

DOI:10.19651/j.cnki.emt.2210387

基于边缘智能的风电机组设备监测系统^{*}

张西晨 徐璐

(浙江理工大学信息科学与工程学院 杭州 310018)

摘要: 在工业物联网中,风力涡轮机的数据采集与监视控制系统(SCADA)产生的海量数据因实时性要求,不适合直接送到云端处理。本文设计并搭建了一套基于边缘智能技术的微型风电机组设备状态监测系统。针对风电机组的转动、振动进行异常监测,对比并分析了 OC-SVM、孤立森林和 HBOS 三种无监督异常检测算法结果。实验结果表明,单类支持向量机算法在转动异常测试集上的 F1 分数为 0.997,在振动异常测试集上的 F1 分数为 0.969,具有较好的实时异常检测效果。本文能够为边缘侧训练与推理方案的落地性验证提供一定的参考价值。

关键词: 风力涡轮机;边缘智能;异常检测;单类支持向量机;孤立森林

中图分类号: TM933 **文献标识码:** A **国家标准学科分类代码:** TK83

Wind turbine equipment monitoring system based on edge intelligence

Zhang Xichen Xu Lu

(School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: In the Industrial Internet of Things, the massive data generated by the SCADA of wind turbines is not suitable for being directly sent to the cloud for processing due to real-time requirements. This paper designs and build a set of Micro-Wind Turbine equipment condition monitoring system based on Edge Intelligence. Three unsupervised anomaly detection algorithms, including OC-SVM, IForest and HBOS, are analyzed and compared with each other. The experimental results show that OC-SVM have the best real-time anomaly detection effect. The F1 scores in the rotation anomaly test set and vibration anomaly test set are 0.997 and 0.969, respectively. This paper can provide some reference value for the landing verification of edge side training and reasoning scheme.

Keywords: wind turbine; edge intelligence; anomaly detection; OC-SVM; iforest

0 引言

风能是当前最受欢迎的新能源之一,然而风电机组设备的骤增也会带来两个新的挑战,一是给云端数据处理决策带来实时性和安全性风险,二是对工控设备状态监测系统的稳定性与可靠性的异常检测算法提出了更高的要求^[1]。因风力涡轮机状态检测需要较高的实时性,集中式的云计算方式已经无法满足这种需求,而边缘智能(edge intelligence)被认为是一种有效的解决方案^[2-3]。本文以工业风扇模拟实际风力场,用微型风电机组模型和数据采集设备模拟可正常运行的风力涡轮机设备,同时结合多种无监督异常检测算法,完成了设备的健康状态监测与预测性维护工作,在一定程度上解决了云计算对大型工控设备检测实时性不足问题,同时也保证了风机数据的安全性。

因风机设备数量的增多以及状态检测传感器种类增多

产生了海量的数据,许多学者对这些数据的处理也做出了很多尝试。文献[4-5]根据终端设备的硬件条件,设计适用于自己终端环境的轻量级 LSTM 故障监测模型,并成功部署于终端设备。文献[6]通过边缘智能技术,在计算机视觉领域利用模型分割方法将模型分别部署在边缘端和云端进行协同推理,适当地解决了资源受限的问题。Du 等^[7]建立了自组织映射(self-organizing map)神经网络模型,将高维变量投射至二维空间,对风电机组性能健康状态数据进行建模;颜永龙等^[8]利用反向传播神经网络和最小二乘支持向量机建立了风电机组与风力轴承的联合工作状态检测模型。

本文在此基础上搭建了模拟运行设备,并在边缘侧对比分析 OC-SVM、孤立森林和 HBOS 三种无监督异常检测算法,发现单类支持向量机模型最适合本系统。同时借助阿里云物联网平台实现云端远程监控,搭建起一套云一

收稿日期:2022-06-20

^{*} 基金项目:浙江省重点研发计划项目(2020C03094)、浙江省自然科学基金青年基金(LQ20F050010)项目资助

管一边一端的全链路系统,在一定程度上验证了边缘智能技术赋能工控设备预测性维护的可行性。

1 系统整体框架设计

因安装在大型工控设备上的传感器数据维度小且变量之间存在较为明显的相关关系,故工控设备的边缘智能技术往往采用边缘训练与推理的服务方式。本系统设计的整体架构也是基于这种服务方式,主要包括终端节点、边缘服务器和云中心网络三层。终端节点层主要为接入边缘服务器

的传感器设备,包括风速计、振动传感器、环境传感器等状态检测传感器,这些终端设备需要采集工控设备状态数据并将获取的信息传送给边缘服务器。边缘服务器层的主要任务是对终端设备传输的数据进行融合、清洗、预处理和训练等,并应用异常检测算法判断设备是否存在异常,甚至是发生异常的具体部位,最后边缘服务器需将检测的结果上报给云数据中心网络。云数据中心网络的任务是对上传到云端的数据实现可视化,当设备出现异常时可通知维修人员采取相应措施来对设备进行维护,整体架构如图1所示。

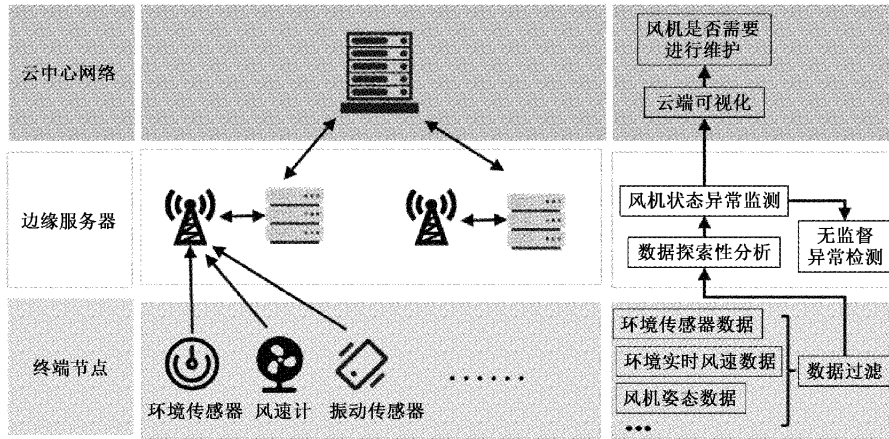


图1 边缘智能工控设备预测性维护架构图

本文搭建的模拟风电机组状态监测系统包括风电机组设备端、边缘设备端和远端三部分。风电机组设备端主要包括微型风机模型、状态监测传感器(姿态传感器、可采集发电电压和风机转速的发电机、风速计)与基于STM32的数据采集系统。系统通电后,数据采集设备会实时采集风机的姿态信息(XYZ轴加速度)、RS485风速计采集环境的实时风速、环境传感器采集环境湿度和大气压数据^[10]。边缘端主要包括边缘设备和触摸屏,边缘设备主要负责传感器数据的清洗、探索性分析、

模型训练以及选取最适合本系统的异常检测模型的任务,串口屏负责将原始数据和检测结果在本地实现可视化。云端部分包括数据上云和数据可视化两部分,利用阿里云物联网平台将边缘设备接收的数据上云,通过在该平台上创建产品、定义产品功能与绑定设备即可。在云端可视化部分,利用IOT Studio平台制作Web端的可视化界面,通过关联设备功能与配置数据源即可实现云端的可视化。模拟风电机组状态监测系统架构如图2所示。

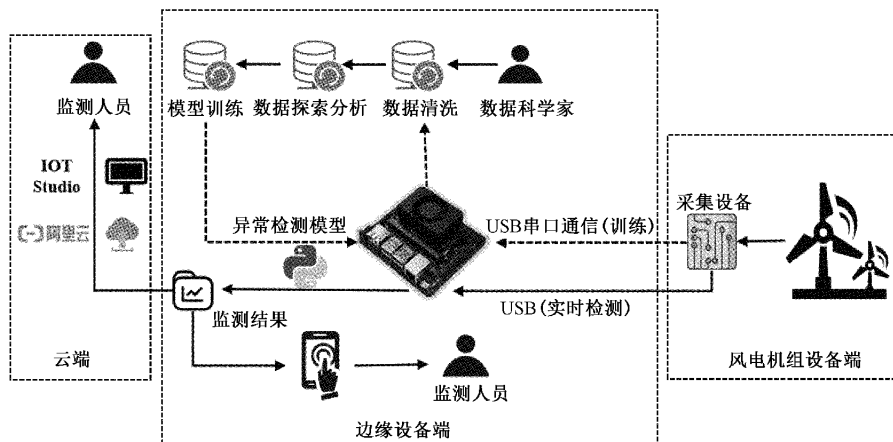


图2 模拟风电机组状态监测系统架构图

本文主要使用OC-SVM、孤立森林、HBOS 3种无监督异常检测算法,同时将本系统需检测的两种异常情况分

别定义为转动异常与振动异常:

- 1) 风机转动异常:这类异常包括风机转速过快与过慢

两种。若风机转速过快可理解为风机转轴与齿轮箱连接过松或有零件折断,若风机转速过慢,可理解为风机转动过程中有异物阻塞。

2) 风机振动异常:这类异常是根据风机的振动数据来进行检测。若风机出现此类异常,在环境风速过大的情况下风机就会产生明显的振动,故系统需要侦测出此类情况。

2 硬件系统与嵌入式软件设计

目前使用最广泛的双馈异步风电机组包括风轮、发电机、变速齿轮箱、偏航系统、液压系统、变桨系统、SCADA 等部分^[9-10],结构较为复杂。然而本文重在研究边缘智能赋能工控设备的可实现性,因此简化风机结构,仅保留发电机、风轮和 SCADA 部分。系统采集的传感器数据特征为 X 轴加速度、Y 轴加速度、Z 轴加速度、发电电压、风机转速、环境实时风速、环境温度、湿度和大气压 9 种数据。

2.1 硬件系统设计

本系统在实际测试场景中的实物如图 3 所示,其中包括提供风力的工业风扇、数据采集设备、用于云端可视化的显示屏三部分。系统通电后,工业风扇可通过调节档速来模拟不同风速条件,风机的状态数据和检测结果可通过触摸屏和 Web 端来实现可视化。数据采集设备与边缘设备部分实物图如图 4 所示,姿态传感器 MPU6050 和 385 发电机安装在模拟风机上,环境传感器 BME280 和主控 STM32 单片机安装在 PCB 板上,三杯口 RS485 风速计来采集环境的实时风速。

边缘设备选择英伟达公司推出的 Jetson Xavier NX,它在功耗和性能方面有着卓越的性能,非常适合嵌入式和边缘系统中的高性能计算。

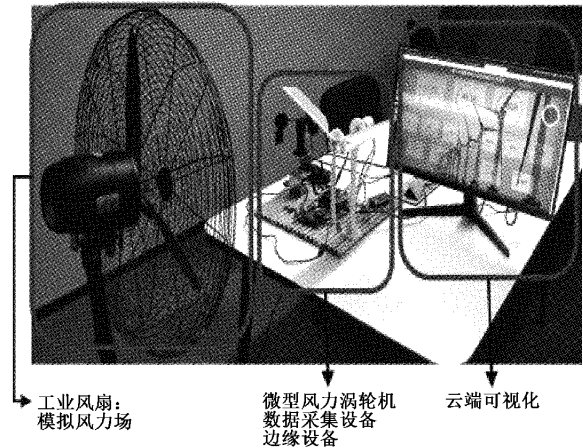


图 3 系统整体实物演示图

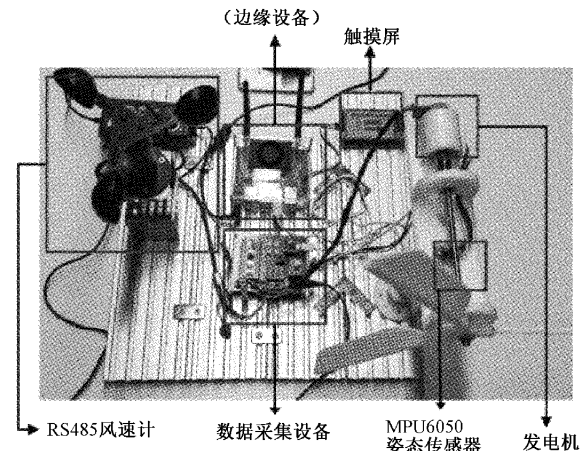


图 4 数据采集设备实物图

本文的硬件系统所用模块与功能如表 1 所示。

表 1 系统所用硬件模块结构

硬件模块	功能	通信协议
3D 打印件	模拟风电机组运行	无
96 线 385 电机	获取风机转速、电压	无
振动传感器(MPU6050)	获取风机姿态数据	IIC
电源适配器	数据采集部分的电源	无
MP1484EN 模块	电源降压模块	无
STM32F103C8T6	数据采集主控单片机	无
RS485 风速计	获取实时风速数据	RS485
RS485 隔离收发模块	辅助风速计与 STM32 通信	无
环境传感器(BME280)	获取温度、湿度、大气压	IIC
CP2102 模块	STM32 与边缘设备串口通信	RS232
Jetson Xavier NX	高性能的边缘设备	无
串口屏	用于数据本地可视化	RS232

2.2 STM32 嵌入式软件设计

本系统嵌入式开发环境选用 STM32CubeMX 和 Keil。

为确保系统能协调多个传感器,同时对实时数据进行卡尔曼滤波(消耗时间长)和串口通信,因此在系统设计时对软

件框架做了精心规划。在 STM32CubeMX 中利用第三方插件 FreeRTOS 嵌入式实时操作系统来实现模拟多线程效果,同时开启传感器读取(以下简称任务 1)和数据滤波与传输(以下简称任务 2)两个任务。为保证传感器连续工作,需将两个任务设置为不同的优先级,任务 1 因负责所有传感器读数故赋予其 OS Priority Real-time,而任务 2 负责实时数据的卡尔曼滤波和串口数据传输故赋予其 OS Priority Normal。

数据采集系统首先需要将所有传感器模块初始化,并打开定时器 1 输入捕获中断、定时器 4 中断和串口空闲中断。至此,进入 FreeRTOS 的两个任务并行阶段。在任务 1 中,调用姿态传感器读取函数来获取三轴加速度数据,之后调用 BME280 驱动函数来获取温湿度和气压等

环境数据;由 STM32 主动发送数据查询帧给 RS485 风速计,根据设备提供的数据应答帧中相关寄存器值来获取风速数据;带有 96 线码盘的 385 电机通过定时器 1 输入捕获和 ADC 电压采样获取风机转速与发电电压,等待所有数据就位之后开始启动任务 2,而任务 1 可以继续循环运行。在任务 2 中,对风机的 9 个数据进行卡尔曼滤波处理(此处卡尔曼滤波的测量噪声方差 Q 为 0.003,观测噪声方差 R 为 0.5),将处理好的数据通过单片机的 USART2 口和信号转换模块发送至边缘设备,传感器数据每 100 ms 上报一次,由此完成整个风机状态数据的采集工作,嵌入式软件设计流程如图 5 所示。图中边缘设备 Jetson Xavier NX 可将原始数据与异常检测结果发送给云端以实现远程监测。

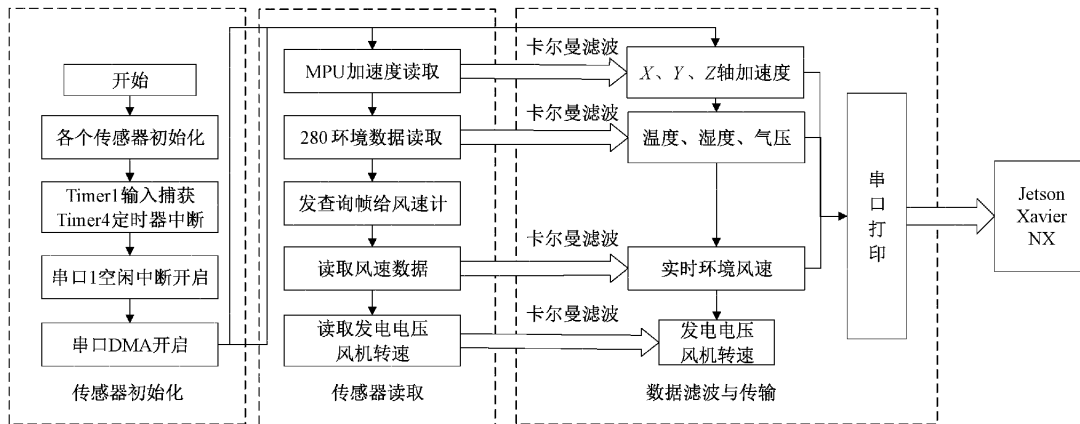


图 5 嵌入式软件设计流程图

3 异常检测模型训练与验证

边缘设备接收风力涡轮机的状态数据后需要对这些数据进行特征工程,其中包括数据的清洗、预处理与探索性分析。特征工程完成之后需要将这些数据送入到异常检测模型来对比不同算法的检测效果。本系统尝试了包括 OC-SVM、孤立森林和 HBOS 算法在内的 3 种无监督异常检测算法,最终发现不论是风机转动异常检测还是振动异常检测,OC-SVM 都优于另外两种算法,表现出良好的检测效能。

3.1 数据探索性分析与预处理

本系统要探索无监督异常检测效果,因此需要大量正常数据集作为训练样本以及少量的异常数据集作为测试样本。系统共选择包括 XYZ 三轴加速度、发电电压、实时风速和风机转速 6 个指标表征系统转动异常与振动异常。其中转动异常评价指标为风机转速(motor_speed)、风机电压(motor_Voltage)和实时风速(wind_speed);振动异常评价指标为 X 轴加速度(Ax)、Y 轴加速度(Ay)与 Z 轴加速度(Az)。在正常数据集的 6 个变量相关关系如图 6 所示,从图中可以发现风机转速、风机电压和实时风速的相关系数均在 0.95 以上,它们具有正相关关系,其中因风机转速

与发电电压数据都来自 385 发电机,故它们的相关系数为 1。从图 6 中也可发现表征风机振动状态的 X 轴加速度、Y 轴加速度与 Z 轴加速度与转动特征数据的相关系数绝对值均在 0.1 以下,故在本系统中可认为振动部分与转动部分无特别明显的相关性,因此可将系统的振动异常和转动异常可分开检测。为使检测更加方便以及有更强的解释结果,本文分别对转动异常检测和振动异常检测地特征数据进行 PCA 降维,由三维数据降到二维数据。

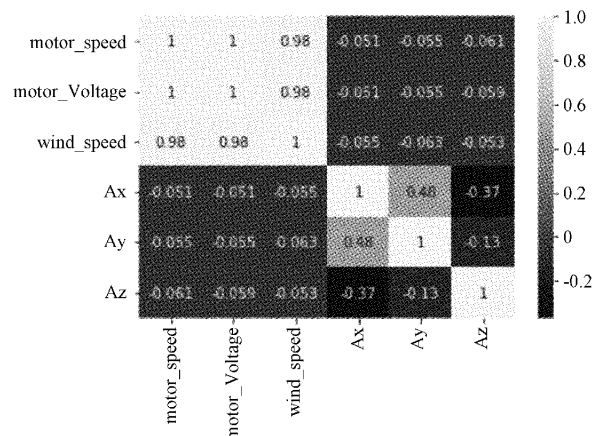


图 6 正常数据集特征变量相关性分析

本文选择特征维度的数据幅度不一致,在机器学习和深度学习中会极大地影响训练的收敛性和准确性,因此在这些特征数据训练之前,需要对其进行归一化处理^[11-12]。此处需要对选取的 6 个数据特征进行处理,使实际观测变量线性映射到[0,1]之间,如式(1)所示。

$$x_i = \frac{x_{ireal} - x_{min}}{x_{max} - x_{min}} \quad (1)$$

式中: x_{ireal} 为某次样本真实值, x_{min} 为该变量最小值, x_{max} 为该变量最大值。

本设计的转动异常检测基于数据集 1,其包括 12 000 组正常转动数据和 1 000 组异常转动数据;振动异常检测基于数据集 2,其包含 12 000 组正常振动数据和 1 000 组振动异常数据,本文对数据集划分如表 2 所示。

表 2 数据集划分

	转动异常	振动异常
	数据集 1	数据集 2
训练集	10 000 组正常数据	10 000 组正常数据
测试集	2 000 组正常数据 1 000 组异常数据	2 000 组正常数据 1 000 组异常数据

3.2 单类支持向量机

用于无监督异常检测的 OC-SVM 扩展了常用于分类的支持向量机的思想。传统的支持向量机的主要思想是把数据隐式地映射到高维空间,再利用线性分类器对其进行分类。而 OC-SVM 是以最大的间距来构造一个独立的超平面,若新数据位于边界之内,则为正常样本,反之,当其位于边界之外时则被视为异常。经典向量机的核心函数通常具有捕获数据非线性的能力,能将数据点从 R^d 中的输入空间映射到 R^D (D 通常是无限维的)中的高维空间。在高维空间中,数据通常是线性可分的,一般都有 3 种基本的核函数^[13],即线性核函数 $K(x, \bar{x}) = x \cdot \bar{x}$, 多项式核函数 $K(x, \bar{x}) = (x \cdot \bar{x} + 1)^q$, 径向基核函数 $K(x, \bar{x}) = \exp(-\gamma \|x - \bar{x}\|^2)$ 。根据以上描述,从运筹学的角度出发,单类支持向量机可用拉格朗日方法表示成二次规划最小化问题,并且是计算密集型的,不具有线性伸缩性^[14-15],如式(2)所示。

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + \frac{1}{vn} \sum_{i=0}^n \xi_i - \rho \\ s.t. (w \cdot K(x_i)) \geq \rho - \xi_i, \xi_i \geq 0 \end{cases} \quad (2)$$

其中, w 为权重向量,其维度与特征空间 $K(x_i)$ 一致, $K(x_i)$ 代表核函数, ξ_i 是非负的松弛因子, n 是数据集样本数量。 $v \in (0, 1)$ 是控制距离原点最大距离于支持向量在样本中的个数平衡的惩罚项,决定着决策边界的平滑程度。 w 和 ρ 是决定决策边界的关键参数,更是实现问题优化求解的关键变量。通过 Hinge 损失函数代替 ξ_i , 可得到无约束的目标函数,如式(3)所示。

$$\min_{w, \rho} \frac{1}{2} \|w\|^2 + \frac{1}{vn} \sum_{i=1}^n \max(0, \rho - w^T K(x_i)) - \rho \quad (3)$$

本文对算法检测结果定义 4 个检测内容,即决策边界、训练集错误率、测试集正常部分错误率和测试集异常部分错误率。决策边界是算法检测正常点与异常点划分线;训练集错误表示将正常点判为异常点的情况;测试集正常部分错误表示将正常点识别为异常点;测试集异常部分错误表示将异常点识别为正常点。OC-SVM 在数据集 1 中转动异常检测检测结果如图 7 所示。

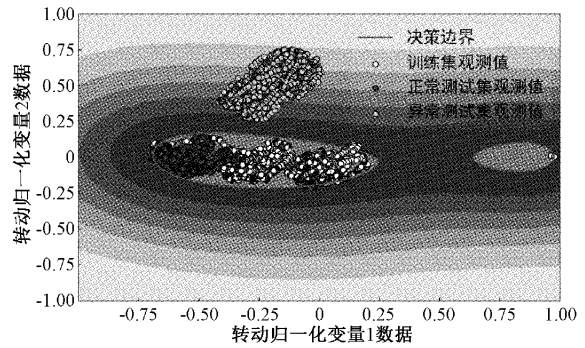


图 7 OC-SVM 转动检测结果

OC-SVM 在数据集 2 中的振动异常检测结果如图 8 所示。

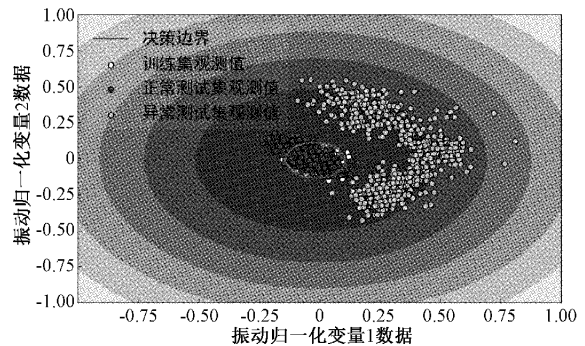


图 8 OC-SVM 振动异常检测结果

OC-SVM 在训练集与测试集中的错误率如表 3 所示。

表 3 OC-SVM 异常检测错误率

	正常训练集	正常测试集	异常测试集
转动异常	0.65%	0.45%	0
振动异常	0.91%	6.05%	0

3.3 孤立森林

孤立森林方法是属于集成学习,它对于异常数据的定义为:“分布稀疏的且远离密度高的群体的点”^[20]。该算法核心思想是寻找这些容易被孤立出来的点,越异常的点越容易被切割出来,孤立森林是以递归方式划分数据集各个树结构的集合。对于该过程中的每次迭代,将选择一个随

机特征,然后根据在所选特征的最大值和最小值之间随机选择一个值对数据进行划分,重复此过程,直到对整个数据集进行划分^[16-17]。一般情况下,异常点与根位置之间的路径比正常点短得多。

孤立森林算法对数据集1进行转动异常检测,检测结果如图9所示。

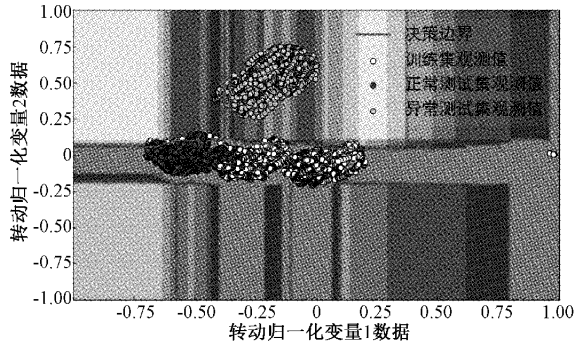


图9 孤立森林转动检测结果

孤立森林算法在数据集2振动异常检测的结果如图10所示。

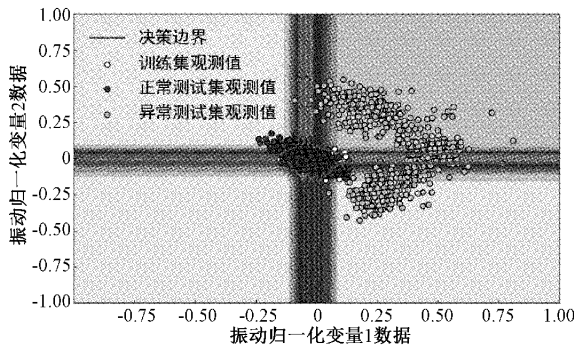


图10 孤立森林振动异常检测结果

孤立森林在训练集和测试集上的错误率如表4所示。

表4 孤立森林异常检测错误率 %

	正常训练集	正常测试集	异常测试集
转动异常	1	4.45	0
振动异常	1	9	4.9

3.4 HBOS 算法

基于直方图的异常检测(HBOS)是一种非参数技术,它的计算速度很快,对高维数据有着良好的适应性。该方法的基本假设是数据集中的每个特征都是独立的,对每个维度进行区间(bin)划分,区间的密度越高,异常评分越低。因为此算法对超参数没有依赖,避免了超参数的选择而导致的潜在误差,具有简单高效的特性。对于具有N维特征的多变量对象,计算每个特征的单个直方图密度,然后求和得出异常分数,算法模型如式(4)所示。

$$HBOS(p) = \sum_{i=0}^d \log\left(\frac{1}{hist_i(p)}\right) \quad (4)$$

其中,HBOS(p)表示异常得分,d表示特征维度,p表示某个样本向量,hist_i(p)是样本向量p的第i个特征的概率密度估计。HBOS算法对数据集1进行转动检测效果如图11所示。

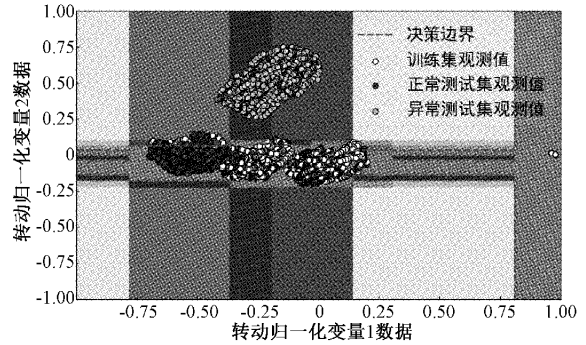


图11 HBOS转动检测结果

HBOS算法对数据集2进行振动检测效果如图12所示。

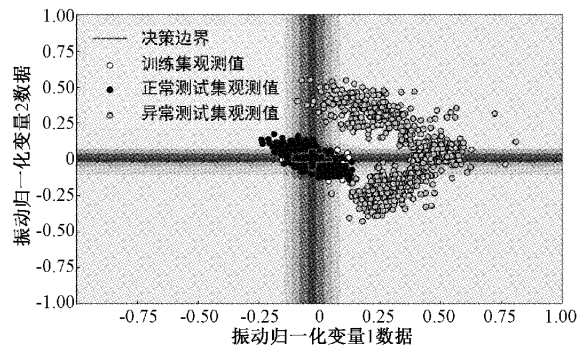


图12 HBOS振动检测结果

HBOS在训练集和测试集上的错误率如表5所示。

表5 HBOS算法异常检测错误率 %

	正常训练集	正常测试集	异常测试集
转动异常	0.76	3.85	0
振动异常	0.98	6.9	7.6

3.5 算法对比分析

准确率、精确率、召回率和F1分数是用来评价机器学习和深度学习模型的重要指标。在无监督的异常检测方法中,因样本的极度不均衡,准确率已无法正确评估算法模型的好坏,在实际中往往使用精准率、召回率和F1分数来对模型进行评价。单类支持向量机(OC-SVM)、孤立森林(IForest)和HBOS三种无监督异常检测模型的转动异常的检测结果如表6所示。

振动异常检测效果如表7所示。

表 6 三种算法转动异常效果对比

	准确率	精准率	召回率	F1 分数
OC-SVM	0.997	1	0.995	0.997
IForest	0.970	1	0.956	0.978
HBOS	0.974	1	0.962	0.981

表 7 三种算法振动异常效果对比

	准确率	精准率	召回率	F1 分数
OC-SVM	0.959	1	0.940	0.969
IForest	0.923	0.974	0.910	0.940
HBOS	0.929	0.961	0.931	0.946

4 结 论

本文搭建了一套室内微型风力涡轮机状态监测系统,验证了基于边缘训练与推理的边缘智能服务方式用于风电机组设备日常监测的可落地性。

系统通过对 XYZ 三轴加速度、发电电压、实时风速和风机转速 6 个指标的检测实现了风电机组系统转动异常与振动异常的日常监测。

在边缘设备上部署并且对比分析了 OC-SVM、孤立森林和 HBOS 三种无监督异常检测算法结果,实验结果表明,边缘设备上部署的 OC-SVM 算法在测试集具有较好的检测性能。

本文的研究可为基于物联网的工控设备的边缘异常检测提供一定参考。当然,对于更复杂的系统结构,可进一步尝试其它效果更佳的算法研究。

参考文献

- [1] 吴金利, 马宏忠, 吴书煜, 等. 基于振动信号的高压并联电抗器故障诊断方法与监测系统研制[J]. 电测与仪表, 2020, 57(1): 113-120.
- [2] 甄鹏华, 韩玉冰. 基于边缘智能的心电信号处理方法及监测系统[J]. 信息技术与网络安全, 2021, 40(7): 108-115.
- [3] 申普. 基于边缘智能的机电设备故障检测服务研究[D]. 桂林: 桂林电子科技大学, 2021.
- [4] PARK D, KIM S, AN Y, et al. LiReD: A light-weight real-time fault detection system for edge computing using LSTM recurrent neural networks[J]. Sensors, 2018, 18(7): 2110.
- [5] SOUALHIA M, FU C, KHOMH F. Infrastructure

fault detection and prediction in edge cloud environments[C]. Proceedings of the 4th ACM/IEEE Symposium on Edge Computing, 2019: 222-235.

- [6] LI E, ZHOU Z, CHEN X. Edge intelligence: On-demand deep learning model co-inference with device-edge synergy[C]. Proceedings of the 2018 Workshop on Mobile Edge Communications, 2018: 31-36.
- [7] DU M, MA S, HE Q. A SCADA data based anomaly detection method for wind turbines[C]. 2016 China International Conference on Electricity Distribution (CICED), IEEE, 2016: 1-6.
- [8] 颜永龙, 李剑, 李辉, 等. 采用信息熵和组合模型的风电机组异常检测方法[J]. 电网技术, 2015, 39(3): 737-743.
- [9] 陈伟. 基于大数据分析的风电机组运行状态监测方法研究[J]. 湖北农机化, 2019(16): 134-135.
- [10] 石睿. 基于大数据分析的风电机组运行状态监测方法研究[D]. 长春: 长春工业大学, 2018.
- [11] 李熠, 李永忠. 基于自编码器及极限学习机的工业控制网络入侵检测算法[J]. 南京理工大学学报, 2019, 43(4): 408-413.
- [12] 金晓航, 许壮伟, 孙毅, 等. 基于生成对抗网络的风电机组在线状态监测[J]. 仪器仪表学报, 2020, 41(4): 68-76.
- [13] 武玉坤, 李伟, 倪敏雅, 等. 单类支持向量机融合深度编码器的异常检测模型[J]. 计算机科学, 2022, 49(3): 144-151.
- [14] 李国全, 高建宇, 白天宇, 等. 基于 SVM 与改进型乌鸦搜索算法的风电功率预测方法[J]. 国外电子测量技术, 2022, 41(2): 40-45.
- [15] 吴君凤. 基于孪生支持向量机的民用航空发动机故障诊断[J]. 电子测量技术, 2020, 43(22): 132-136.
- [16] 刘李梦玮. 基于孤立森林的网络数据流在线异常检测方法研究[D]. 北京: 北京邮电大学, 2021.
- [17] 马俊彦, 杨仲江, 王昊. 基于孤立森林算法的电涌保护器异常参数监测研究[J]. 电子测量与仪器学报, 2020, 34(7): 58-63.

作者简介

张西晨, 本科, 主要研究方向为边缘智能、异常检测。

E-mail: 2802433362@qq.com

徐璐, 博士, 讲师, 硕士研究生导师, 主要研究方向为成像激光雷达, 软件无线电。

E-mail: xlhit@126.com