC)将异常定义为“航天器发生的不符合期望或非计划的事件”<sup>[4]</sup>。受到空间环境、系统自身性能等因素,卫星会在其在轨服役期间发生包含非关键系统突发故障、硬件受损、软件故障等多种异常事件。相应地,异常事件发生的全过程中,对应分系统的遥测数据会出现非预期变化<sup>[5]</sup>。因此,对卫星遥测数据进行异常检测,能够有效发现卫星在轨运行过程中的异常事件,及早发现卫星子系统中的非预期的故障或性能衰减为卫星的在轨运维提供有效的信息支撑。

卫星遥测数据异常检测方法可分为4类,即:基于阈值的方法、基于专家经验的方法、基于物理模型的方法和数据驱动的方法<sup>[6]</sup>。

基于阈值的异常检测方法需要判断总线电流、温度等各类遥测数据值是否在规定的上下限的预定范围内,当遥测参数值在阈值外时认为此时系统出现异常。Decoste<sup>[7]</sup>通过逐步生成连续的上下界阈值包络实现对航天器遥测数据的异常检测。基于阈值的异常监测方法参数构成较简单实际应用中较易实现,但由于阈值固定在某一范围内,对一些因传感器误差造成的短期波动可能异常敏感从而出现较高的虚警率。

基于专家系统的遥测数据异常检测方法通过卫星的设计师在设计和研制阶段建立的知识库来检测遥测数据中包含的异常。此类方法需要在多维遥测数据之间建立相应的规则库,当遥测数据之间的关联关系不满足既定规则时认为出现异常,这样的系统包括美国国家航空航天局(National Aeronautics and Space Administration, NASA)构建的MOORE<sup>[8]</sup>、SCARES<sup>[9]</sup>、StarPlan<sup>[10]</sup>,法国电信卫星的DIAMS<sup>[11]</sup>以及日本宇宙科学研究所(Japan Aerospace Exploration Agency, JAXA)的ISACS-DOC<sup>[12]</sup>等。它们能对特定型号的卫星系统进行良好的分析建模,但过于依赖历史经验,判定规则庞大繁杂,难以自行发现卫星工作状态变化及时更新规则库,同时也难以迁移至类似型号卫星系统。

基于模型的遥测参数异常检测方法通过建立与实际系统运行机理保持一致的物理模型实现对卫星在轨运行过程的仿真,通过仿真数据与真实数据之间的对比实现异常检测。如Shangguan等<sup>[13]</sup>就采用MWorks作为底层

框架使用Modelica语言对卫星电源系统进行物理仿真建模,实现了对真实遥测数据的异常检测。这类方法检测结果有良好的可解释性,但随着卫星运行任务、模式复杂度的提升,系统难以被简化的模型精确描述,模型参数的辨识和更新难度也限制了此类方法的应用能力。

数据驱动的异常检测方法则利用地面站积累的历史遥测数据作为训练数据,仅关注数据间的数学关系,无需预先构建规则库和模型库,对复杂任务场景下卫星遥测参数之间的动态关联耦合具有良好的适应能力。同时随着机器学习、深度学习等相关理论的快速发展,通过智能算法对遥测数据建立分类、聚类、时序重构或回归预测模型实现对卫星遥测数据的异常检测,已成为国内外相关领域的研究重点。例如NASA埃姆斯研究中心<sup>[14]</sup>提出了一种基于聚类的感应监控系统;JAXA提出了基于混合概率主成分分析与隐马尔科夫模型的遥测数据异常检测模型<sup>[15-16]</sup>;欧空局则引入统计学习模型计算遥测数据的局部分布实现异常片段的识别<sup>[17]</sup>。Li等<sup>[18]</sup>提出了霍夫曼多尺度熵提高不同数据类型之间的区分度并提出一种改进的自适应粒子群优化方法来训练支持向量机(support vector machine, SVM)对卫星动量轮电压遥测数据进行异常检测;Lesouple等<sup>[19]</sup>提出一种基于孤立森林方法根据二叉树中的根节点和叶子节点的距离实现异常点判读。此外,最小二乘支持向量机(least square support vector machine, LSSVM)<sup>[20-21]</sup>、高斯过程回归(gaussian process regression, GPR)<sup>[22-23]</sup>、极限学习机(extreme learning machine, ELM)<sup>[24-25]</sup>、长短期记忆网络(long short term memory, LSTM)<sup>[26-28]</sup>、时间卷积网络(time convolutional network, TCN)<sup>[29]</sup>、图神经网络(graph neural network, GNN)<sup>[30]</sup>、生成对抗网络(generative adversarial network, GAN)<sup>[31]</sup>、自编码器(autoencoder, AE)<sup>[32]</sup>等相关算法,凭借自身优异的非线性拟合能力也广泛应用于遥测数据异常检测模型的构建中。上述方法大多利用历史正常遥测数据建立遥测参数自身的时序演化模型,或多类遥测参数间的关联映射模型,对被检测的遥测物理量进行准确预测或重构,进而计算其与实际值之间的残差,通过判断残差是否在既定的阈值内实现遥测数据的异常检测。

但在实际中卫星任务模式变化、子系统关联关系的动态耦合以及遥测链路中存在的干扰、噪声等,使上述异常检测模型对残差阈值相对敏感检测结果易出现虚警;另外卫星遥测数据中的正常、异常样本存在显著不平衡情况,正常样本远多于异常样本,难以获取精确的分类模型,因而现有模型往往只采用正常样本进行训练,忽略了少量异常样本带来的有效信息,导致模型可能在已知异常上表现仍欠佳,因而如何充分利用有限异先验使正常、异常样本之间的差异更加显著,是提升卫星遥测数据异常检测精度、降低虚警率的关键。

针对现有模型存在的上述问题,本文提出了基于对比序列重构的卫星遥测数据异常检测方法,以变分自编码器(variational autoencoder, VAE)时间序列重构模型为基础,对多维航天器遥测数据进行重构和异常检测。VAE作为一个时序重构模型,相对于分类、聚类模型能够达到更好的未知异常模式检测效果,同时相对于回归预测模型能够减少模型参数所需要占用的空间资源,适用于本文所描述资源受限、数据不平衡场景。时序重构往往基于AE,而VAE将AE中的隐层向量变成了具有不确定性的概率分布,能避免模型过拟合增强其泛化性,提高模型对训练样本中未出现片段的重构性能。所提出方法首先将历史遥测数据拆分成相同长度的工作模式片段,利用VAE对遥测数据进行重构,其中还引入了对比学习策略,构造包含异常先验的对比学习训练集,使编码器能够在隐层空间中扩大正常遥测与异常遥测间的差异实现对未知状态遥测数据的差异化重构。在此基础上基于核密度估计方法(kernel density estimation, KDE)学习异常判读阈值,实现基于重构误差分布特性的卫星遥测数据异常检测,有效约减卫星遥测数据异常判读对判读阈值的敏感度,保证较高检出率的同时降低虚警率。

本文的具体安排如下:第1章明确卫星遥测数据的生成方式和异常模式;第2章详细介绍基于序列重构误差的异常检测模型。第3章采用真实卫星遥测数据对所提出方法进行实验验证。第4章对全文进行总结并对未来研究工作进行展望。

## 1 卫星遥测数据异常模式分析

### 1.1 卫星遥测数据的产生

卫星是由多个子系统组成的复杂系统,通过各个子系统之间的协同配合完成遥感、通信、导航等复杂任务,典型卫星系统的组成结构如图1所示。

其中有效载荷是保证卫星能够完成特定任务的分系统。平台分系统是保证卫星能够正常工作的子系统。各子系统在星载计算机的控制下协同完成各类任务。通过传感器采集各系统重要物理量,经模数转换、数据压缩、调制解调等形式形成地面可用的遥测数据,支撑地面管理和卫星运行状态判断便于运维和任务决策。

### 1.2 卫星遥测数据的异常分类

卫星的非预期运行、软硬件退化损等均会反映到遥测数据中。按遥测数据的表现形式分类可分为单点异常、关联异常和集合异常3类<sup>[33]</sup>。

#### 1) 单点异常

点异常指一段遥测数据中超出其余所有数据点取值范围的数据点。如图2所示是某卫星母线输出电流遥测

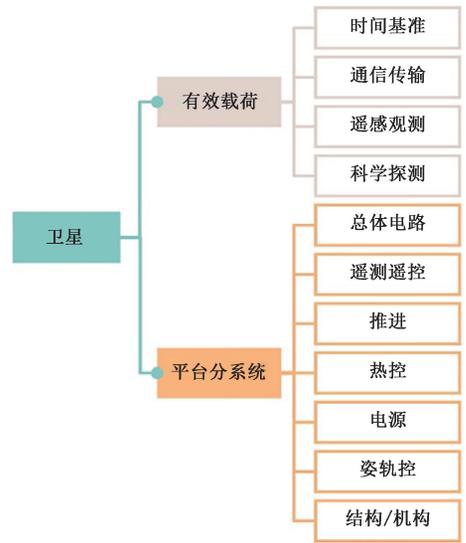


图1 卫星子系统的基本构成

Fig. 1 Basic subsystems of satellite

数据(归一化后)的变化情况,可以发现图中存在单点离群现象,这些数据为点异常。

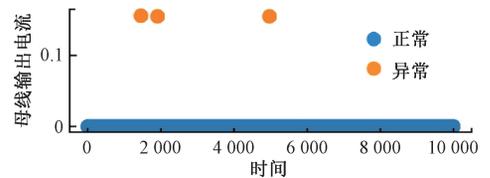


图2 卫星遥测数据单点异常模式

Fig. 2 Single-point anomaly of satellite

#### 2) 关联异常

关联异常指在一段数据中与临近时间中其他片段存在显著偏差的片段。如图3所示的卫星蓄电池压力数据中,第600~700s的压力显著偏离原有的156 Pa左右的稳定区间呈现较大的先下降后上升的异常过程。对此阶段的分析结果表明卫星太阳电池阵模块发生了对日定向驱动故障而无法对帆板姿态进行调整导致太阳帆板输出功率降低,蓄电池出现异常充放电,相对于其前后片段单体压力呈现出异常的变化模式。

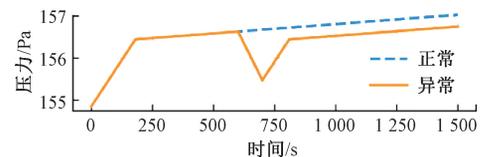


图3 卫星关联异常

Fig. 3 Contextual anomaly of satellite

#### 3) 集合异常

集合异常指与其他数据相比相似数据点的集合作为

一个整体被视为异常。在如图4所示的卫星蓄电池荷电状态遥测数据中第600s开始出现异常下降。此阶段相的分析结果表明卫星充电调节器的输出功率降低而无法保持对蓄电池的充电电流。遥测数据自身未出现明显离群现象,与上下文也未出现显著偏差,但整合在一起后其趋势发生了变化,是一种集合异常。

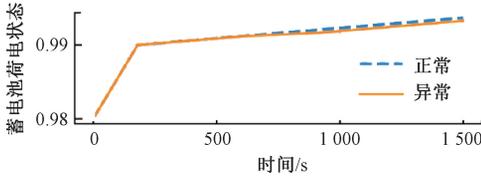


图4 卫星集合异常

Fig. 4 Collective anomaly of satellite

卫星遥测数据的变化受到工作模式和系统运行状态等多方面因素的共同影响呈现出不同的异常模式。部分异常模式下,异常、正常数据间的差异相对较小,且异常数据的变化过程呈现长时缓慢的趋势,难以设置合适的阈值实现遥测数据异常的判读,过于严格的阈值会降低准确率,而过于宽松的阈值会提高虚警率。因此对于正常、异常样本而言,将样本输入异常检测模型后输出应具有显著差异,从而降低异常判读结果对检测阈值的敏感度,同步提升异常检测准确率、降低虚警率。

## 2 基于序列重构误差的遥测数据异常检测

为充分利用先验异常使异常检测模型对正常、异常样本的输出差异更加显著,本文提出了基于序列重构误差的卫星遥测数据异常检测模型,模型原理如图5所示。所提出方法包括离线训练和在线检测两个阶段。离线训练过程中,基于对比学习理论,利用正常样本和异常样本训练VAE模型在其隐层空间内实现正常和异常样本的差异化表征;在线异常检测中,将未知状态异常遥测样本输入到VAE模型中,对遥测数据进行重构后计算遥测数据的重构误差,利用KDE计算和提取重构误差的分布特征,从而实现异常遥测数据的判读。

### 2.1 基于变分自编码器的卫星遥测数据重构

自编码器通过学习遥测数据变化的时序规律实现对遥测数据的特征表示。对于卫星遥测数据而言,将原始数据输入,经编码器后形成隐层表示后再被输入到解码器复原原始输入<sup>[34]</sup>。相比于普通的AE,VAE的隐层表示变成了具有不确定性的概率分布,模型泛化性更强,同时在训练编码器的过程中引入对比学习方法放大正常数据与异常数据间隐层表示的差异,增大异常样本的重构误差,便于后续异常检测部分有效降低虚警率。

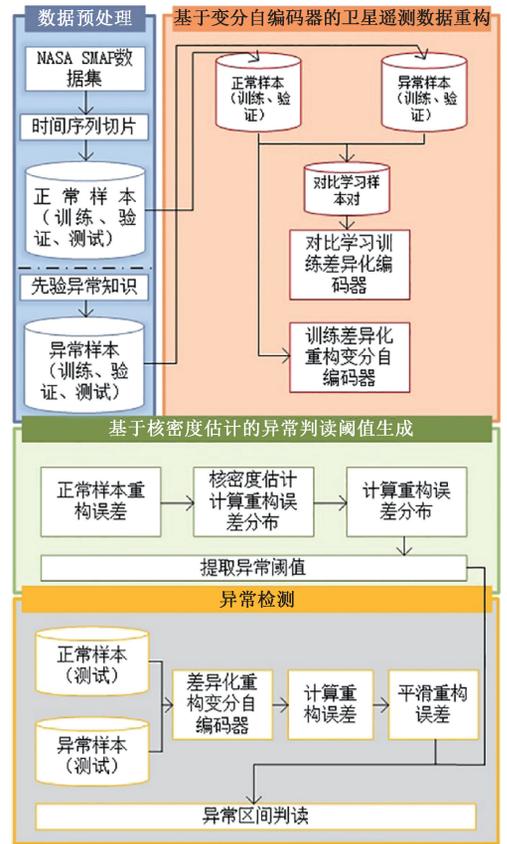


图5 基于重构误差分布异常检测方法原理

Fig. 5 Schematic of anomaly detection method based on reconstruction error distribution

VAE利用遥测数据 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 拟合相应遥测物理量的实际分布。首先假设一个满足标准正态分布的 $p(\mathbf{z})$ ,即:

$$p(\mathbf{z}) \sim N(0, 1) \quad (1)$$

再用此正态分布对实际遥测样本概率分布 $p(\mathbf{X})$ 逼近:

$$p(\mathbf{X}) = \sum p(\mathbf{X}|\mathbf{z})p(\mathbf{z}) \quad (2)$$

在式(2)所示的过程中对于遥测数据中的样本 $\mathbf{x}_k$ 在隐层空间的分布 $p(\mathbf{z}|\mathbf{x}_k)$ ,可以用均值 $\mu$ 和方差 $\sigma^2$ 表征,引入神经网络对这两个参数进行拟合:

$$\mu_k = f_1(\mathbf{x}_k) \quad (3)$$

$$\log \sigma_k^2 = f_2(\mathbf{x}_k) \quad (4)$$

其中, $f_1(\cdot)$ 、 $f_2(\cdot)$ 分别是对均值 $\mu$ 、方差 $\sigma^2$ 的对数进行拟合的映射。在训练过程中为了避免网络参数退化VAE使 $p(\mathbf{z}|\mathbf{x}_k)$ 都向标准正态分布近似,从而有:

$$p(\mathbf{z}) = \sum_{\mathbf{x}} p(\mathbf{z}|\mathbf{x})p(\mathbf{x}) = \sum_{\mathbf{x}} N(0, 1)p(\mathbf{x}) = N(0, 1) \sum_{\mathbf{x}} p(\mathbf{x}) = N(0, 1) \quad (5)$$

由此,可以满足式(1)所示先验假设。

VAE用已知的遥测样本逼近实际遥测的分布来实现遥测数据重构。对于重构的遥测数据序列 $\hat{\mathbf{X}}$ ,通过其与实际遥测数据样本 $\mathbf{X}$ 之间的距离 $L_{rc}$ 定VAE训练的损失函数:

$$L_{rc} = \|\mathbf{X} - \hat{\mathbf{X}}\|^2 \quad (6)$$

为使 $p(\mathbf{z}|\mathbf{x}_k)$ 逼近标准正态分布函数,在式(3)定义损失函数基础上建立正则化函数:

$$L_{\mu} = \|f_1(\mathbf{x}_k)\|^2 \quad (7)$$

$$L_{\sigma^2} = \|f_2(\mathbf{x}_k)\|^2 \quad (8)$$

由此可将正态分布逼近的问题转换为使分布 $p(\mathbf{z}|\mathbf{x}_k)$ 与标准正态分布 $N(0,1)$ 距离最小的问题,于是损失函数进一步变为:

$$L_{VAE} = L_{rc} + L_{\mu, \sigma^2} = \|\mathbf{X} - \hat{\mathbf{X}}\|^2 + \frac{1}{2}(-\log\sigma^2 + \mu^2 + \sigma^2 - 1) \quad (9)$$

其中, $L_{rc}$ 为通过数据重构误差定义的损失函数, $L_{\mu, \sigma^2}$ 逼近标准正态分布的损失函数。

至此得到可完成遥测数据重构VAE模型,输入遥测数据样本 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 能重构得到对应的重构序列 $\hat{\mathbf{X}} = \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_n\}$ 用于后续遥测数据的异常判读。

## 2.2 基于对比学习的变分自编码器训练

上述的VAE能通过拟合和逼近遥测数据的分布实现遥测数据的重构。但对于与正常遥测样本相近的异常样本,对其重构后仍难以辨识或扩大其与正常样本的差异。因此需要对VAE的训练过程进行进一步优化,在隐层空间中扩大正常样本与异常样本间的差异。

对比学习是近年来兴起的一种无监督学习方法,其目的在于使相似样本在隐层空间中尽可能接近同时使不同样本之间的距离尽可能远<sup>[35]</sup>。因此可以在VAE的训练中引入对比学习策略扩大正常数据与异常数据之间的差异便于实现对异常遥测的高精度判读。

首先对包含有异常样本的历史遥测数据进行划分:定义参考正常遥测数据序列 $\mathbf{X}^A$ 、正常遥测数据序列 $\mathbf{X}^P$ 、异常遥测数据序列为 $\mathbf{X}^N$ ,建立用于对比学习训练的正样本对 $\{\mathbf{X}^A, \mathbf{X}^P\}$ 和负样本对 $\{\mathbf{X}^A, \mathbf{X}^N\}$ 。再将VAE中的编码器部分变成有共享权重的孪生神经网络(siamese neural network)架构<sup>[36]</sup>,损失函数形式为:

$$L_{Encoder} = \frac{1}{N} \sum_{i=1}^N y_i \log p(y_i) + (1 - y_i) \log(1 - p(y_i)) \quad (10)$$

其中, $N$ 是一次输入孪生神经网络架构编码器的样本对个数, $y_i$ 是当前第 $i$ 个对比学习样本对的标签, $p(y_i)$ 是孪生神经网络架构编码器的输出。上述训练过程如图6所示。

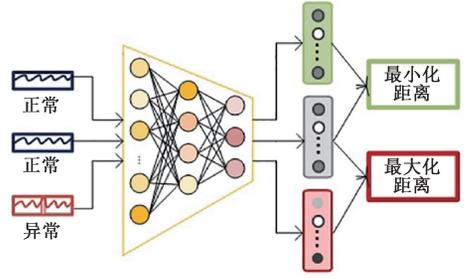


图6 基于对比学习的变分自编码器编码器部分训练  
Fig.6 The contrastive learning diagram of encoder in VAE

进而按照2.1节所述过程使用正常样本数据集同时作为输入和目标输出对整个VAE进行训练可实现对正常样本的精准重构。

## 2.3 基于核密度估计的异常判读阈值生成

异常判读的阈值是影响遥测数据异常检测结果的一个主要因素。如何利用已有遥测数据生成可信的检测阈值是本文所提出方法需要解决的另一个关键问题。本文采用KDE方法拟合经过VAE重构后的遥测数据重构误差分布,提取用于后续异常检测的判读阈值。对于遥测数据序列片段 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ ,对应的重构误差为 $\mathbf{E} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ ,通过KDE方法拟合序列重构误差的概率分布函数 $\hat{f}_h(\mathbf{e})$ 可以表示为:

$$\hat{f}_h(\mathbf{e}) = \frac{1}{Nh} \sum_{n=1}^N K\left(\frac{\mathbf{e} - \mathbf{e}_n}{h}\right) \quad (11)$$

其中, $\mathbf{e}$ 为选定的重构误差点, $h$ 为带宽, $K$ 为高斯核函数。带宽 $h$ 是重构误差点 $\mathbf{e}$ 附近的邻域,采用高斯核函数时,带宽可通过式(12)进行计算:

$$h = \left(\frac{4\sigma^5}{3N}\right)^{\frac{1}{5}} \approx 1.06\sigma N^{-\frac{1}{5}} \quad (12)$$

由此,即可在完成卫星遥测数据重构的基础上,通过统计被检测遥测数据经过VAE的重构误差并拟合其误差分布,实现对遥测数据中异常片段的检测和判读。

## 3 实验验证评估

### 3.1 实验设置

#### 1) 实验数据

利用NASA收集的土壤水分主被动探测卫星(soil moisture active and passive, SMAP)的遥测数据<sup>[26]</sup>进行验证。该遥测数据集是经过归一化处理之后的数据,并不具有明确的物理含义,且整个数据集主要包含点异常和关联异常两类遥测数据异常,因此在本文中,采用A-7、D-7、D-9三个通道的遥测数据来代表相应的数据集进行相关的实验验证本文所提出的方法。这三个通道数据集中正常、异常样本的分布情况如表1所示。

表1 遥测数据集中异常样本的分布情况

Table 1 Distribution of anomalies in telemetry data

数据通道	样本区间	正常样本	异常样本	异常类型
A-7	(0~8 630)	(0~6 199) (8 601~8 630)	(6 200~8 600)	关联异常
D-7	(0~7 641)	(0~4 939)	(4 940~7 641)	点异常
D-9	(0~7 405)	(0~6 249)	(6 250~7 405)	点异常

## 2) 评价指标

本文通过建立卫星遥测数据的序列重构模型,计算卫星遥测数据的重构误差统计分布实现对异常遥测序列的判读。对于一个待检测的遥测数据样本而言,其可能出现的检测结果如表2所示。

表2 遥测参数异常检测效果评价指标

Table 2 Evaluation index of telemetry data

真实状态	检测结果	结果评价
异常样本	异常样本	真阳性(true positive, TP)
异常样本	正常样本	假阴性(false negative, FN)
正常样本	异常样本	假阳性(false positive, FP)
正常样本	正常样本	真阴性(true negative, TN)

通过检出率(true positive rate, TPR)、虚警率(false positive rate, FPR)、准确度(accuracy, ACC)3个指标评价所提出方法的异常检测性能。

## (1) 检出率

检出率表示检测正确的异常样本占所有样本的比率,其计算方式为:

$$TPR = \frac{TP}{TP + FN + FP + TN} \quad (13)$$

## (2) 虚警率

虚警率表示检测结果为假阳性的样本占所有样本的比率,其计算方式为:

$$FPR = \frac{FP}{TP + FN + FP + TN} \quad (14)$$

## (3) 准确率

准确率表示所有正确检测的样本占全部样本的比率,其计算方式为:

$$ACC = \frac{TP + TN}{TP + FN + FP + TN} \quad (15)$$

一般而言,检出率和准确率越高而虚警率越低代表所建立的异常检测模型的检测性能越好。

## 3) 对比实验设计

对比实验中采用孤立森林以及NASA针对此数据提出的非监督异常检测方法(Telemanom)<sup>[26]</sup>进行对比分

析。孤立森林是一种解决时间序列异常检测的经典方法,通过对比遥测数据点的离群特性,并量化遥测数据点与全部数据分布的距离,实现异常数据的判读。Telemanom算法基于LSTM构建,通过预测遥测数据的变化趋势实现异常的判读。

## 3.2 实验结果分析

利用本文所提出的方法,对NASA SMAP卫星遥测数据中A-7、D-7和D-9三个遥测数据样本进行异常检测,并与数据集中人工标记的异常样本进行比对。3个遥测数据样本中,A-7样本中存在的是相比点异常更加复杂的关联异常,因此从遥测序列重构和异常判读两方面以A-7遥测数据为例对异常检测结果进行讨论。

## 1) 遥测数据序列重构结果分析

采用基于对比学习的VAE对NASA SMAP卫星A-7遥测数据序列进行重构,重构结果如图7所示。可以看到在正常的遥测数据片段内重构精度较高,而对于异常样本而言则有更大的重构误差。

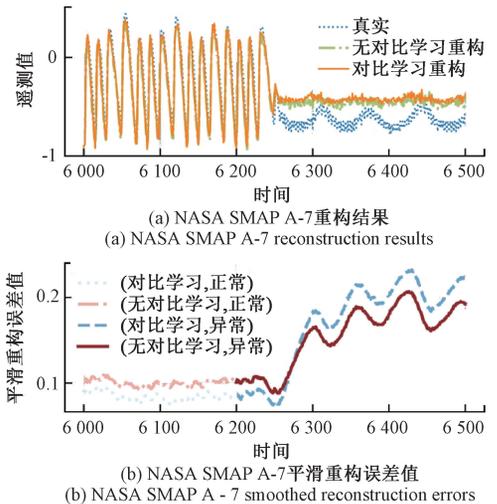


图7 NASA SMAP A-7 遥测序列重构效果图

Fig. 7 Reconstruction results of NASA SMAP A-7

如图7(b)所示,对A-7遥测序列的重构误差可以看出,对于正常遥测数据样本,采用对比学习训练的VAE的重构误差相对更低,而在异常区间内,引入对比学习后的异常序列重构误差也要显著高于未引入对比学习后的重构误差,表明采用对比学习方法,能够有效利用遥测数据中的异常样本,扩大正常和异常遥测间差异。

对于NASA SMAP卫星A-7、D-7和D-9三种遥测数据,对应重构误差分析结果如表3所示。

从表3中数据可以看到在各个数据集上,本文采用的对比学习方式训练的VAE相较于基本的VAE能够在对正常时间序列样本以较高精度重构,同时又能够在一定程度上放大异常时间序列样本的重构误差,两者的平

表 3 遥测数据重构误差统计结果

Table 3 Reconstruction errors results of telemetry data in normal and abnormal intervals

数据样本	重构模型	样本状态	平均绝对误差	正常/异常样本差异
A-7	本文方法	正常	0.182	<b>0.029</b>
		异常	0.211	
	VAE	正常	0.169	0.014
		异常	0.183	
D-7	本文方法	正常	0.003	<b>1.002</b>
		异常	1.005	
	VAE	正常	0.005	1.000
		异常	1.005	
D-9	本文方法	正常	0.004	<b>1.533</b>
		异常	1.537	
	VAE	正常	0.005	1.519
		异常	1.524	

均绝对误差均高于 0.029,使用相同数据集的前提下,至少比 VAE 高 0.002。

2) 异常判读结果分析

利用所提出的重构误差分布特性统计方法,对序列中存在的异常进行判读。对于 NASA SMAP 数据集中 A-7 遥测数据的异常检测结果如图 8 所示。

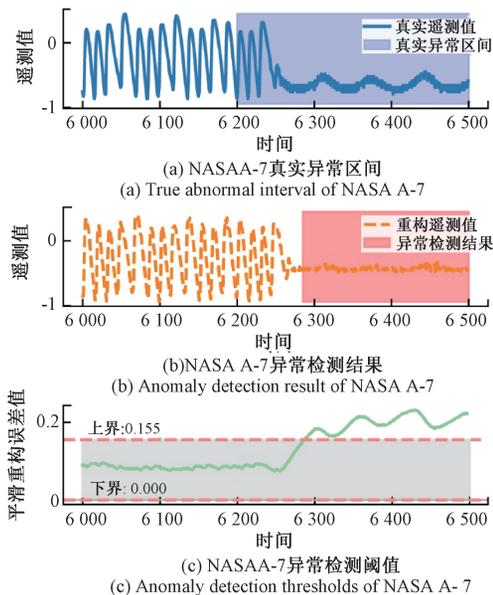


图 8 NASA A-7 数据集异常检测效果图

Fig. 8 Anomaly detection for NASA SMAP A-7 dataset

对于 A-7 数据,根据本文提出的异常判读阈值生成方法设置其异常检测判据为(0~0.155),当某个时刻计

算的平滑重构误差超出这个范围时将被判定为异常。从图 8 中的异常检测结果对比可以看出,人为标记的 A-7 遥测数据的异常区间为第 6 200~6 500 点,所提出方法的异常检测区间为第 6 285~6 500 点。且所提出的方法在对遥测数据进行异常检测的过程中,未有明显虚警情况出现,表明通过引入对比学习方法训练得到的遥测数据重构模型,能够对正常样本和异常样本记性有效的差异化表示,从而降低异常检测结果对阈值的敏感度。

虽然所提出方法能够降低异常检测过程的虚警率,但也仍存在部分异常样本未完全检出的现象。结合图 8(c)中的重构误差曲线可以看出,对于第 6 200~6 250 点,遥测数据自身的变化趋势、变化范围等与正常样本间差异不大,而在对异常样本的人工标注中,也结合了对应时刻的指令对数据进行了联合判读;对于第 6 250~6 285 点,可以看出遥测数据的重构误差处于连续变化的阶段,对应数据的分布特征也未出现显著变化,因此造成此段数据未准确判读。

对于 NASA SMAP 卫星 A-7、D-7 和 D-9 三种遥测数据,所提出方法的异常检测结果统计如表 4 所示。

表 4 遥测数据异常检测结果

Table 4 Anomaly detection results of telemetry data

数据样本	异常样本区间	异常样本检测结果
A-7	(6 200~8 600)	(6 285~8 614)
D-7	(4 940~7 641)	(5 058~7 625)
D-9	(6 250~7 405)	(6 475~7 389)

从表 4 中可以看到,各个数据集上的异常样本检测结果区间与真实的异常样本区间重合率高,且基本上没有出现真实异常样本区间之外的区间被标记,检出率较高而虚警率较低。

上述试验结果表明,本文所提出的方法能够有效学习遥测数据中正常样本和异常样本之间的差异,通过重构未知状态的遥测数据并统计重构误差,有效检测遥测数据中的异常样本,并降低检测结果的虚警率。

3.3 实验结果对比

利用 NASA SMAP 数据集中的 A-7、D-7 和 D-9 遥测数据集,将所提出方法与孤立森林以及 NASA 所提出的 Telemnom 方法进行对比,并通过检出率、虚警率和准确率对各类方法进行对比。利用不同方法对 3 个数据集进行异常检测的结果如表 5 所示。

通过对比不同方法对 NASA SMAP 数据集中 A-7、D-7 和 D-9 遥测数据的异常检测结果可以看出,对于检出率指标(TPR),所提出方法与期望值之间分别仅相差 0.018(A-7)、0.016(D-7)和 0.026(D-9),而 NASA 所提出的 Telemnom 方法在各个数据中达到的效果——与

表5 遥测数据异常检测结果对比

Table 5 Comparison of telemetry data anomaly detection

数据样本	评价指标	指标期望值	孤立森林	Telemanom	本文方法
A-7	TPR	0.287	0.232	0.052	0.269
	FPR	0.000	0.139	0.000	0.002
	ACC	1.000	0.814	0.854	0.988
D-7	TPR	0.353	0.000	0.101	0.337
	FPR	0.000	0.000	0.000	0.000
	ACC	1.000	0.646	0.749	0.985
D-9	TPR	0.156	0.000	0.043	0.124
	FPR	0.000	0.000	0.000	0.000
	ACC	1.000	0.844	0.889	0.970

期望值分别相差 0.235 (A-7)、0.252 (D-7) 和 0.113 (D-9), 本文提出方法的最差效果 0.026 (D-9) 相较于其中的最好效果 0.113 (D-9), 至少提升了 0.087, 说明本文所提出方法能够一定程度上提高异常检出率; 对于误检率而言, 如表 5 所示本文所提出方法在 3 个异常遥测数据集中虚警率 (FPR) 都低于 0.002, 均未出现明显虚警现象, 而孤立森林方法则存在较为明显的虚警现象, 说明本文所提出的方法在保证较高异常检出率的同时也能达到降低虚警率的效果; 检测准确率 (ACC) 方面, 所提出方法相较于孤立森林和 NASA 所提出的 Telemanom 方法均至少提升了 8.1%。

综合上述 3 个方面指标的分析, 可以看出, 本文所提出方法检出率和准确率较高而虚警率较低, 证明所提出的异常检测模型能够有效避免由于异常样本不足而产生的虚警问题。

## 4 结 论

卫星是一类被广泛应用于通信、导航、遥感等民生、国防领域的航天器, 由于其所处工作环境复杂, 在地面站利用遥测数据信息对其工作状态进行监测及时发现异常工作模式并进行相关的调整是十分有必要的, 但针对现有模型仍存在对已有异常样本应用不充分、异常检测模型虚警率高的问题, 本文提出了基于对比序列重构的卫星遥测数据重构和基于 KDE 的异常判读阈值生成方法, 采用对比学习的模型构建方式实现了目标重构遥测时间序列正常区间和异常区间的差异化重构, 基于此又提出了基于 KDE 的异常判读阈值生成方法完成异常判读, 经过实际卫星遥测数据的验证, 该方法在保证一定异常检测精度的前提下, 有效降低了模型的虚警率, 具有良好的实际应用前景。

## 参考文献

- [1] BARBA N, AUSTIN A, BANFIELD D, et al. High science value return of small spacecraft at mars [J]. Planetary Science and Astrobiology Decadal Survey (PSADS), 2020: 2023-2032.
- [2] GE H, LI B, JIA S, et al. LEO enhanced global navigation satellite system (LeGNSS): Progress, opportunities, and challenges [J]. Geo-spatial Information Science, 2022, 25(1): 1-13.
- [3] ZHANG L. Development and prospect of Chinese lunar relay communication satellite [J]. Space: Science & Technology, 2021, 1: 241-254.
- [4] NATIONAL RESEARCH COUNCIL. Limiting future collision risk to spacecraft: An assessment of NASA's meteoroid and orbital debris programs [M]. National Academies Press, 2011.
- [5] JI X Y, LI Y Z, LIU G Q, et al. A brief review of ground and flight failures of Chinese spacecraft [J]. Progress in Aerospace Sciences, 2019, 107: 19-29.
- [6] 庞景月. 基于概率性预测的航天器遥测数据异常检测方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2019.  
PANG J Y. Probabilistic prediction based anomaly detection method for spacecraft telemetry data [D]. Harbin: Harbin Institute of Technology, 2019.
- [7] DECOSTE D. Automated learning and monitoring of limit functions [C]. International Symposium on Artificial Intelligence, Robotics, and Automation in Space, 1997.
- [8] HOWLIN K, WEISSERT J, KRANTZ K. MOORE: A prototype expert system for diagnosing spacecraft problems [C]. NASA, the 1988 Goddard Conference on Space Applications of Artificial Intelligence, 1988.
- [9] HAMILTON M. SCARES: a spacecraft control anomaly resolution expert system [C]. NASA, Marshall Space Flight Center, Second Conference on Artificial Intelligence for Space Applications, 1988.
- [10] HEHER D, POWNALL P. StarPlan: A model-based diagnostic system for spacecraft [C]. NASA, Lyndon B. Johnson Space Center, Third Annual Workshop on Space Operations Automation and Robotics (SOAR 1989), 1990.
- [11] HAZIZA M. Towards an operational fault isolation expert

- system for French telecommunication satellite Telecom 2[J]. *Ground Data Systems for Spacecraft Control*, 1990; 385-392.
- [12] TAKAKI R, HASHIMOTO M, HONDA H, et al. ISACS-DOC: Automatic monitoring and diagnostic system for spacecraft [C]. *Proceeding of the 6th International Symposium on Reducing the Cost of Spacecraft Ground Systems and Operations*, 2005.
- [13] SHANGGUAN D, CHEN L, DING J. A digital twin-based approach for the fault diagnosis and health monitoring of a complex satellite system[J]. *Symmetry*, 2020, 12(8): 1307.
- [14] IVERSON, DAVID L. Inductive system health monitoring[C]. *International Conference on Artificial Intelligence*, 2004.
- [15] YAIRI T, ODA T, NAKAJIMA Y, et al. Evaluation testing of learning-based telemetry monitoring and anomaly detection system in SDS-4 operation [C]. *Proceedings of the International Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS)*, 2014.
- [16] 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.
- [17] MARTÍNEZ-HERAS J A, DONATI A. Enhanced telemetry monitoring with novelty detection [J]. *AI Magazine*, 2014, 35(4): 37-46.
- [18] LI Y, LEI M, LIU P, et al. A novel framework for anomaly detection for satellite momentum wheel based on optimized SVM and Huffman-multi-scale entropy [J]. *Entropy*, 2021, 23(8): 1062.
- [19] LESOUPLE J, BAUDOIN C, SPIGAI M, et al. Generalized isolation forest for anomaly detection [J]. *Pattern Recognition Letters*, 2021, 149: 109-119.
- [20] GE S, JUN L, LIU D, et al. Anomaly detection of condition monitoring with predicted uncertainty for aerospace applications [C]. *Proceedings of 12th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, 2015, 1: 248-253.
- [21] YU W, TAO Z, JIANJIANG H U I, et al. An anomaly detection method for spacecraft solar arrays based on the ILS-SVM model[J]. *Journal of Systems Engineering and Electronics*, 2023, 34(2): 515-529.
- [22] PANG J, LIU D, PENG Y, et al. Collective anomalies detection for sensing series of spacecraft telemetry with the fusion of probability prediction and Markov chain model[J]. *Sensors*, 2019, 19(3): 722.
- [23] WANG Z, GAO Y, PANG J, et al. Satellite power system anomaly detection based on spatio-temporal uncertainty fusion [C]. *2023 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS)*, 2023: 1-5.
- [24] 史欣田, 庞景月, 张新, 等. 基于集成极限学习机的卫星大数据分析[J]. *仪器仪表学报*, 2018, 39(12): 81-91.
- SHI X T, PANG J Y, ZHANG X, et al. Satellite big data analysis based on bagging extreme learning machine[J]. *Chinese Journal of Scientific Instrument*, 2018, 39(12): 81-91.
- [25] BAIREDDY S, CHAN M W, DESAI S R, et al. Spacecraft time-series online anomaly detection using extreme learning machines [C]. *2022 IEEE Aerospace Conference (AERO)*, 2022: 1-9.
- [26] HUNDMAN K, CONSTANTINOU V, LAPORTE C, et al. Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding [C]. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018: 387-395.
- [27] ZENG Z, JIN G, XU C, et al. Satellite telemetry data anomaly detection using causal network and feature-attention-based lstm [J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-21.
- [28] YANG L, MA Y, ZENG F, et al. Improved deep learning based telemetry data anomaly detection to enhance spacecraft operation reliability [J]. *Microelectronics Reliability*, 2021, 126: 114311.
- [29] KANG S, YANG L, SONG Y, et al. Satellite power system state prediction based on online learning with parameter association rules [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023(72): 2521613.

- [30] XIE L, PI D, ZHANG X, et al. Graph neural network approach for anomaly detection [J]. *Measurement*, 2021, 180: 109546.
- [31] YU J, SONG Y, TANG D, et al. Telemetry data-based spacecraft anomaly detection with spatial-temporal generative adversarial networks [J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-9.
- [32] DE RYCK T, DE VOS M, BERTRAND A. Change point detection in time series data using autoencoders with a time-invariant representation [J]. *IEEE Transactions on Signal Processing*, 2021, 69: 3513-3524.
- [33] 彭喜元, 庞景月, 彭宇, 等. 航天器遥测数据异常检测综述 [J]. *仪器仪表学报*, 2016, 37(9): 1929-1945.
- PENG X Y, PANG J Y, PENG Y, et al. Review on anomaly detection of spacecraft telemetry data [J]. *Chinese Journal of Scientific Instrument*, 2016, 37(9): 1929-1945.
- [34] KINGMA D P, WELING M. An introduction to variational autoencoders [J]. *Foundations and Trends® in Machine Learning*, 2019, 12(4): 307-392.
- [35] HADSELL R, CHOPRA S, LECUN Y. Dimensionality reduction by learning an invariant mapping [C]. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 2006, 2: 1735-1742.
- [36] CHICCO D. Siamese neural networks: An overview [J]. *Artificial Neural Networks*, 2021: 73-94.

## 作者简介



李桢煜, 2022 年于山东大学获得学士学位, 现为哈尔滨工业大学自动化测试与控制研究所硕士研究生, 主要研究方向为数据分析、异常检测和机器学习等。

E-mail: hit\_lizhenyu@163.com

**Li Zhenyu** received her B. Sc. degree in 2022 from Shandong University, now she is a master in Automatic Test and Control Institute, Harbin Institute of Technology. Her main research interests include data analysis, anomaly detection and machine learning, etc.



宋宇晨 (通信作者), 2015 年于电子科技大学获得学士学位, 2022 年于哈尔滨工业大学获得博士学位, 现为哈尔滨工业大学助理教授, 主要研究方向为复杂系统智能感知与评估、智能测试信息处理、卫星星座状态

监测与运维等。

E-mail: songyuchen@hit.edu.cn

**Song Yuchen** (Corresponding author) received his B. Sc. degree in 2015 from University of Electronic Science and Technology of China, received his Ph. D. degree in 2022 from Harbin Institute of Technology, now he is an assistant professor in Harbin Institute of Technology. His main research interests include complex system intelligent sensing and evaluation, intelligent testing information processing, satellite constellation state monitoring and operating.