电力负荷时间序列分析的异常值检测方法

石 蓉[1] 尹 璐[1]
SHI Rong YIN Lu

摘 要 为有效提高电力系统的运行效率和经济性，文章提出一种电力负荷时间序列分析的异常值检测方法。通过对电力负荷时间序列数据实施聚类与标记操作，以此作为异常值检测模型的输入数据。运用卷积神经网络与 DCE-LAE 模型，构建并训练异常值检测模型，最终输出异常值的检测结果。实验数据表明，所提方法在准确性、检测效率等评估维度上均表现优异，可高效且精准地检测出时间序列中的异常值，能够为电力系统的安全稳定运行和科学决策提供有力支持。

关键词 异常值检测；时间序列；能源管理；电力系统；电力负荷

doi：10.3969/j.issn.1672-9528.2026.01.013

0 引言

电力负荷作为电力系统运行的核心指标之一，反映了在特定时刻电力用户对电能的需求总量。准确掌握电力负荷的变化规律，对于电力系统的规划、调度以及能源管理等诸多方面都具有不可忽视的重要意义。电力负荷数据通常以时间序列的形式呈现，具有显著的时间相关性和动态变化特征。在电力负荷时间序列中，异常值的存在是一个常见但又难以忽视的问题。这些异常值会严重扭曲电力负荷的真实变化趋势，干扰对负荷规律的准确把握，进而影响电力系统的决策制定和运行控制。因此，研究有效的电力负荷时间序列异常值检测方法具有重要的现实意义。

文献 [1] 通过对异常检测模型进行设计，该模型的编码器部分由挤压激励注意力模块与概率稀疏自注意力模块共同构成。结合稀疏自注意力机制，进而成功捕获时间序列中的长期依赖关系。不过，此方法在实际应用中存在检测效率不高的问题。文献 [2] 则将 Transformer 网络作为生成对抗网络的基础模型，还运用 patch 技巧，使模型能够敏锐捕捉时间窗口内的异常细节信息。然而，该方法在检测精度方面表现欠佳。

本文基于机器学习算法中的监督学习理念，利用已标记好的数据来构建异常值检测模型，进而输出电力负荷时间序列的异常值结果。

1 电力负荷时间序列分析的异常值检测方法的设计

1.1 电力负荷时间序列数据的聚类与标记处理

鉴于电力负荷数据规模极为庞大，在对电力负荷时间序列开展聚类分析工作之前，对原始时间序列数据开展预处理工作是极为关键且不可或缺的步骤 [3]。鉴于电力负荷曲线具有多样性特征，有必要将数据划分为工作日数据和非工作日数据这两大类别。此外，鉴于本次聚类过程是基于时间序列之间的相似性来开展，由于并非关注数据点间的绝对距离情况，所以有必要对时间序列数据实施归一化处理。在完成上述预处理操作后，采用滑动窗口方法分步对预处理后的数据实施聚类分析。

在识别非标准时间序列间隔时，需要分别计算各个类簇的聚类评估值以及每个数据点的距离评估值。此时，聚类评估值主要用于衡量类簇中包含的数据点数量，其关键作用在于识别出数据量较少的类簇——这些类簇往往与异常时间序列具有相关性。该聚类评估值的数学公式为：

$$\mathrm{score}(\alpha)_i = \sum_{j=1}^{N} j \tag{1}$$

式中：N 代表数据量；$\mathrm{score}(\alpha)_i$ 代表第 i 个类簇的聚类分数。

距离分数这一指标表示数据点与某类簇中心点的距离同该簇内全部距离均值的比例关系，其计算公式为：

$$\mathrm{score}(\chi)_j = \frac{\chi_j}{\frac{1}{N}\sum_{j=1}^{N}\chi_j} \tag{2}$$

式中：χ_j 代表第 j 个数据到当前类簇中心的距离；$\mathrm{score}(\chi)_j$ 代表第 i 个类簇中第 j 个数据的距离分数。

聚类分数与距离分数的计算过程，均聚焦于滑动窗口内的各个数据点以及它们所归属的簇。在完成相关计算工作之后，依据这两个分数来构建消费端分数，此消费端分数可表示为：

$$\mathrm{score}_j = \mathrm{score}(\alpha)_i \cdot \mathrm{score}(\chi)_j \tag{3}$$

1. 国网陕西营销服务中心（计量中心） 陕西西安 710075

基于计算得出的消费端分数，能够高效地把每个数据划分成少数类和多数类，随后依据分类情况对数据进行标记。

1.2 异常值检测模型的建立与训练

依据上述已标记完成的电力负荷时间序列数据，将其作为异常值检测模型的输入。鉴于卷积神经网络能够深入挖掘电力负荷数据中隐含的时序特性[4]，可借助这一特性构建异常值检测模型。此模型不仅能降低数据维度，还能有效提升检测效率。

针对电力负荷时间序列数据，卷积神经网络架构中的DCE编码器主要负责完成深层时序特征的捕捉与提取工作。编码模块末端的隐藏状态将被用作初始状态，传递至LAE编码器，以推进隐藏状态的深度学习进程。在DCE解码器之后，引入一个神经元数量和预测步长相匹配的TimeDistributed封装层，以此达成预测结果的输出，生成与输入数据特征相近的检测曲线。在该异常值检测模型中，编码器与解码器将协同开展训练过程，借助持续减小模型综合损失的方式，最终得到参数配置最优的模型。

设定 $X=[x_1, x_2,\cdots, x_n]$ 为经过标记处理后的电力负荷时间序列，其卷积过程为：

$$H'=\partial\left(X*\boldsymbol{d}^{p}+\boldsymbol{b}\right) \tag{4}$$

式中：∂代表激活函数；$\boldsymbol{d}^{p}$代表DCE编码器中的一维卷积核；*代表一维卷积符号；$\boldsymbol{b}$代表偏置向量。

其反卷积过程为：

$$H=\partial'\left(k^{\beta}*\widetilde{\boldsymbol{d}}^{p}+\boldsymbol{b}'\right) \tag{5}$$

式中：k^{β}代表反卷积的编码后数据；$\widetilde{d}^{p}$代表反卷积的DCE解码器中的一维卷积核；∂'代表反卷积的激活函数；$\boldsymbol{b}'$代表反卷积的偏置向量。

为了使得DCE-LAE模型生成的目标概率分布能够与输入的时序卷积层分布相匹配，防止卷积嵌入的精细调整引发检测空间的畸变现象[5]，本文采用KL散度作为卷积嵌入层的损失度量指标。基于此，该异常值检测模型的具体表达公式为：

$$F_{\mathrm{KL}}=\left(H\|H'\right)=\sum_{i=1}^{n}H_i\log\left(\frac{H_i}{H_i'}\right) \tag{6}$$

KL散度在衡量数据分布差异方面展现出卓越的性能，若两个数据分布之间的差异程度越低，KL散度的计算结果便会随之降低。最终，DCE-LAE网络整体采用的损失函数具体表达式为：

$$f=f_0+\eta f_{\mathrm{KL}} \tag{7}$$

式中：f_0代表检测损失；f_{KL}代表卷积损失；η代表控制扭曲嵌入层系数。

模型构建完毕后，其训练步骤如下：

（1）针对不包含检测损失项的卷积网络实施预训练操作，通过该过程获取降维处理后的深层特征。将提取的深层特征输入到LAE编码器，使其能够学习输入特征在各个时间步所对应的隐藏状态。

（2）将最终时间步的隐藏状态作为初始状态进行设置，并将其馈入LAE解码器。随后，借助反卷积操作来执行时间序列检测任务。

（3）将η值设定为0.1，并采用持续最小化联合损失函数f的方法，针对DCE-LAE模型开展迭代式的微调优化操作。在模型训练阶段，每当完成5次迭代计算，便执行一次联合损失的计算。

（4）当相邻两次计算所获得的损失值变化幅度低于预先设定的阈值，当迭代次数达到预设的100次这一限定阈值时，训练流程即刻终止，同时DCE-LAE模型停止训练。

1.3 输出异常值检测结果

对于已训练好的电力负荷时间序列样本，编码器会将其输入序列压缩为固定维度的特征向量。实际应用中，时间序列的异常状态通常难以提前识别。为应对这一挑战，编码器采用了随机隐式解密编码机制。训练时，输入时间序列对应的解密编码若具有正常序列特征，则表明其属于正常时序数据；反之，若解密编码序列出现混乱等异常，则提示该输入存在异常时序数据。此时需激活模型的异常检测模块，并重构多标签持续参数误差模型。

创建长度为r的滑动时间窗口，其数据点表示为单位向量。基于数据标签特征，将训练数据集划分为多个长向量，并对其进行标记，形成子时间序列段 $X'=[x_{1:r}, x_{2:r+1},\cdots, x_{n-r+1:n}]$。鉴于包含隐含解密代码的多标签时序数据处于加密状态，直接检测异常值较为困难。因此，将这些带有随机隐含解密代码的数据输入编码器进行解密，将其转换为异常检测所需的训练空间向量 $G_{m,\,m+r}$。利用电力负荷时间序列训练模型后，将待测时序数据输入异常检测模型，最终输出时间序列的异常检测结果：

$$F_{\mathrm{KL}}\left(G_{m:m+r},\mathrm{score}_j\right)=\lg\frac{f\left(G_{m+r},\mathrm{score}_j\right)}{f\left(G_m,\mathrm{score}_j\right)}\circ\eta\left(X'\right) \tag{8}$$

式中：$f(G_m, \mathrm{score}_j)$代表隐含解密代码时序参数的检测误差；$f(G_{m+r}, \mathrm{score}_j)$代表输入电力负荷时间序列参数的检测误差。

本研究对多种异常值检测模型开展对比分析，其核心目标是度量异常估计时间序列相对于原始训练样本序列的检测误差。基于计算得到的误差值，将其转换为异常检测比率，再与原始序列预设的空间阈值作比较。异常检测比率若超过阈值，则判定相应点为异常。同时，检测误差的增大意味着向量序列中异常值超出空间阈值的程度越大（二者正相关），

误差值越大，表明异常值相对于空间阈值的偏离程度越高。最终，通过从异常值检测模型中提取异常数值，再回溯到原始数据中找出异常值对应的电力负荷数据来源，就能精准且高效地识别海量电力负荷数据时间序列中的异常值。

2 实验测试与分析

2.1 实验准备

为验证本文设计的电力负荷时间序列异常值检测方法的有效性，通过与文献 [1-2] 方法进行对比，评估本文方法在检测精度、检测效率等指标上的表现，同时考察在不同数据规模下的适用性。

本文方法的仿真及对比测试均在表 1 的实验环境中进行。

表 1 实验参数

配置	参数
实验框架	TensorFlow3.0
CPU	8 核 32 GB
编程语言	Python
操作系统	Windows8.0
GPU	1×P200

依托上述的实验环境，本次测试所采用的数据集源自某供电台区的历史供电负荷数据与用电负荷数据。该数据集的采集时间间隔设定为 15 s，时间跨度达 120 天，涵盖了波动异常与极值异常这两种典型的负荷特征。并依据 7:3 的比例，对数据集进行随机划分操作，将其拆分为训练集与测试集。

2.2 实验结果与分析

基于上述实验准备，首先，以供电负荷数据的极值异常为例，进行时间序列异常值检测分析，并分别使用本文方法与文献 [1-2] 方法进行对比，其结果如图 1 所示。

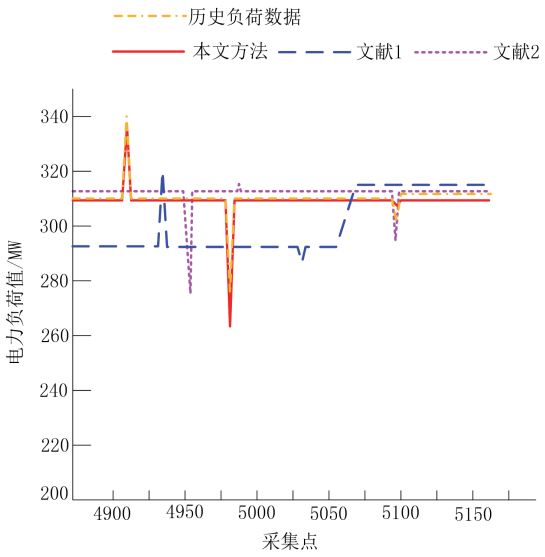

图 1 供电负荷数据时间序列异常值检测结果

根据图 1 结果可以能够清晰地观察到 3 种方法的检测序列与历史负荷数据之间存在显著差异。基于此，可明确得出结论：本文所提出的方法具备精准检测并识别供电负荷数据时间序列中极值异常的能力。

其次，对 3 种方法的异常值检测效率进行测试分析，其结果如图 2 所示。

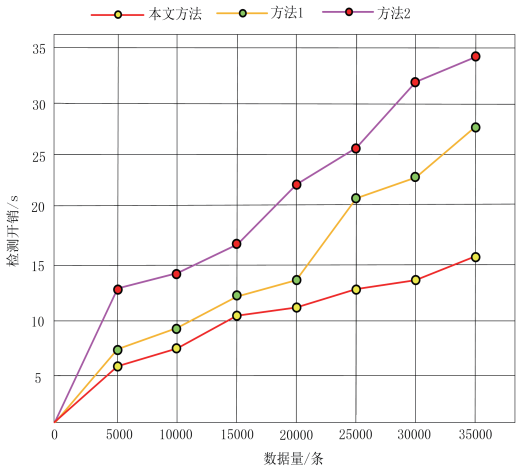

图 2 检测效率对比

综合上述实验结果可以看出，本文方法在准确性、检测效率等评估指标上均表现出色，能够有效地检测出时间序列的异常值。与其余两种方法相比，本文方法具有更强的适应性和准确性，能够更好地应对电力负荷时间序列的非线性、非平稳性以及复杂异常模式的特点。因此，电力负荷时间序列分析中的异常值检测方法在实际电力系统的异常值检测中具有较高的应用价值，能够为电力系统的安全稳定运行和科学决策提供有力支持。

3 结语

电力负荷时间序列分析中的异常值检测是一项兼具挑战性与重要性的工作，其对保障电力系统平稳运行、优化电力资源调配以及提升电力服务水准而言，意义不可小觑。在实际应用中，需要根据具体的电力负荷数据特点、应用场景以及资源条件等因素，综合选择合适的异常值检测方法。对于数据量较小、分布相对简单的场景，基于统计的方法可能是一个不错的选择。而对于数据量大、分布复杂且对检测精度要求较高的场景，基于机器学习的方法可能更具优势。此外，还可以考虑将多种方法进行融合，以充分发挥各自的优势，提高异常值检测的准确性和可靠性。未来随着电力系统的不断发展和智能化水平的提升，电力负荷数据将更加丰富和复杂，对异常值检测的要求也将越来越高。需要进一步探索更加高效、准确、可解释性强的异常值检测方法，结合大数据、人工智能等先进技术，不断提升电力负荷时间序列分析的水平，为电力系统的安全稳定运行提供更有力的支持。

基于 SoC 的 GNSS 信号采集回放系统设计

于嘉杰[1] 宋剑鹏[1] 夏文龙[1] 郭庆功[1*]
YU Jiajie SONG Jianpeng XIA Wenlong GUO Qinggong

摘 要 为满足高精度导航接收机测试及算法验证对卫星信号采集与复现能力的需求，文章设计了一种基于 SoC 的全球导航卫星系统信号采集与回放系统。系统通过直接射频采样和数字信号处理技术，不依赖传统的模拟中频结构，在宽带条件下实现了对多种卫星信号的同时采集与高保真回放，并在可编程数字域中灵活完成下变频和信号处理。实验测试表明，该系统能够稳定完成 GPS L1CA 和 BDS B1I 等典型导航信号的采集与回放任务，模拟真实导航信号环境，配合 Pocket SDR 软件接收机可实现采集信号解调与米级定位精度。

关键词 射频直采；Pocket SDR；信号采集回放；以太网；SDCARD

doi：10.3969/j.issn.1672-9528.2026.01.014

0 引言

现代全球导航卫星系统（global navigation satellite system, GNSS）不仅具有军事战略意义，也带来了显著的社会经济效益，推动各国加快自主导航系统建设[1]。随着导航系统应用的不断扩大，导航终端种类增多、规模提升，对其功能与性能的测试变得至关重要。导航终端测试已成为导航应用的核心技术，对推动产品标准化和提升经济社会效益具有重要意义。

在实际环境中，GNSS 信号常受到射频干扰、电离层扰动和多径效应等因素影响，具有随机性和时变性，难以实现信号传播条件的完全复现。为了能够在可控环境下对导航接收机进行系统性测试与性能评估，亟需一种能够记录并复现实际信号的手段。GNSS 信号采集与回放系统正是在此背景下发挥关键作用：该系统可采集实际环境中的信号并保存为数据文件，反复回放真实射频场景，为接收机提供稳定、可控的测试输入，同时也便于对接收机算法进行对比与验证，提升其准确性与可靠性[2]。

近年来，GNSS 信号采集与回放系统因其在导航接收机测试与算法验证中的重要作用，逐渐成为国内外研究的热点方向。现有系统在架构设计上主要可分为基于传统超外差结构的中频采样方式[3-7]和基于新型高速数据转换器的直接射频采样方式[8-12]。2012 年，Di 等[3]基于 USRP-N210 主板与配对的 RFX 子板，设计了一款低成本、便携且灵活的 GNSS 信号采集与回放系统。2023 年，于得水[7]利用 AD9361 与

1. 四川大学电子信息学院 四川成都 610021
[基金项目] 电信科学技术第五研究所项目“分布式时空基准测试设备 WBZB01”（Y23-Z-S81-1）

参考文献：

[1] 杨超城，严宣辉，陈容均，等 . 融合双重注意力机制的时间序列异常检测模型 [J]. 计算机科学与探索，2024, 18(3):740-754.

[2] 蔡美玲，汪家喜，刘金平，等 . 基于 Transformer GAN 架构的多变量时间序列异常检测 [J]. 中国科学 : 信息科学，2023, 53(5): 972-992.

[3] 衡红军，刘静 . 基于混合方法的多维时间序列驾驶异常点检测 [J]. 计算机工程，2020,46(3):99-104.

[4] 王苏杭，王一棠，庞勇，等 . 面向盾构机实测数据的滑动窗口分层异常值检测及修正方法 [J]. 仪表技术与传感器，2023(9): 93-99.

[5] 王志坚 . 一种 GARCH 模型异常值的稳健检测法及其应用 [J]. 统计与决策，2020,36(10):41-44.

【作者简介】

石蓉（1976—），女，陕西安康人，本科，高级工程师，研究方向：电力营销。

尹璐（1990—），女，山西灵石人，本科，工程师，研究方向：负荷管理。

（收稿日期：2025-07-07 修回日期：2025-12-16）