

DOI:10.12404/j.issn.1671-1815.2303584

引用格式: 易爽, 贺俊杰, 郑胜, 等. 基于VAE的核电运行状态监测方法[J]. 科学技术与工程, 2024, 24(19): 8109-8114.

Yi Shuang, He Junjie, Zheng Sheng, et al. VAE-based operation state monitoring of nuclear power plants[J]. Science Technology and Engineering, 2024, 24(19): 8109-8114.

## 原子能技术

# 基于VAE的核电运行状态监测方法

易爽<sup>1</sup>, 贺俊杰<sup>2</sup>, 郑胜<sup>1,2\*</sup>, 杨森权<sup>3,4</sup>, 曾曙光<sup>2</sup>

(1. 三峡大学电气与新能源学院, 宜昌 443002; 2. 三峡大学理学院, 宜昌 443002;

3. 中核武汉核电运行技术股份有限公司, 武汉 430074; 4. 中核集团核工业仿真技术重点实验室, 武汉 430074)

**摘要** 由于核反应堆发电的特殊性, 核电厂对于生产安全的敏感度远胜于普通电厂。作为日常运维的重要环节, 核电机组运行状态监测, 对于核电厂的安全稳定运行具有重要意义。当前核电机组状态监测主要采用预设固定阈值报警结合人工巡检的方式, 这种方式无法发现低于报警阈值的异常状态, 同时存在一定程度的漏报风险。核电运行数据作为高维海量时序数据, 具有正常样本和异常样本分布不均以及数据缺乏标签的问题, 这限制了有监督深度学习方法的使用。提出了一种基于变分自编码器 (variational autoencoders, VAE) 构建的无监督深度学习模型对真实运行数据进行异常检测, 通过正常运行数据学习正常模式下数据在隐空间的分布, 并基于异常数据无法被良好重构的原理, 通过重构误差的大小来判别当前状态是否异常。实验以核电机组化学和容积控制系统 (removal-chemical and volume control system, RCV) 中的上充泵为例, 使用真实运行数据结合插入异常的方式对模型进行了验证, 并与经典机器学习方法进行了对比。实验结果表明基于变分自编码器的模型能够有效检测到核电真实数据中的异常数据片段及离群点, 检测精确率和召回率均高于90%, 检测性能相对孤立森林和支持向量机等经典机器学习算法具有优势, 具备一定的实用价值和研究意义。

**关键词** 核电; 运行状态; 异常检测; 变分自编码器

**中图分类号** TL48;

**文献标志码** A

## VAE-based Operation State Monitoring of Nuclear Power Plants

YI Shuang<sup>1</sup>, HE Jun-jie<sup>2</sup>, ZHENG Sheng<sup>1,2\*</sup>, YANG Sen-quan<sup>3,4</sup>, ZENG Shu-guang<sup>2</sup>

(1. College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China;

2. College of Science, China Three Gorges University, Yichang 443002, China;

3. China Nuclear Power Operation Technology Corporation, Wuhan 430074, China;

4. Key Laboratory of Nuclear Industry Simulation Technology of CNNC, Wuhan 430074, China)

**[Abstract]** Due to the unique power generation method used in nuclear reactors, nuclear power plants are more safety-sensitive compared to conventional power plants. Therefore, daily monitoring of the operating state of nuclear power units is critical for ensuring operational safety. Currently, status monitoring of nuclear power plants is conducted through automated alarms with preset fixed thresholds, and manual supervision. However, this method cannot detect anomalies below the alarm thresholds, which may lead to risks of underreporting. Nuclear power operational data, characterized by high-dimensional time series, faces challenges of imbalanced distributions between normal and abnormal samples as well as the lack of labeled data. These factors limit the application of supervised deep learning methods. An unsupervised deep learning model based on variational auto-encoders (VAE) was proposed for anomaly detection in real operational data. This model learned the distribution of data in the latent space under normal operating conditions and relied on the principle that abnormal data cannot be reconstructed effectively. Anomalies were detected by evaluating the magnitude of reconstruction error. The experiment focused on the upper charging pump in the removal-chemical and volume control system (RCV) of a nuclear power plant. It involved the validation of the model using real operational data with deliberately inserted anomalies and compared it to classical machine learning methods. The results of the experiment show that the model based on variational autoencoders effectively detects abnormal data segments and outliers in real nuclear power plant data. It achieves precision and recall rates both exceeding 90%. In terms of detection performance, it outperforms classical machine learning algorithms like isolation forest and support

收稿日期: 2023-05-16; 修订日期: 2023-12-21

基金项目: 中核集团核工业仿真技术重点实验室对外开放基金 (B220631)

第一作者: 易爽 (1989—), 男, 汉族, 湖北宜昌人, 博士研究生。研究方向: 核电运行数据异常检测。E-mail: yishuang@ctgu.edu.cn。

\* 通信作者: 郑胜 (1965—), 男, 土家族, 湖北恩施人, 教授。研究方向: 人工智能, 图像处理, 异常检测。E-mail: zsh@ctgu.edu.cn。

vector machine(SVM). This demonstrates its practical value and research significance.

[**Keywords**] nuclear power; operation status; anomaly detection; VAE

随着核电厂工业水平的不断进步,工业规模的不断扩大,数字化、智能化的核电运维正成为日益重要的发展方向。为了降低事故发生的风险,保证核电厂设备运行过程的安全稳定运行,确保运行状态满足给定的性能指标,迫切需要对核电机组运行过程实现高效可靠的在线监控和异常检测<sup>[1]</sup>,从而在异常状态发生时及时预警并处理以避免或减少可能带来的损失。

当前,核电厂主要通过监控与数据采集(supervisory control and data acquisition, SCADA)系统<sup>[2]</sup>以及集散控制系统(distributed control system, DCS)的监测数据对各类原因引起的变量超限进行监控,当测量值超过预设的固定报警阈值时,便发出警报,由监盘人员进行判断和进一步处理。然而,预设报警阈值的方式无法发现低于阈值的异常状态,同时,作为针对单维数据的检测方式,不完全适用于多传感器系统的状态偏移检测<sup>[3]</sup>。因此,基于核电真实运行数据的特点,开发有效可靠的异常检测模型对于核电机组日常运行状态实现状态监测非常必要且具有实用价值。

异常检测是对运行过程中数据模式是否正常的判别,也是故障诊断的起点和基础<sup>[4]</sup>。国内外学者对于工业过程异常检测的理论以及方法进行了广泛的研究,尤其是基于数据驱动的方法,逐渐成为相关领域的研究热点<sup>[5-6]</sup>。数据驱动的异常检测方法主要可以分为基于统计的方法,基于经典机器学习的方法,以及深度学习方法。基于统计的方法如主成分分析方法(principal components analysis, PCA)<sup>[7-9]</sup>、偏最小二乘法(partial least squares, PLS)<sup>[10-11]</sup>、多元统计过程控制方法(multivariate statistical process control, MSPC)<sup>[12]</sup>等,往往对于目标对象的数据分布具有一定的假设,如线性系统、高斯分布、单一模态等等,真实工业场景往往无法满足这些假设,导致会有较多误报和漏报发生。经典机器学习的方法,如KNN( $K$  nearest neighbors)<sup>[13]</sup>、SVDD(support vector domain description)<sup>[14]</sup>、SVM(support vector machine)<sup>[15]</sup>等,基于距离、密度、分类等规则将异常数据与正常数据区分开,从而实现异常检测。但是对于高维海量数据,经典机器学习方法可能会受限于计算复杂度和计算成本等问题。近年来,深度学习在学习复杂结构数据的特征方面展现出了强大的能力,大量的深度异常检测方法已经被应用于各类具有挑战性的实际问题,并显示出明显优于传统异常检测方法的性能。

变分自编码器(variational auto-encoder, VAE)是一种无监督深度学习方法,该方法可以将输入的多维时序数据映射到一个隐空间,并将采样得到的隐变量重构为原始输入,基于异常数据重构为异常数据的概率很低的假设,通过重构误差就可以判断输入数据是否正常,实现状态监测。核电领域的运行数据异常检测是核电产业数字化和智能化发展的重要任务之一,对于提升核电运行安全性,实现自动化智能监盘具有重要意义。然而,核电运行数据属于高维海量时序数据,同时存在着正负样本不均衡以及缺乏数据标记的问题。VAE方法作为融合了概率分布估计以及深度学习降维的方法,已被诸多研究证明对于解决这些问题非常有效<sup>[16-17]</sup>。现基于变分自编码器构建了运行数据异常检测模型,通过正常运行数据训练得到正常模式下的隐变量的分布,建立检测模型。能够有效解决核电领域正常样本多、异常样本少,且缺乏数据标签的问题。由于特征降维和重构概率估计的能力,基于VAE的异常检测模型能够及时有效检测到数据中的异常。

## 1 问题描述

核电系统具有数量众多、种类不同的传感器对机组运行的各项参数进行实时监测,在长久运行过程中积累了大量的高维时序历史数据。由于测点众多且不同系统、不同设备间的数据特征存在较大差异,通常会将按照机组-系统-设备-传感器的层级去对状态监测任务进行细分,分模块实现。以本文研究的化学与容积控制系统中的上充泵为例,上充泵属于设备层级,结合机理结构选择设备相关的传感器(如温度、电流、振动等)来组成设备的相关参数并获取对应的历史运行数据。

工业过程状态监测或异常检测任务的目标,是发现与正常工况或正常数据模式不一致的监测数据,并结合实际情况制定相应的报警策略。从数据驱动的算法层面考虑,只要是数据特征(如幅值、梯度、趋势、相关性等)与正常模式不同的数据点或数据片段,都应当被识别为“异常”,而无需考虑这种异常是否等同于实际需要产生报警的真实异常工况。数据特征层面的异常,根据其来源可以被分为三类,如图1所示。其中,数据错误指的是因为数据采样、传输过程中由于噪声、扰动、数据传输等因素导致的测量数值异常但真实设备状况无异常的情况,如数据缺失,少量离群点、脉冲噪声等。这一类情况在真实数据中很常见,可以通过数据清洗和预

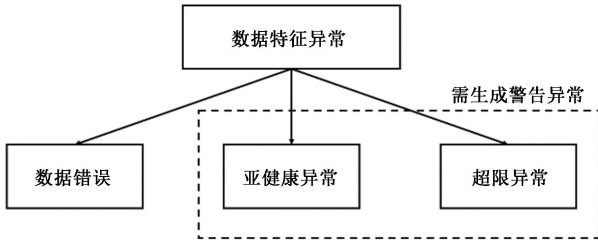


图1 异常数据来源分类

Fig.1 Classification of anomalies by source

处理进行修复,在实时监测中无需产生告警;亚健康异常和超限异常是针对当前预设固定报警阈值的监测方式来分类的,即低于报警阈值但设备状态已经产生了异常的状态被称为亚健康状态,高于报警阈值的则被称为超限异常。亚健康状态可能是不那么“严重”的异常,也可能是超限异常的初期状态。对于亚健康异常的监测,是智能运维发展的重要方向,相较于传统阈值超限报警,可以检测到更微小的异常,并提早实现预警。

由于数据量庞大,标记成本高,大多数情况下历史运行数据是无标签数据。但是基于正常运行数据远多于异常数据这一基本假设,可以通过数据特征和变化趋势,结合专家经验和数据预处理,得到相对“正常”的历史数据集并用于训练无监督异常检测模型。在实际检测过程中,为了实现实时监测,降低计算复杂度,通常会使用时间窗口而不是完整序列作为模型输入,异常检测任务可以被概括如下:对于某一  $d$  维多传感器系统采集的时间序列数据记作:  $\mathbf{X} = \{\mathbf{x}^{(i)} \in \mathbf{R}^{d \times n}, i \in \{1, 2, \dots, n\}\}$ , 其中  $\mathbf{x}^{(i)} = \{x_1^{(i)}, x_2^{(i)}, \dots, x_d^{(i)}\}$  表示某一时刻的多维样本。选取尺寸为  $W$  的时间窗口作为输入,送入通过正常数据训练的无监督异常检测模型,并最终输出相应输入对应的标签  $C_i \in \{0, 1\}$ , 以实现异常状态的检测。

## 2 VAE 异常检测模型

### 2.1 变分自编码器原理

变分自编码器是一种有向概率图模型 (deep probabilistic graphical model, DPGM), 它利用一个神经网络来估计后验概率, 形成了类似于自编码器

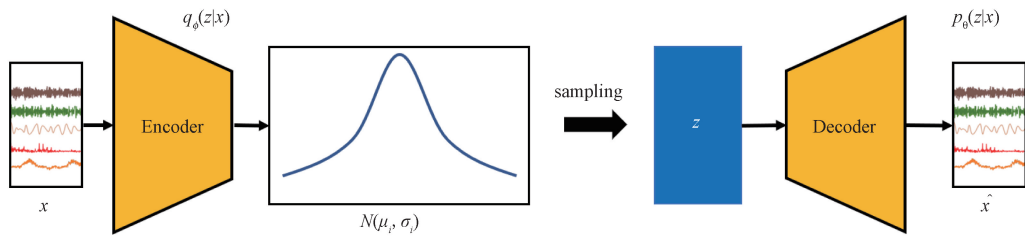


图2 变分自编码器的结构

Fig.2 Structure of variational auto-encoder

(auto-encoder, AE) 的结构。与传统的自编码器相比,VAE除了具有自编码器对输入数据进行编码和解码的能力外,还引入了概率推断的思想,将输入向量编码为隐空间的概率分布,并在此概率分布上随机采样并解码重构得到与原始输入相对应的重构概率。与自编码器使用重构误差判断异常不同,VAE通过概率编码器建模的是隐变量分布的参数而非输入信号,因此具有更强的可变性和泛化表达能力。VAE的结构如图2所示,是与自编码器类似编码-解码结构,同时包含一个重采样过程。对于一个给定的  $d$  维样本时间序列数据集  $\mathbf{X} = \{\mathbf{x}^{(i)} \in \mathbf{R}^{d \times n}, i \in \{1, 2, \dots, n\}\}$ , 输入变量  $x$  编码器转换为隐变量  $z$  的后验概率分布记为  $q_\phi(z|x)$ , 其中  $\phi$  代表该分布的参数,可以用均值  $\mu_x$  和方差  $\sigma_x^2$  描述并满足标准正态分布假设。通过对该分布随机采样,得到隐变量  $z$ , 并由解码器解码得到重构输出  $\hat{x}$ 。

VAE的训练过程就是为了使  $q_\phi(z|x)$  尽可能接近真实分布  $p_\theta(z)$ , 其中  $\theta$  代表该分布的参数。要使似然函数最大化,则

$$p_\theta(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(n)}) = \log p_\theta(\mathbf{x}^{(i)}) \quad (1)$$

变换可得

$$\log p_\theta(\mathbf{x}^{(i)}) = D_{KL}[q_\phi(z|\mathbf{x}^{(i)}, p_\theta(z|\mathbf{x}^{(i)}))] + \mathcal{L}(\theta, \phi; \mathbf{x}^{(i)}) \quad (2)$$

式(2)中:  $D_{KL}$  为两个分布的散度;  $\mathcal{L}$  为分布参数的目标函数。

由于似然函数的优化较难,通常转换为对似然函数下界的优化,即变分推断中的 ELBO (evidence lower bound), 记作  $E$ 。则有

$$\log p_\theta(\mathbf{x}^{(i)}) = -D_{KL}[q_\phi(z|\mathbf{x}^{(i)}, p_\theta(z))] + E_{q_\phi(z|\mathbf{x}^{(i)})} \log p_\theta(\mathbf{x}^{(i)}|z) \quad (3)$$

基于以上推断,VAE的损失函数同时使用重构损失和KL散度损失。损失函数定义为

$$\text{Loss} = D_{MSE}(\mathbf{X}, \hat{\mathbf{X}}) + D_{KL}[q(z|x) || p(z)] \quad (4)$$

该损失函数主要由两部分组成,重构损失描述的是重构数据  $\hat{\mathbf{X}}$  和原始数据  $\mathbf{X}$  之间的差异,用均方误差 (mean square error, MSE) 来衡量;而 KL 散度损



失则是衡量真实分布  $p(z)$  与  $q(z|x)$  间相似度的指标,可以通过最小化 KL 散度来使两个分布尽可能相似。对于通过正常数据训练好的 VAE 模型,当送入异常数据时,异常点的重构概率小,重构效果相对较差,因而能实现异常检测。

## 2.2 检测流程及实现

依据 VAE 的算法原理和核电运行状态监测的需求,可将模型的检测流程归纳如图 3 所示。

由于 SCADA 系统采集的不同传感器数据的量纲和采样周期等存在差异,同时数据采集和传输过程中也会存在缺失、重复等问题,因此需要进行数据清洗和预处理,得到数据长度一致、幅值范围统一的正常运行数据,用于模型训练。训练得到的异常检测模型用于对处理后的实时数据进行检测,判断数据是否异常,并结合报警策略生成告警。

## 3 实验

### 3.1 实验数据

本次实验所采用的数据是由 1 000 MW 压水堆核电机组的 SCADA 系统采集的真实核电运行数据,由于核电厂各类传感器参数众多,为了便于分析,实验部分选用了核电厂的化学和容积控制系统 (removal-chemical and volume control system, RCV) 中上充泵设备的 17 个相关传感器的历史数据,其时间长度约为 3 800 min,采样间隔统一为 1 min。传感器包括电流、温度、振动三类信号。在选择的历

史数据中,包含了上充泵的不同工况,即:泵运行、停泵、泵停运三种工况,其中停泵为瞬态工况,其他为稳态工况。

图 4 所示为所选择的传感器中的部分数据,从变化趋势可以发现,电流和振动在停泵时几乎是瞬时变化的,温度则是缓慢变化的,这与不同物理量变化的相应快慢有关,实际机理相符。

### 3.2 实验结果与分析

由于真实核电数据中同一设备的发生异常的数据较少且标记,为了验证算法的检测效果并给出量化的指标,采取真实数据插入人造异常数据的方法对模型进行验证。

首先将数据集划分为训练数据和测试数据,之后使用未插入异常的训练数据训练 VAE 检测模型,并插入了异常数据的测试集上进行验证。

为了量化检测性能,实验使用以下指标对于检测结果进行评价:即常用的评价指标精确率  $P$  (precision)、召回率  $R$  (recall) 和  $F_1$  分数 ( $F_1$ -score) 对异常检测结果进行评价,具体计算公式为

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$F_1 = \frac{2PR}{P + R} \quad (7)$$

式中:TP、FP、TN、FN 分别为真阳性、假阳性、真阴性、假阴性的数量。精确率描述的是检测出的异常

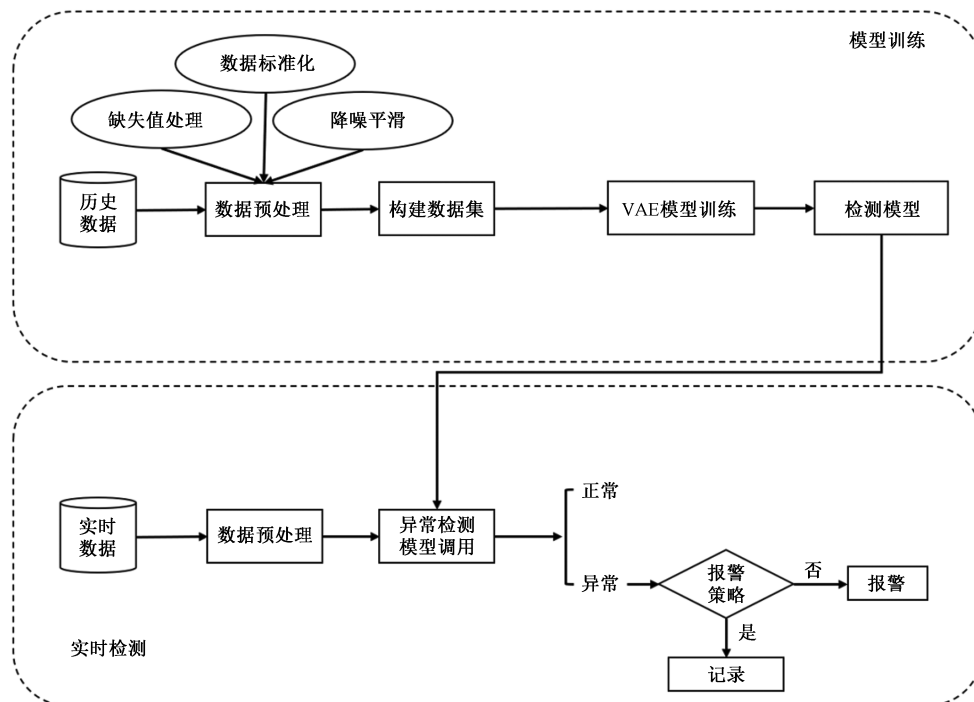


图 3 异常检测模型流程图

Fig. 3 Anomaly detection procedure with VAE

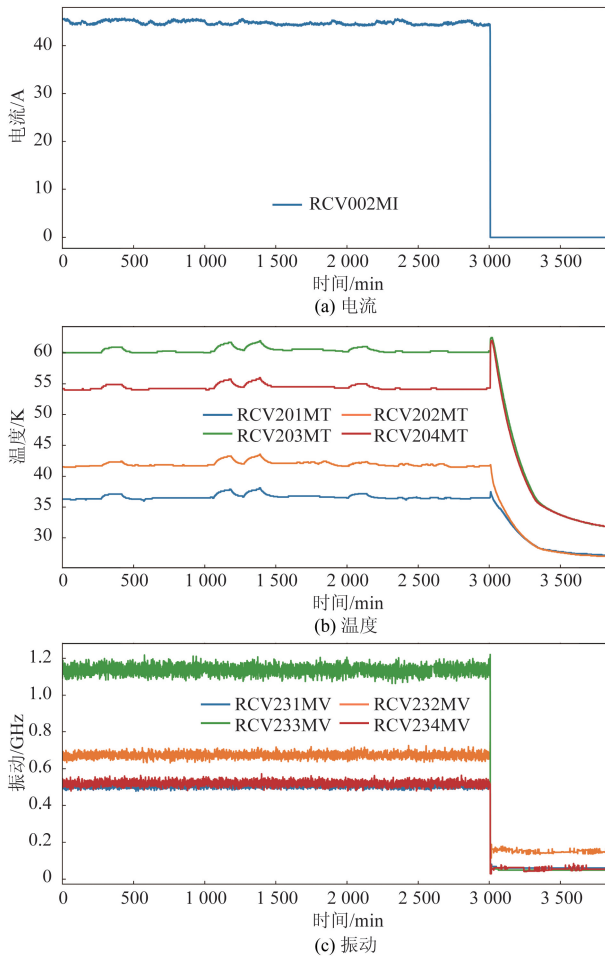


图 4 上充泵相关传感器实验数据

Fig. 4 Experimental data of the sensor related to the upper charge pump

点中真实异常点所占的比例,精确率越高说明误报率低;召回率描述的是所有异常点中成功检测到的异常点所占的比例,召回率高说明漏检率低; $F_1$  分数则是结合了精确率和召回率的综合评价指标。

图 5 所示为 VAE 对异常数据的检测结果,为便于显示,图 5(a)中仅列出了某一维插入了异常片段的温度数据。图 5(b)中红色虚线对应设定的判断阈值,实例分数高于红色虚线的部分即为检测为异常的数据。根据结果可以看出,VAE 检测模型能够成功检测出插入的稳态和变化态的异常数据片段。需要说明的是,实验中异常判断的阈值是根据设定的异常点占比百分数设置的,为了使检测结果符合实际需求,实际使用需要通过更多数据进行交叉验证来确定最佳的阈值,从而在误检和漏检中找到最符合状态监测需求的平衡。

为了进一步验证 VAE 的异常检测效果,在此引入常见的异常检测算法孤立森林和支持向量机,将这两种算法与 VAE 进行对比,计算得到量化指标如表 1 所示。

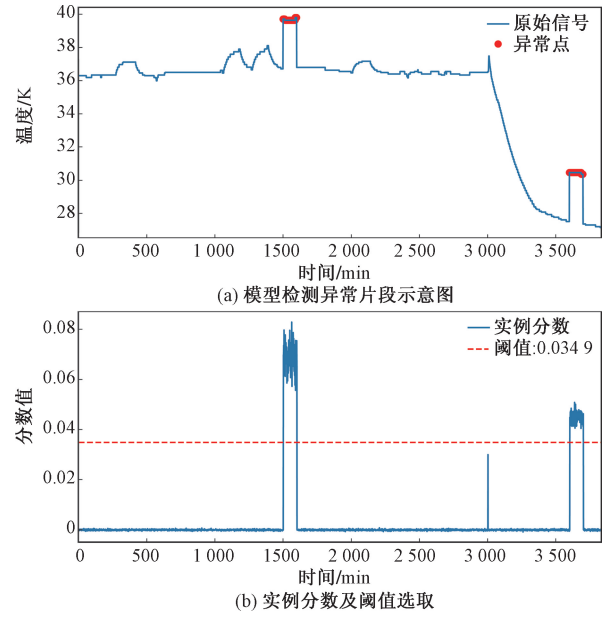


图 5 基于 VAE 异常检测模型的检测结果

Fig. 5 Results of anomaly detection by VAE-based model

表 1 VAE 与经典异常检测算法的结果对比

Table 1 The comparison of detection performance by VAE and classical anomaly detection models

算法名称	精确率	召回率	$F_1$
孤立森林	0.171 0	0.995 0	0.291 8
SVM	0.450 0	0.990 0	0.309 4
VAE	0.908 3	0.990 0	0.947 4

通过表 1 中展示指标结果对比可以看出,虽然三种方法的召回率都达到了 99%,但是孤立森林方法和 SVM 方法的检测精准率都相对较低,这意味着大量的误报,对于实际应用来说会大大降低可用性。经分析,使得孤立森林和 SVM 误检数量较多的是由于测试数据中包含了不同工况的数据,数据特征存在较大变化,导致不能很好地将不同工况的正常数据与异常数据分开。相比而言,VAE 方法的异常检测准确率和召回率都能达到 90% 以上,并没有受到不同工况数据的明显干扰,取得了较好的检测结果。

## 4 结论

由于核电厂智能运维的发展趋势和对于核电大数据的发掘需求,基于深度学习的核电运行状态监测方法正成为未来的发展方向。基于变分自编码器提出了一种可用于核电数据异常检测的算法模型,通过正常运行数据学习正常模式下的数据分布建立检测模型,并通过异常数据的重构误差大的原理实现对于异常状态的判断。该模型可以通过深度学习提取海量时序数据中的数据特征,并

且基于重构的方式可以在缺乏数据标记的情况下实现对于不同种类异常情况的有效检测。真实核电数据运行数据上的实验结果证明,该方法具有较好的检测性能,并可以实现跨工况的异常检测,具备一定的实用前景。

#### 参 考 文 献

- [1] Ma J, Jiang J. Applications of fault detection and diagnosis methods in nuclear power plants: a review[J]. *Progress of Nuclear Energy*, 2011, 53(3): 255-266.
- [2] Bailey D, Wright E, Practical SCADA for industry[M]. Oxford: Newnes, 2003.
- [3] Mcavoy T. Fault detection and diagnosis in industrial systems[J]. *Journal of Process Control*, 2002, 12(3): 453-454.
- [4] 赵春晖, 王福利. 工业过程运行状态智能监控: 数据驱动方法[M]. 北京: 化学工业出版社, 2019: 8-9.  
Zhao Chunhui, Wang Fuli. Intelligent monitoring of industrial process operation status: data-driven approach [M]. Beijing: Chemical Industry Press, 2019: 8-9.
- [5] 赵劲松, 王梓齐, 刘长良. 基于 Bagging 集成策略和多元状态估计的风电机组齿轮箱状态监测[J]. *科学技术与工程*, 2020, 20(20): 8180-8186.  
Zhao Jinsong, Wang Ziqi, Liu Changliang. Windturbine gearbox condition monitoring based on bagging ensemble strategy and MSET [J]. *Science Technology and Engineering*, 2020, 20(20): 8180-8186.
- [6] 邢尚鹏, 隋阳. 核电厂设备冷却水系统动态可靠性评估方法[J]. *科学技术与工程*, 2022, 22(23): 10092-10101.  
Xing Shangpeng, Sui Yang. An approach for dynamic reliability assessment of equipment cooling water system for nuclear power plant [J]. *Science Technology and Engineering*, 2022, 22(23): 10092-10101.
- [7] 梁海玲, 白森, 李坚. 基于鲁棒稀疏 PCA 的工业异常检测[J]. *科学技术与工程*, 2022, 22(15): 6164-6171.  
Liang Hailing, Bai Sen, Li Jian. Industrial anomaly detection based on robust sparse PCA [J]. *Science Technology and Engineering*, 2022, 22(15): 6164-6171.
- [8] Russell E L, Chiang L H, Braatz R D. Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis [J]. *Chemometrics and Intelligent Laboratory Systems*, 2000, 51(1): 81-93.
- [9] Zhao C H, Gao F R. Fault-relevant principal component analysis (FPCA) method for multivariate statistical modeling and process monitoring [J]. *Chemometrics & Intelligent Laboratory Systems*, 2014, 133(1): 1-16.
- [10] Burnham A J, Viveros R, MacGregor J F. Frameworks for latent variable multivariate regression [J]. *Journal of Chemometrics*, 1996, 10(1): 31-45.
- [11] Jong S D. SIMPLS: An alternative approach to partial least squares regression [J]. *Chemometrics and Intelligent Laboratory Systems*, 1993, 18(3): 251-263.
- [12] Mahadevan S, Shah S L. Fault detection and diagnosis in process data using one-class support vector machines [J]. *Journal of Process Control*, 2009, 19(10): 1627-1639.
- [13] Tran T N, Wehrens R, Buydens L. KNN-kernel density-based clustering for high-dimensional multivariate data [J]. *Computational Statistics & Data Analysis*, 2006, 51(2): 513-525.
- [14] Xie L, Liu X Q, Zhang J M, et al. Non-Gaussian process monitoring based on NGPP-SVDD [J]. *Acta Automatica Sinica*, 2009, 35(1): 107-112.
- [15] Li Y, Li X. Anomaly detection of nuclear power plant based on SVM optimized by genetic algorithm [J]. *Journal of Nuclear Science and Technology*, 2018, 55(10): 1075-1082.
- [16] Zhou Y, Liang X, Zhang W, et al. VAE-based deep SVDD for anomaly detection [J]. *Neurocomputing*, 2021, 453: 131-140.
- [17] Bergamin L, Carraro T, Polato M, et al. Novel applications for VAE-based anomaly detection systems [C]// 2022 International Joint Conference on Neural Networks (IJCNN). New York: IEEE, 2022: 1-8.