

文章编号:1001-5078(2024)08-1277-09

· 红外技术及应用 ·

# 红外特征提取与关联分析的设备故障预警方法

杨磊<sup>1</sup>, 王国丽<sup>1</sup>, 朱丽晓<sup>1</sup>, 李云红<sup>2</sup>, 李丽敏<sup>2</sup>, 苏雪平<sup>2</sup>, 王梅<sup>2</sup>

(1. 国家能源集团新疆吉林台水电开发有限公司, 新疆 伊犁 835000; 2. 西安工程大学电子信息学院, 陕西 西安 710048)

**摘要:**水电站电气设备的安全稳定运行至关重要。为了实现设备故障的自动预警,通过红外特征提取并进行灰色关联分析,提出建立主成分分析(PCA)和基于密度的聚类算法(DBSCAN)的设备故障预警模型。首先,通过数据预处理补全缺失数据,剔除异常数据后进行主成分分析降维并提取新的主成分特征。其次,将新的主成分采用DBSCAN算法构建特征样本集,建立灰色关联模型,计算灰色关联系数,然后,通过灰色关联系数的变化程度突变点进行故障预警。实验结果表明,所提方法能有效提取红外特征,并在设备异常状态下实现设备故障预警,故障预警准确率达到97.88%。

**关键词:**红外特征提取;PCA降维;DBSCAN聚类;灰色关联分析;故障预警

**中图分类号:**TP391;TN219 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2024.08.015

## Infrared feature extraction and correlation analysis for early warning of equipment faults

YANG Lei<sup>1</sup>, WANG Guo-li<sup>1</sup>, ZHU Li-xiao<sup>1</sup>, LI Yun-hong<sup>2</sup>, LI Li-min<sup>2</sup>, SU Xue-ping<sup>2</sup>, WANG Mei<sup>2</sup>

(1. CHN ENERGY Xinjiang Jilintai Hydropower Development Co., Ltd, Yili 835000, China;

2. School of Electronics and Information, Xi'an Polytechnic University, Xi'an 710048, China)

**Abstract:** The safe and stable operation of electrical equipment in hydropower plants is critical. In order to achieve the automatic early warning of equipment faults, through infrared features extraction and gray-associated analysis, it is proposed to establish an early warning model of equipment failure with principal component analysis (PCA) and density-based clustering algorithm (DBSCAN). Firstly, the missing data are made up through the data pre-processing, the abnormal data are eliminated, and the principal component analysis is performed to reduce the dimensionality and extract the new principal component features. Secondly, the new principal components are used to construct the feature sample set by DBSCAN algorithm, establish a gray association model, calculate the gray associated coefficient, and then fail to warn the degree of change point of the gray association coefficient carry out the fault early warning through the change degree of the grey correlation coefficient of the mutation points. The experimental results show that the proposed method can effectively extract infrared characteristics and achieve equipment fault warning under the abnormal state of the equipment, and the fault warning accuracy rate reaches 97.88%.

**Keywords:** infrared feature extraction; PCA dimensionality reduction; DBSCAN clustering; grey correlation analysis; fault warning

**基金项目:**国家自然科学基金项目(No. 62203344);陕西省自然科学基金基础研究重点项目(No. 2022JZ-35);

国家级大学生创新创业计划项目(No. 202210709012);陕西高校青年科技创新团队项目资助。

**作者简介:**杨磊(1985-),男,工程师,主要从事控制及检测相关工作。E-mail:12077181@ceic.com

**通讯作者:**李云红(1974-),女,满族,教授,博士,主要从事红外热像测温技术、人工智能、深度学习、图像处理研究。

E-mail:hitliyunhong@163.com

**收稿日期:**2023-11-03; **修订日期:**2023-12-28

## 1 引言

电气设备故障预警是确保水电站、变电站等电力系统关键设备稳定运行的关键环节。在实践中,及时发现并诊断设备故障能够避免系统停机、事故发生,从而降低经济损失和安全隐患。传统的电气设备故障预测通常依靠人工来分析数据,工作量大且耗时<sup>[1]</sup>,因此红外热像检测技术被广泛应用于变电站中,有效改善了结果的可靠性并提高了安全系数。此技术具备多项优点,如非接触<sup>[2]</sup>、高测温效率以及适用于广泛的检测范围。通过及时判断设备的故障,并迅速定位潜在的危险区域<sup>[3]</sup>。灰色关联分析在数据信息缺乏等方面具有优势,且计算量小,在高压断路器<sup>[4]</sup>等设备故障中均有应用。采集分析电气设备的运行数据,在设备故障发生之前及时发现异常,并提出预警信息,帮助企业或个人采取相应的维护措施,尽可能避免设备故障对生产和安全造成的影响。

A. Dragomir<sup>[5]</sup>等人提出了电负荷修正手段,应用红外热像技术测定电力设备热应力时考虑电气负载影响,其需在最佳外部环境条件下多次测试来调整校正参数。李宝树<sup>[6]</sup>等人提出了一种通过分析高压电力设备的红外图像进行故障分析的系统。该系统利用红外成像技术,以非接触式实时监测方式来确定高压设备的运行状态和故障位置,并能够预测未来的状态。杨曹<sup>[7]</sup>等人从国内电力系统日常巡检角度,介绍了热红外成像仪在常见电力设备热辐射故障诊断、红外图像获取和分析过程方面的应用。提出了一些红外诊断与日常检测结合的运行方法。邹辉<sup>[8]</sup>等人提出了一种新型的电力设备智能热缺陷诊断方法。该方法旨在解决人工分析热图像所需的大量时间和精力,以及可能导致的判断误差和错误诊断问题。它采用了智能化的方式来进行热缺陷诊断,能够更快速、更准确地识别热缺陷。李鑫<sup>[9]</sup>等人采用粒子群算法 PSO 和 Niblack 算法从电力设备红外图像中分割出设备热像并提取故障特征参量,通过 SVM 和优化 BA 进行参数配置实现设备故障诊断。聊方伦<sup>[10]</sup>等人采用 PCA 算法对火电厂电气设备故障进行故障诊断,此方法虽提高了诊断精度,可能会增加训练时间和计算复杂度。朱明皓<sup>[11]</sup>采用改进主元分析对燃煤电厂除尘设备的早期故障预警。刘小强<sup>[12]</sup>等人采用层次聚类建立了

具有自学习能力和自适应能力的变压器故障预警模型,但层次聚类对于数据中的噪声和异常值比较敏感,会影响最终聚类结果。李艳萍<sup>[13]</sup>等人基于长短时记忆和堆栈去噪自编码器网络,提出了一种自适应状态监测和故障预警框架,通过马氏距离和切比雪夫不等式来反映设备状态并进行早期异常检测。白杰<sup>[14]</sup>等人利用主成分分析和灰色关联分析结合,通过权重确定和数据处理,以诊断电力变压器故障类型,但是依赖于大量的样本数据以获得准确的权重和关联系数,因此在数据不足的情况下可能表现不佳。此外,算法的实际应用可能会受到噪声和干扰的影响,需要进一步的鲁棒性改进。

针对以上算法存在的问题,论文通过分析电气设备故障原因,论文针对电气设备故障,提出了一种结合 PCA-DBSCAN 特征提取与灰色关联分析相结合的故障预警方法。

## 2 电气设备故障及预处理

### 2.1 电气设备故障原因分析

电气设备在安全隐患初期,可能会因为大风、高温高湿等外部环境因素导致设备散热不良,引发设备过热、烧毁等问题。设备正常运转时由于电流、电压、功率、水位流量等作用而导致发热<sup>[15]</sup>。当输电线路短路或电气设备负载过大时,电流可能会超过设备的承载能力,导致设备过热甚至损坏。水电站电压水平较高,电源电压受到干扰或波动时,电气设备的工作电压可能会不稳定,导致设备无法正常启动或运行。功率超过设备的负荷能力可能会导致设备过载发热,引发故障或损坏。此外,水位流量信号在控制系统中发挥着重要作用,如调节水轮机转速、控制阀门开度等。如果水位流量信号不稳定或异常,可能会导致自动化控制系统失灵,从而影响设备的正常运行。这些故障往往发生在设备内部,在前期很难发现,因此要定期检查和维护电气设备,确保设备电压稳定、电缆线路良好、水力波动和负荷波动在正常范围内,同时注意环境因素对设备的影响,采取相应的措施进行防护和保养。

### 2.2 结合 PCA-DBSCAN 的特征提取

PCA-DBSCAN 方法原理:将计算机获取的时间序列数据,通过主成分分析<sup>[16-17]</sup> (Principal Component Analysis, PCA) 进行数据预处理,提取得到新的主成分解决数据冗余问题,根据累积方差贡献率对

主成分进行筛选,降低数据特征变量的维度。再利用筛选后的主成分使用基于密度的聚类算法<sup>[18]</sup> (Density-Based Spatial Clustering of Applications with Noise, DBSCAN) 进行聚类建模,得到电气设备故障预警模型前期的数据。

设备实际测量的电流、电压、功率、水位流量等数据中都会存在一定的相关性,导致了信息重叠,增加了数据处理的难度。其中对数据进行降维可以进行人工特征筛选,但往往费时又费力,效果还有可能不好,因此我们可以采用一些模型来进行数据降维。常用的就是主成分分析法,它通过线性变换将原始特征转换为新的特征向量,使得新的特征向量按照其对样本差异的解释能力从大到小排序。通过对原始数据压缩和降维,减少数据分析所需的时间,并将复杂的问题简化,此外,降维后的数据尽可能保留了原始数据的信息并最大程度地反映了真实情况,PCA 通过对数据进行标准化和投影变换,将高维数据映射到低维空间中,以提取出数据的主要特征。

### 2.2.1 PCA 理论

主成分分析法的几何意义是通过线性变换将原始数据集投影到新的坐标空间。假设在二维空间中数据由变量  $x_1$  和  $x_2$  构成,通过主成分分析对数据进行正交变换后得到新的坐标系,由变量  $y_1$  和  $y_2$  表示,在这个变换中,  $y_1$  是第一个主成分,它对应于原始数据集中的最大方差方向。也就是说,  $y_1$  是  $x_1$  和  $x_2$  的线性组合,且在  $y_1$  方向上的方差最大。这意味着  $y_1$  能够最大限度地保留原始数据集中的方差信息。通过这样的线性变换,我们将原始的二维数据集投影到一个新的二维坐标空间中,其中每个维度都代表一个主成分。多维原始数据降维方式与二维类似,原始坐标与新坐标之间关系如式:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (1)$$

PCA 算法基本流程如下:

(1) 假设计算机监控系统采集了指标数据为  $m$  维向量  $x = (x_1, x_2, \dots, x_m)^T$ ,  $n$  个特征变量  $x = (x_{i1}, x_{i2}, \dots, x_{im})^T, i = 1, 2, \dots, n$ , 对原始指标数据进行标准化变换:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, i = 1, 2, \dots, n; j = 1, 2, \dots, m \quad (2)$$

$$\bar{x}_i = \frac{\sum_{j=1}^n x_{ij}}{n} \quad (3)$$

$$s_j^2 = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n - 1} \quad (4)$$

(2) 计算协方差矩阵:

$$C = \frac{1}{n} Z^T Z \quad (5)$$

协方差矩阵  $C$  是  $m \times m$  实对称矩阵。

(3) 对协方差矩阵  $C$  进行特征分解,求特征值  $\lambda_i$  以及  $\lambda_i$  对应的特征向量  $W_i$ 。

(4) 主成分个数的选择:

由贡献率大小选取前的  $k$  个特征值以及对应的

特征向量。本文使用累积贡献率满足  $\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^n \lambda_i} \geq 85\%$  以上,取前  $k$  个主成分形成新的主元空间进行分析。 $k$  为主成分个数,  $n$  为所有主成分个数。

(5) 计算主成分得分:

$$F_i = Z_1 u_{1m} + Z_2 u_{2m} + \dots + Z_m u_{im}, i = 1, 2, \dots, m \quad (6)$$

其中,  $F_i$  为第  $i$  主成分得分。

主成分得分是数据在新坐标系中的表示,它反映了数据在不同主成分方向上的变化情况。主成分分析通过旋转坐标系,获得一个新的坐标系。新坐标系可以以矩阵形式表示,矩阵的特征向量可以将新坐标系表示出来,协方差矩阵的特征向量称为主成分,原始数据投影到新坐标系上对应的值称为主成分得分。

### 2.2.2 DBSCAN 聚类算法

DBSCAN 算法是常用的聚类算法之一,在降维后的主成分空间中,DBSCAN 会根据每个数据点的邻近性和密度来判断其是否为核心点、边界点或噪声点。基于这些核心点和邻近关系,DBSCAN 会将降维后的数据点分配到不同的聚类中,形成数据的聚类结构。基本原理如下:

DBSCAN 聚类算法包含两个邻域参数,  $\varepsilon$  半径参数和  $\text{minPts}$  密度阈值,定义了一个样本点周围的邻域的大小。对于给定的一个样本点,如果在  $\varepsilon$  的范围内存在至少  $\text{minPts}$  个样本点,则认为这些点属于同一个簇。 $\varepsilon$  参数控制了簇内的样本点的紧密程

度。较小的  $\varepsilon$  值将导致更多的簇被发现,而较大的  $\varepsilon$  值可能将一些本应属于同一簇的样本点划分到不同的簇中。 $\text{minPts}$  指定了作为核心点所需的最小样本数。核心点是位于  $\varepsilon$  领域内具有至少  $\text{minPts}$  个样本点的数据点。通过设置适当的  $\text{minPts}$  值,可以控制簇内的丰富程度,过小的值可能导致过多的噪声点,而过大的值则可能使得簇无法发现。定义如下:

(1)  $\varepsilon$  邻域:  $N_\varepsilon(a)$  为数据点  $a$  的  $\varepsilon$  邻域;邻域内点的集合可表示  $N_\varepsilon(a) = \{b \in D \mid \text{dist}(a,b) \leq \varepsilon\}$ ,  $D$  为所有点的集合,  $\text{dist}(a,b)$  是  $a,b$  两点之间的距离。

(2) 核心对象:对给定  $a$ ,  $a$  的邻域内包含数据点个数大于等于  $\text{minPts}$ , 称点  $a$  为核心对象;

(3) 边界点:对给定  $a$ , 若点  $a$  不是核心点,且  $a \in N_\varepsilon(b)$ , 则  $p$  是边界点;

(4) 噪声点:  $a$  不是核心点也不是边界点,称  $a$  为噪声点;

(5) 直接密度可达:若  $a \in N_\varepsilon(b)$ ,  $b$  是核心对象,则点  $a$  到  $b$  直接密度可达;

(6) 密度可达:若存在一系列点  $a_1, a_2, \dots, a_k, \dots, a_1 = b, a_k = a$ , 且  $a_k$  到  $a_{k+1}$  直接密度可达,则点  $a$  到点  $b$  密度可达;

(7) 密度相连:若存在数据点  $c \in D$  到点  $a$  和点  $b$  都密度可达,则  $a$  到  $b$  密度相连。

DBSCAN 算法工作流程如图 1 所示,若数据集  $D$  中所有聚类簇为  $I_1, I_2 \dots I_n$ , DBSCAN 聚类结果示意图如图 2 所示。

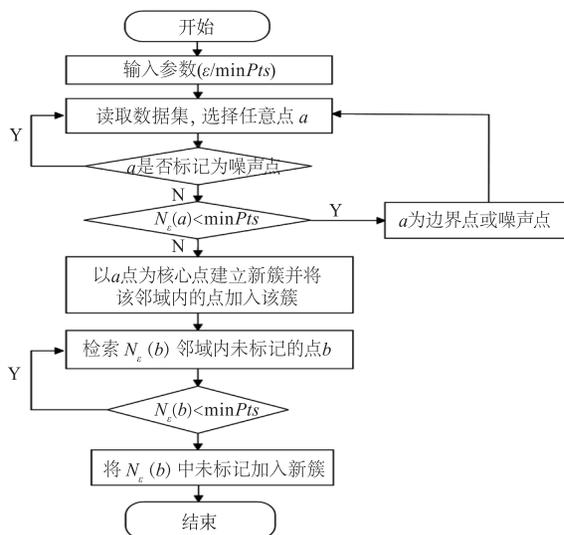


图 1 DBSCAN 算法工作流程

Fig. 1 Workflow of DBSCAN algorithm

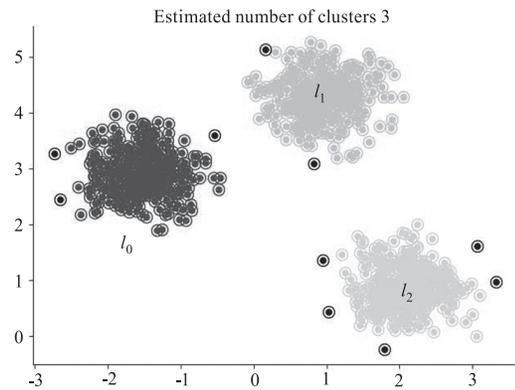


图 2 DBSCAN 聚类结果示意图

Fig. 2 Schematic representation of DBSCAN clustering results

### 2.3 特征提取

DBSCAN 聚类算法中包含两个邻域参数,  $\varepsilon$  是半径参数,  $\text{minPts}$  是密度阈值。图 3 是选择 6 个不同邻域半径进行对比实验的 DBSCAN 聚类结果图, 通过  $I_0, I_1, I_2$  三个簇的形状分布及噪声情况可以看出,  $\varepsilon$  取值 0.3 时, 噪声最小, 聚类效果最优。

### 3 基于灰色关联的设备故障预警

在数据采集过程中由于内部和外部干扰因素所采集到的数据往往具有不确定性, 根据灰色关联分析思想, 如果系统中两个因素之间的动态变化趋势程度较高, 则这两个因素关联性较高, 反之亦然。在本节中对于设备故障分析, 对聚类后的数据, 选择设备在故障前后测得的数据为参数, 通过计算出设备在故障前后两组数据之间的灰色关联系数, 分析灰色关联系数的变化程度, 通过检测其变化程度的突变位置进行设备故障识别。

电气设备故障发生后, 温度也会随之变化, 可以根据温度的变化对故障进行识别。在设备正常运行状态下得到的时间序列数据作为参考序列, 故障时获取的时间序列数据为比较序列, 通过计算出灰色关联系数的变化程度, 通过变化程度的突变点进行故障识别。公式(7)给出设备故障变化前后的灰色关联系数:

$$r(j) = \frac{\min_i \min_j |\varepsilon_0(j) - \varepsilon_i(j)| + \alpha \max_i \max_j |\varepsilon_0(j) - \varepsilon_i(j)|}{|\varepsilon_0(j) - \varepsilon_i(j)| + \alpha \max_i \max_j |x_0(j) - x_i(j)|} \quad (7)$$

灰色关联度分析中, 可以将时间序列数据在几何上绘制成曲线, 通过比较曲线形状之间的相似性来确定各因素之间的相关性。在上式中,  $\varepsilon_0(j)$  为设备正常运行时测得  $j$  点的数据,  $\varepsilon_i(j)$  是设备故障时刻  $j$  点数据。

灰色关联度计算步骤如下:

(1) 序列初始化, 确定参考序列为  $\varepsilon_0$ , 比较序列为  $\varepsilon_i$ , 由于原始序列单位不同, 数据之间差别性较大, 会影响数据分析结果, 因此要对数据进行无量纲处理, 使其大小落在  $[0, 1]$  之间。

(2) 求序列差:

$$\Delta_i(j) = |\varepsilon_0(j) - \varepsilon_i(j)|, j = 1, 2, \dots, n; i = 1, 2, \dots, m \quad (8)$$

(3) 求两极差, 最大值为  $A = \Delta_{\max}$ , 最小值为  $a = \Delta_{\min}$ 。

(4) 求设备故障前后  $j$  点的关联系数:

$$r_i(j) = \frac{(a + \alpha A)}{\Delta_i(j) + \alpha A}, j = 1, 2, \dots, n; i = 1, 2, \dots, m \quad (9)$$

其中,  $\alpha$  为分辨系数, 一般取 0.5。根据公式(9) 计算得到设备故障前后得到的时间序列数据之间的灰色关联系数曲线图, 计算曲线的变化程度, 通过变化程度的突变点来进行故障预警, 变化程度可由式表示:

$$V_{ar} = \frac{r_{j+1} - r_j}{r_j} \quad (9)$$

其中,  $j$  为原始时间序列长度, 使用  $V_{ar}$  突变点进行故障预警。

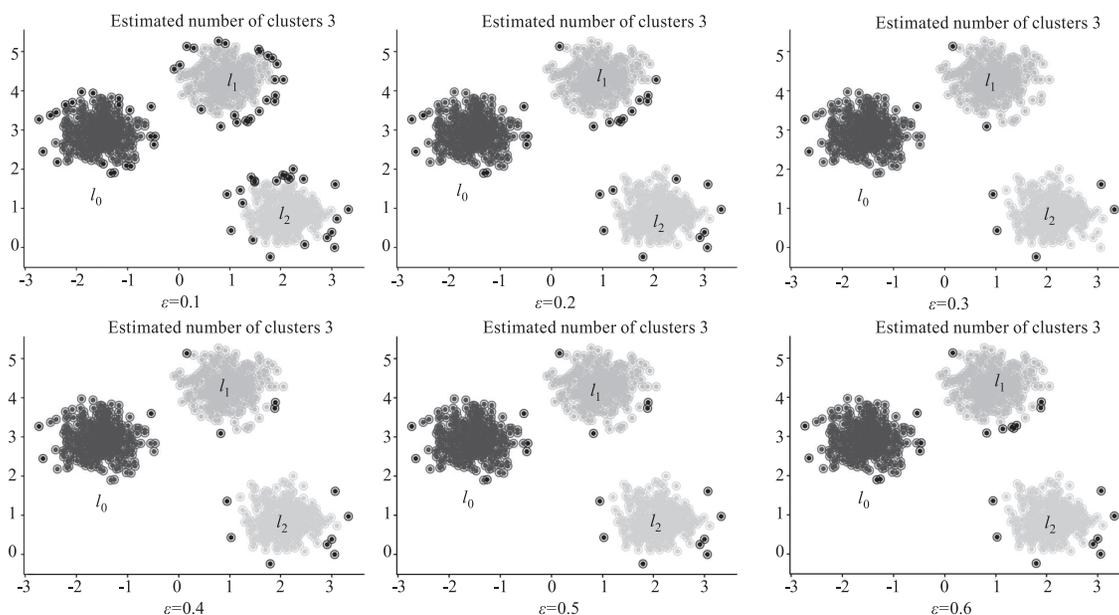


图3 DBSCAN 聚类结果图

Fig. 3 DBSCAN clustering result map

## 4 实验结果与讨论

### 4.1 故障预警流程

图4 为电气设备故障检测与预警的流程, 包括数据采集、数据预处理、数据建模、关联分析故障预警四个步骤。

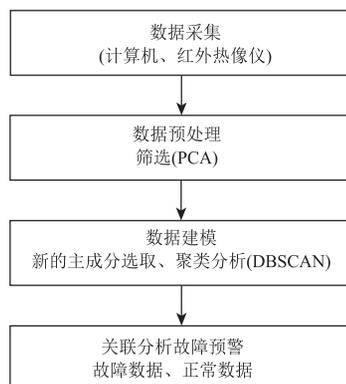


图4 故障预警流程

Fig. 4 Fault warning process

#### (1) 数据采集

由图4, 首先通过计算机监控系统和红外热像仪实时采集电气设备的相关变量数据后上传到计算机监控系统, 然后由软件平台对数据进行整合。

#### (2) 数据预处理

直接采集到的数据质量不理想, 需要经过预处理以免影响准确的诊断结果。首先, 考虑到传感器故障、采集错误或其他因素, 数据可能存在缺失、偏移等异常情况, 因此需要对数据进行整理和补充。此外, 直接将初始采集的数据用于建模可能导致模型的精度下降。因此, 在构建模型之前, 需要对数据中的变量进行筛选和加工。最后, 为了解决数据中可能存在的冗余信息, 以及减少变量的维度, 选择应用主成分分析法对数据进行降维, 以获得新的主成分变量。

#### (3) 数据建模

选取新的主成分变量的累计贡献率大于等于85%的主成分,这样可以在降低维度的同时尽量保留数据的重要特征。其次采用 DBSCAN 聚类算法对数据聚类,通过将数据点分配到不同的聚类中,识别数据中的不同模式和异常情况。

#### (4) 关联分析故障预警

在没有数据标签的情况下,通过聚类分析在聚类模型中进行故障判定。在聚类后,将聚类结果与灰色关联分析相结合,通过计算关联系数变化程度的突变点,以实现故障预警。

### 4.2 状态运行参量筛选

传感器上传的原始时间序列数据一般会存在缺失、噪声、异常等,需要进行预处理恢复数据的完整性后再对数据进行分析,具体处理步骤如下:

(1) 缺失值处理:检查数据中是否存在缺失值,如果有缺失值,采取填充或删除的方式进行处理。使用线性插值方法填充,若缺失值较多或无法填充则删除该样本。

(2) 噪声处理:分析数据中的离群值和错误值。对于离群值,根据领域知识决定是否删除或修正。

(3) 异常值筛选:根据已经筛选出的变量,对时间序列数据进行异常值检测,使用箱线图检测识别和过滤异常值,然后对异常数据点使用数据组平均值修正。

### 4.3 结合 PCA 特征降维与 DBSCAN 特征聚类

#### 4.3.1 PCA 降维

以新疆公司吉林台流域梯级水电站电气设备为例进行实例分析,从该水电站计算机监控系统系统中获取的从2023年8月20日,8点至2023年8月27日8点,正常和故障共计1008个状态的水电站电气设备历史数据,其中772条用于故障预警模型训练,236条用于故障模型验证,采样间隔均为10 min。传感器实时记录电压、电流、水位流量、功率、温度、湿度、开关状态等与电气设备运行状态有关的7个状态参量,如表1所示,通过对这7个状态参量分析,运用主成分分析法对数据进行降维处理,得到新的主成分。

表2为采用PCA算法对水电站电气设备7个运行状态参量进行数据降维后的特征值和累积贡献率。选择PCA算法中模型参数  $n\_components = 4$ ,提取了累积贡献率为85%以上4个主成分。前4

个主成分能够以89.27%的准确度代表原始数据的大部分信息,保留了原始数据的关键特征,成功实现了对原始样本数据的属性降维操作。由表2分析结果可视化后得到图5,可以看出第一主成分累积贡献率约为96.29%,第二主成分累积贡献率约为92.43%,第三主成分累积贡献率约为88.31%,第四主成分累积贡献率约为89.27%,4个主成分包含了数据的绝大多数信息,因此可以对其聚类,能达到降维和简化计算过程的效果。

表1 电气设备状态运行参量

Tab. 1 Conditional operating parameters of electrical equipment

序号	参数名称	单位
1	电压	V
2	电流	A
3	水位流量	m <sup>3</sup> /s
4	功率	W
5	温度	℃
6	湿度	%
7	开关状态	/

表2 PCA 分析结果

Tab. 2 Results of PCA analyses

主成分数量/个	特征值	累积贡献率
成分1	0.7569	0.9629
成分2	0.8094	0.9243
成分3	0.8453	0.8830
成分4	0.8580	0.8927

#### 4.3.2 DBSCAN 聚类

DBSCAN 是无监督聚类,不需要数据标签,因此,按照正常数据和故障数据进行分类,通过上述PCA降维后,论文选取前2个主成分构成特征样本集,然后采用DBSCAN算法分别对正常样本和故障样本数据进行聚类。正常数据特征样本集和故障数据特征样本集聚类结果如图6所示,由图5、图6可知,特征样本集中不同状态参量的正常和故障的样本数据经过DBSCAN聚类很好的形成聚类簇,不同状态参量在正常和故障情况的下的聚类具有明显的界限,聚类簇的形成说明了样本数据在特征空间中具有相似的特性和较好的分离

性和可区分性。

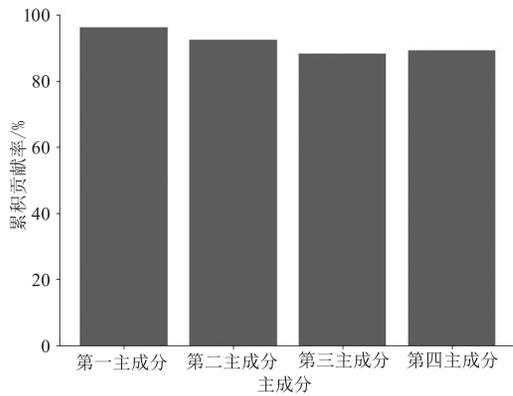
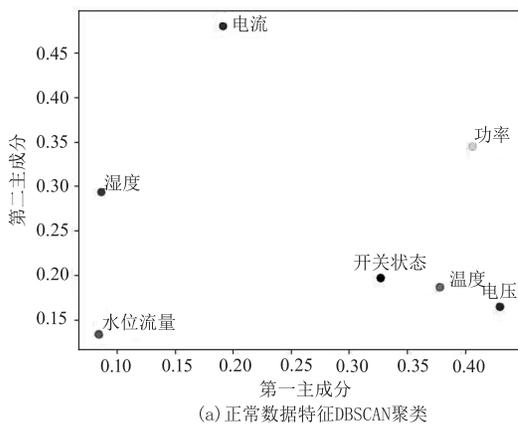
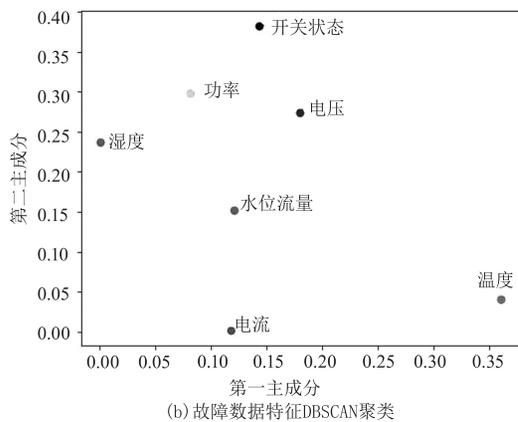


图5 主成分累计贡献率

Fig. 5 Cumulative contribution of principal components



(a) 正常数据特征DBSCAN聚类



(b) 故障数据特征DBSCAN聚类

图6 正常数据和故障数据特征 DBSCAN 聚类结果图

Fig. 6 DBSCAN clustering result plot of normal and fault data features

为了评估待检测数据与特征样本集中数据的相似性,论文选取了特征样本集的聚类中心,通过使用聚类中心的代表性向量计算待测数据和特征样本集之间的欧氏距离。由于7个状态参量的运行状态均会对电气设备温度产生不同程度的影响,因此计算了待检测数据与特征样本集之间的聚类中心,结果如表3所示。

表3 正常-故障数据聚类中心

Tab. 3 Normal-fault data clustering centre

状态参量	正常数据聚类中心		故障数据聚类中心	
	成分1	成分2	成分1	成分2
电压	0.430	0.165	0.180	0.274
电流	0.191	0.480	0.118	0.002
水位流量	0.084	0.134	0.121	0.152
功率	0.406	0.345	0.081	0.298
温度	0.378	0.187	0.361	0.041
湿度	0.086	0.294	0.001	0.236
开关状态	0.327	0.197	0.143	0.382

由于电气设备在运行过程中有时会处于故障的临界位置,选取电压和电流两种状态参量来测试验证数据的有效性。表4为待测数据到聚类中心的欧氏距离,电压经过降维后与特征样本集中电压对应聚类中心的欧氏距离为1.22,与其他特征参量相比相距的欧氏距离值最小。电流经过降维后与特征样本集中电流对应聚类中心的欧氏距离为2.57,其他特征参量相比相距的欧氏距离值最小,实验结果表明,结合PCA-DBSCAN算法能够实现数据特征降维和聚类以及较好的完成待测数据与特征样本集之间的特征匹配。

表4 待测数据到聚类中心的欧氏距离

Tab. 4 Euclidean distance from the data to be measured to the clustering centre

状态参量	待测数据到聚类中心的欧氏距离	
	电压	电流
电压	<b>1.22</b>	<b>2.57</b>
电流	5.65	3.18
水位流量	4.50	3.43
功率	6.24	9.68
温度	4.70	3.22
湿度	4.87	8.55
开关状态	5.15	4.79

#### 4.4 灰色关联分析故障预警

完成了DBSCAN聚类后,已经得到了待测数据与特征样本集之间的特征匹配,为了进一步评估各状态参量对设备故障的影响,采用灰色关联分析方法建立灰色关联模型计算灰色关联系数,以在设备正常运行的状态下得到的时间序列数据作为参考序列,故障时获取的时间序列数据为比

较序列,利用公式(9)和(10)计算出两组数据对应点之间的灰色关联系数的变化程度,通过变化程度的突变点进行故障识别。图7为特征样本集电压和电流数据灰色关联系数变化程度曲线图。可以看出,在一个小时时间段内,设备故障前后电压的灰色关联系数变化率前后灰色关联系数变化率在第13 min左右出现了明显的一个向上的突变,电流的灰色关联系数变化率在第18 min左右出现了明显的一个向上的突变。

为进一步验证论文所提算法对设备故障的预警效果,分别选取上述实验涉及到的7种状态参量236个样本数据进行实验验证,使用混淆矩阵来评估故障预警的分类情况,准确率指标如表5所示。

由表5故障识别结果可知,各类状态参量结果的准确率各项指标都达到了95%以上,图8为7种状态参量的故障预警验证图,可以从图中看出,7种参量的状态在第10~20 min之间都会有一个突变,这也验证了所提结合PCA和DBSCAN与灰色关联分析的设备故障预警具有较好的预警效果。

表5 故障识别结果评价

Tab. 5 Evaluation of fault warning results

状态参量	电压	电流	水位流量	功率	温度	湿度	开关状态
准确率/%	98.82	99.30	98.43	98.11	97.12	97.66	95.75

## 5 结论

通过对电气设备故障前期历史数据的分析,实现了对设备故障的预警。论文提出了结合PCA-DBSCAN的红外特征提取与灰色关联分析的设备故障预警模型。采用主成分分析法对水电站电气设备运行状态相关的7个状态参量降维特征提取至4个主成分,可涵盖原有运行状态的信息,从而减少了算法的运行工作量。然后对降维后的数据采取无监督聚类DBSCAN算法,构建正常数据和故障数据特征样本集,提取聚类中心完成待测数据与特征样本集之间的特征匹配。根据特征样本集之间灰色关联系数变化率的突变点来进行故障预警。结果表明,论文提出的方法可以对电气设备故障状态进行预警,对水电站电气设备故障预警具有重要意义。

## 参考文献:

[1] Liao Panpan, Zhang Jiamin. Research on influence factors

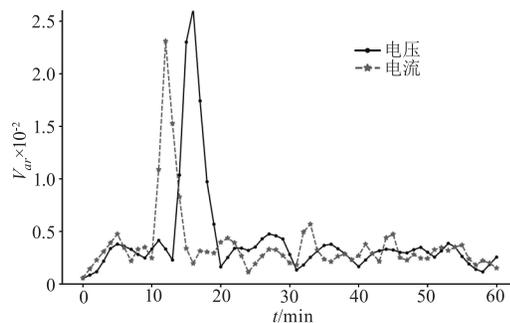


图7 特征样本集灰色关联系数变化率

Fig. 7 Rate of change of grey correlation coefficient of the feature sample set

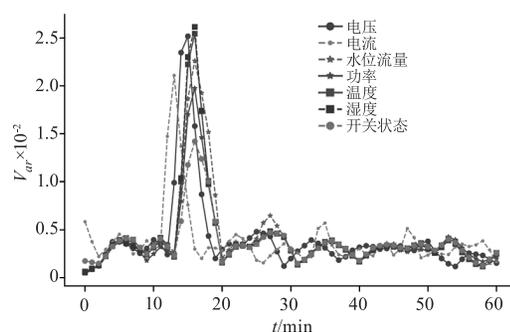


图8 故障预警验证图

Fig. 8 Fault warning validation chart

for measuring and method of correction in infrared thermometer[J]. Infrared Technology, 2017, 39(2): 173-177. (in Chinese)

廖盼盼,张佳民. 红外测温精度的影响因素及补偿方法的研究[J]. 红外技术, 2017, 39(2): 173-177.

- [2] Lü Ang, Chen Yi, Fang Jinyong, et al. Defect detection for circuit-board-mounted devices based on infrared imaging[J]. Laser & Infrared, 2018, 48(5): 579-584. (in Chinese)  
吕昂,陈怡,方晋甬,等. 基于红外热成像的电路板载器件故障检测[J]. 激光与红外, 2018, 48(5): 579-584.

- [3] Hu Fankui. Classification and thermal fault diagnosis of substation equipment based on infrared images[D]. Harbin: Heilongjiang University, 2020. (in Chinese)  
胡凡奎. 基于红外图像的变电站设备分类及热故障诊断[D]. 哈尔滨: 黑龙江大学, 2020.

- [4] Li Bangyan, Qi Weiqiang, Gui Yuan, et al. Mechanical fault diagnosis of high voltage circuit breaker based on Relief F and grey relation analysis[J]. High Voltage Apparatus, 2020, 56(6): 197-203, 210. (in Chinese)

- 李邦彦,齐伟强,桂媛,等.基于 Relief F 和灰色关联分析的高压断路器机械故障诊断[J].高压电器,2020,56(6):197-203,210.
- [5] Dragomir A, Adam M, Andruçă M, et al. Considerations regarding infrared thermal stresses monitoring of electrical equipment[C]//2017 International Conference on Electromechanical and Power Systems (SIELMEN). IEEE, 2017:100-103.
- [6] Li B, Zhu X, Zhao S, et al. HV power equipment diagnosis based on infrared imaging analyzing[C]//2006 International Conference on Power System Technology. IEEE, 2006:1-4.
- [7] Cao Y, Gu X, Jin Q. Infrared technology in the fault diagnosis of substation equipment[C]//2008 China International Conference on Electricity Distribution. IEEE, 2008:1-6.
- [8] Hui Z, Fuzhen H. An intelligent fault diagnosis method for electrical equipment using infrared images [C]//2015 34th Chinese Control Conference (CCC). IEEE, 2015:6372-6376.
- [9] Li xin, Cui Haoyang, Xu Yongpeng, et al. R-earch on infrared image feature extraction and fault diagnosis of power equipment[J]. Laser & Infrared, 2018, 48(5):659-664. (in Chinese)  
李鑫,崔昊杨,许永鹏,等.电力设备 IR 图像特征提取及故障诊断方法研究[J].激光与红外,2018,48(5):659-664.
- [10] Liao Fanglun, Zhou Ping. Fault diagnosis method of electrical equipment in thermal power plant based on PCA [J]. Techniques of Automation and Applications, 2022, 41(3):39-43. (in Chinese)  
聊方伦,周平.基于 PCA 的火电厂电气设备故障诊断方法[J].自动化技术与应用,2022,41(3):39-43.
- [11] Zhu Minghao. Failure warning method of dust removal equipment in power plant based on improved PCA-SARMA and its application [J]. Inner Mongolia Electric Power, 2023, 41(3):86-90. (in Chinese)  
朱明皓.基于改进 PCA-SARMA 的电厂除尘设备故障预警方法及应用[J].内蒙古电力技术,2023,41(3):86-90.
- [12] Liu X, Li J, Shao L, et al. Transformer fault early warning analysis based on hierarchical clustering combined with decision trees[J]. Energies, 2023, 16(3):1168.
- [13] Li Y, Hong F, Tian L, et al. Condition monitoring and early fault warning of power plant auxiliary equipment using LSTM-SDAE model with an adaptive threshold[J]. Frontiers in Energy Research, 2023, 10:1037539.
- [14] Bai J, Cong P, Yu L, et al. Research on transformer fault diagnosis method based on principal component analysis and grey correlation analysis [C]//2021 IEEE International Conference on Power, Intelligent Computing and Systems(ICPICS). IEEE, 2021:508-510.
- [15] Wang Manran. Study of substation equipment fault detection and recognition based on infrared thermal image [D]. Enshi: Hubei Minzu University, 2020. (in Chinese)  
王曼然.基于红外热像图的变电设备故障检测与识别研究[D].恩施:湖北民族大学,2020.
- [16] Zhang Qiuxin, Zhou Jinjun, Chen Feng, et al. Fault diagnosis method of rolling bearings of packaging equipment based on DS-PCA Model [J]. Packaging Engineering, 2023, 44(9):224-231. (in Chinese)  
张秋昕,周进军,陈锋,等.基于 DS-PCA 模型的包装设备滚动轴承故障诊断方法研究[J].包装工程,2023,44(9):224-231.
- [17] Sarita K, Devarapalli R, Kumar S, et al. Principal component analysis technique for early fault detection[J]. Journal of Intelligent & Fuzzy Systems, 2022, 42(2):861-872.
- [18] Ma Liangyu, Sun Jiaming, Yu Shilei, et al. DBSCAN and SDAE-based abnormal condition early warning for wind turbine unit[J]. Journal of Chinese Society of Power Engineering, 2021, 41(9):786-793, 808. (in Chinese)  
马良玉,孙佳明,於世磊,等.基于 DBSCAN 和 SDAE 的风电机组异常工况预警研究[J].动力工程学报,2021,41(9):786-793,808.