

DOI: 10.19666/j.rlfed.202212221

# 基于 PCA-FINCH 的磨煤机故障诊断方法

钱虹<sup>1,2</sup>, 张现涛<sup>1</sup>

(1.上海电力大学自动化工程学院, 上海 200090;  
2.上海市电站自动化技术重点实验室, 上海 200072)

**[摘要]** 针对类似磨煤机故障等小概率发生事件难以提取且用于机器学习分类导致的故障诊断精度低的问题, 提出了一种基于小样本的 PCA-FINCH 高精度故障诊断方法。首先, 基于主元分析 (PCA) 对表征设备运行状态的历史数据进行故障检测, 通过  $T^2$  控制限与  $Q$  控制限来检测故障的发生并识别故障样本, 提取故障样本从而组成小样本故障集; 然后, 基于 FINCH 分类器, 对获取的小样本故障集进行精确分类, 实现对设备的故障诊断; 最后, 使用包含有磨煤机故障的历史数据集对该方法进行验证。结果表明, 提出的 PCA-FINCH 故障诊断方法能够对小样本故障实现高精度分类, 其在精确度上, 较决策树 CART、随机森林 RF、支持向量机 SVM 分别提高了 2.61 百分点、1.74 百分点、1.85 百分点, 其在收敛速度上也表现优异。

**[关键词]** 磨煤机; 故障诊断; 小样本; FINCH 聚类; 主元分析

**[引用本文格式]** 钱虹, 张现涛. 基于 PCA-FINCH 的磨煤机故障诊断方法[J]. 热力发电, 2023, 52(9): 147-154. QIAN Hong, ZHANG Xiantao. Fault diagnosis method of coal mill based on PCA-FINCH[J]. Thermal Power Generation, 2023, 52(9): 147-154.

## Fault diagnosis method of coal mill based on PCA-FINCH

QIAN Hong<sup>1,2</sup>, ZHANG Xiantao<sup>1</sup>

(1.College of Automation Engineering, Shanghai University of Electric Power, Shanghai 200090, China;  
2.Shanghai Key Laboratory of Power Station Automation Technology, Shanghai 200072, China)

**Abstract:** Aiming at the problem of low probability of occurrence events such as coal mill failures that are difficult to extract and used for machine learning classification, resulting in low fault diagnosis accuracy, a PCA-FINCH high-precision fault diagnosis method for small samples is proposed. Firstly, based on principal component analysis PCA, fault detection is carried out on the historical data that characterizes the operating state of the equipment, and the occurrence of faults is detected and the fault samples are identified through the  $T^2$  control limit and the  $Q$  control limit, and the fault samples are extracted to form a small sample fault set; Secondly, based on the FINCH classifier, the obtained small sample fault set is accurately classified to realize the fault diagnosis of the equipment. Finally, the method is verified using a historical data set containing coal mill faults. The results show that the PCA-FINCH fault diagnosis method proposed can achieve high-precision classification of small-sample faults, and its accuracy is 2.61 percentage points, 1.74 percentage points and 1.85 percentage points higher than that of decision tree CART, random forest RF and support vector machine SVM, respectively, and its convergence speed is excellent.

**Key words:** coal mill; fault diagnosis; small sample; FINCH clustering; principal component analysis

磨煤机是火力发电机组燃烧制粉系统的重要设备, 由于磨煤机工作环境较为恶劣, 其比较容易出现故障, 若磨煤机设备状态异常或出现故障, 将会直接影响到机组出力, 甚至出现严重事故<sup>[1-3]</sup>。在复杂环境下实现对磨煤机设备准确、快速的故障诊

断, 能够帮助检修人员采取及时有效的修复措施, 这对保证电厂安全稳定运行、提高生产效率具有重要意义。

目前, 有关电站设备的故障诊断研究大多是基于数据驱动建模的机器学习方法。文献[4-6]基于传

修回日期: 2022-12-18 网络首发日期: 2023-04-27

基金项目: 上海市自然科学基金项目 (19ZR1420700)

Supported by: Natural Science Foundation of Shanghai (19ZR1420700)

第一作者简介: 钱虹 (1967), 女, 博士, 教授, 主要研究方向为故障诊断、过程控制等, qianhong.sh@163.com。

通信作者简介: 张现涛 (1996), 男, 硕士研究生, 主要研究方向为智能故障诊断, 15222590800@163.com。

统 BP 神经网络建立分类器模型,对电厂锅炉、凝汽器等设备进行了故障诊断。但 BP 神经网络在训练过程中容易陷入局部极小值,且当样本数量过大时神经网络的收敛速度较慢。文献[7-8]采用概率神经网络代替 BP 神经网络来构建故障诊断分类器,降低了神经网络的训练时间,一定程度上提高了诊断准确率。文献[9-10]对 BP 神经网络的学习算法进行改进,采用动量法与自适应学习率法进行误差反向传播,有效提高了网络训练速度。文献[11]采用遗传算法优化 BP 网络的初始权值、阈值以及学习速率等结构参数,从而构建出具有最佳结构参数的神经网络。支持向量机是机器学习在故障诊断领域应用的又一代表性算法<sup>[12-13]</sup>。文献[14-15]基于支持向量机分类器对磨煤机、汽轮机进行了故障诊断,并利用仿真实验数据对分类器性能进行验证。支持向量机依赖于核函数参数与惩罚因子等模型参数,一旦参数选择不当将严重影响分类性能。文献[16]采用蝙蝠算法优化支持向量机的惩罚因子与核参数。遗传算法<sup>[17]</sup>、天牛群算法<sup>[18]</sup>等寻优算法也被应用到支持向量机的参数优化当中以获得模型的最佳性能。此外,随机森林 RF<sup>[19]</sup>、极限学习机 ELM<sup>[20]</sup>等机器学习算法也被尝试应用到电厂设备的故障诊断。然而上述诊断方法都存在一定局限性,对故障样本数量需求过大,以 BP 神经网络、支持向量机 SVM 等分类器为代表的机器学习算法需要充足数量的故障样本来进行模型训练。但现场故障发生的概率小、故障样本少且难以获取,这导致模型泛化性能低,故障诊断准确率低,实时性差。因此,机器学习算法不能够很好地应用到电厂实际场景之中。

针对当前机器学习分类器所需的大样本故障难以获取或导致分类精度低的问题,本文提出一种主元分析(PCA)与 FINCH 聚类相结合的故障诊断方法,基于 PCA,从表征设备运行状态的海量历史数据中利用 2 个控制限  $T^2$  与  $Q$  对故障样本的甄别来获取小概率发生的故障样本集;再将故障样本集送入 FINCH 分类器,利用邻接矩阵与有向图的小样本聚类能力,实现精确的故障分类。将本方法用于包含有磨煤机故障的历史数据样本集中,能够实现磨煤机故障类型的准确分类,研究结果表明该方法有效。

## 1 基于 PCA-FINCH 的故障诊断方法

### 1.1 基于 PCA 的故障检测与故障特征提取

PCA 是将数据矩阵分解成一个主成分空间

和一个残差子空间<sup>[21-23]</sup>。对于磨煤机的样本数据集  $\mathbf{X} \in \mathbb{R}^{N \times m}$ , 可以按式(1)分解:

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}_1^T + \mathbf{t}_2 \mathbf{p}_2^T + \cdots + \mathbf{t}_l \mathbf{p}_l^T + \mathbf{E} = \mathbf{TP}^T + \mathbf{E} \quad (1)$$

式中:  $N$  为时序样本的长度;  $m$  为监测变量个数;  $\mathbf{T} \in \mathbb{R}^{N \times l}$  为得分矩阵;  $l$  为所选取主元的个数;  $\mathbf{t}_i$  为得分向量;  $\mathbf{P} \in \mathbb{R}^{m \times l}$  为负载矩阵;  $\mathbf{p}_i$  为负载向量;  $\mathbf{E}$  为残差矩阵。

PCA 通过监测每一个时序样本的  $T_i^2$  值与  $Q_i$  值来判断磨煤机设备是否异常或故障。 $T_i^2$  值与  $Q_i$  值由式(2)、式(3)给出:

$$T_i^2 = \mathbf{t}_i \boldsymbol{\lambda}^{-1} \mathbf{t}_i^T = \mathbf{x}_i \mathbf{p} \boldsymbol{\lambda}^{-1} \mathbf{p}^T \mathbf{x}_i^T \quad (2)$$

式中:  $\boldsymbol{\lambda}$  为由  $l$  个特征值构成的对角矩阵;  $T^2$  用于评价被观测量的主要变化,如果潜在变量的变化大于正常情况下的变化,则表示设备状态异常。

$$Q_i = \mathbf{e}_i^T \mathbf{e}_i = \mathbf{x}_i^T (\mathbf{I} - \mathbf{p} \mathbf{p}^T) \mathbf{x}_i \quad (3)$$

式中:  $Q$  是统计量,为预测误差的平方,用来监控新样本对模型的适应性。

PCA 监测到样本的  $T_i^2$  值或者  $Q_i$  值超过监测阈值  $T_\alpha^2$ 、 $Q_\alpha$  时,即可判断磨煤机设备异常或故障。监测阈值  $T_\alpha^2$ 、 $Q_\alpha$  由式(4)、式(5)给出:

$$T_\alpha^2 = \frac{k(m^2 - 1)}{m(m - k)} F(k, m - k, \alpha) \quad (4)$$

式中:  $F(k, m - k, \alpha)$  是一个具有  $m - k$  个自由度和置信度为  $\alpha$  的  $F$  分布;  $T_\alpha^2$  为置信度  $\alpha$  下的控制限。

$$Q_\alpha = \theta_1 \left[ \frac{\eta_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1} \right]^{\frac{1}{h_0}} \quad (5)$$

式中:  $\theta_i = \sum_{j=1}^m \lambda_j^i, (i = 1, 2, 3)$ ;  $h_0 = 1 - 2\theta_1 \theta_2 / 3 \theta^2$ ;  $Q_\alpha$

为置信度  $\alpha$  下的控制限。

PCA 检测到故障发生后,利用  $T^2$  贡献图与  $Q$  贡献图来评价每一变量对于当前故障的重要性,从而筛选出当前故障的故障特征。

$$\text{Cont}_i^Q = (\boldsymbol{\xi}_i^T \mathbf{C} \mathbf{x})^2 \quad (6)$$

式中:  $\text{Cont}_i^Q$  为每个变量对  $Q$  统计量的贡献度;

$\tilde{\mathbf{C}} = \mathbf{I} - \mathbf{P} \mathbf{P}^T$ ;  $\boldsymbol{\xi}_i^T$  为单位矩阵的第  $i$  列。

$$\text{Cont}_i^{T^2} = (\boldsymbol{\xi}_i^T \mathbf{D}^{1/2} \mathbf{x})^2 = \mathbf{x}^T \mathbf{D}^{1/2} \boldsymbol{\xi}_i \boldsymbol{\xi}_i^T \mathbf{D}^{1/2} \mathbf{x} \quad (7)$$

式中:  $\mathbf{D} = \mathbf{P}^T \mathbf{A}^{-1} \mathbf{P}$ ;  $\text{Cont}_i^{T^2}$  为每个变量对  $T^2$  统计量的贡献度。

利用  $T^2$  贡献图与  $Q$  贡献图,从 PCA 监测的全部状态变量中筛选出贡献度较大的若干变量  $\{s_1, s_2, \dots, s_p\}$  作为故障特征集合,提取故障时段内的故

障特征样本, 组成磨煤机故障样本集  $\mathbf{X}_{\text{fault}}^{L \times p} = \{(x_{i,1}, x_{i,2}, \dots, x_{i,p}) | 1 \leq i \leq L\}$ , 将其作为磨煤机诊断模型的输入。

### 1.2 基于 FINCH 聚类的故障诊断

机器学习算法的诸多网络结构参数决定了其需要充足数量的故障样本来进行模型训练, 因此, 基于机器学习的分类器在面对小样本故障时难以实现高精度分类。FINCH 算法是一种高效且完全无参数的无监督聚类算法, 其无需预先设置相似度阈值等超参数, 且不需要关于数据分布本身的先验知识, 具有对数据自然聚类的潜力。将 FINCH 聚类应用于小样本故障诊断比较适合。

FINCH 算法核心思想是通过每个数据点的最近邻去进行连接, 从而得到数据划分<sup>[24-26]</sup> (图 1)。

FINCH 聚类	
输入:	磨煤机故障样本集 $\mathbf{X}_{\text{fault}} = \{x_1, x_2, \dots, x_p\}$ , $\mathbf{X}_{\text{fault}} \in \mathbb{R}^{L \times p}$ , 其中, $L$ 为样本数量, $p$ 为特征维度
过程:	
1.	通过精确距离或快速近似最近邻方法计算第一邻居整数向量 $\mathbf{k}^1 \in \mathbb{R}^{N \times 1}$
2.	给定 $\mathbf{k}^1$ , 根据式(8)进行聚类获得第一分区 $\Gamma_1 = \{C_1, C_2, \dots, C_{r_1}\}$ , $C_{r_1}$ 为第一分区故障类型个数
3.	cluster_num = $C_{r_1}$ , $i = 1$
4.	while cluster_num $\neq 1$ do
5.	计算当前分区 $\Gamma_i$ 内每一聚类簇 $C_i$ 的簇平均向量 $\mu_i$ , 形成簇平均向量矩阵 $\mathbf{M}$ $\mathbf{M} = \{\mu_1, \mu_2, \dots, \mu_{C_i}\}$ , $\mathbf{M} \in \mathbb{R}^{C_i \times m}$
6.	由簇平均向量矩阵 $\mathbf{M}$ , 更新第一邻居整数向量 $\mathbf{k}^1 \in \mathbb{R}^{C_i \times 1}$
7.	根据式(8)进行聚类, 得到分区 $\Gamma_M$ $\Gamma_M = \{C_1, C_2, \dots, C_{r_M}\}$ , $\Gamma_M \supseteq \Gamma_i$
8.	if $C_{r_M} = 1$ then
9.	break
10.	else
11.	更新簇的故障标签: $\Gamma_M \rightarrow \Gamma_{i+1}$
12.	$i = i + 1$
13.	end if
14.	end while
输出:	分区集 $\mathcal{L} = \{\Gamma_1, \Gamma_2, \dots, \Gamma_L\}$ , 每一分区 $\Gamma_i$ $\Gamma_i = \{C_1, C_2, \dots, C_{r_i}   C_{r_i} > C_{r_{i+1}}, \forall i \in \mathcal{L}\}$

图 1 FINCH 聚类的迭代过程  
Fig.1 The iterative process of FINCH clustering

对于磨煤机故障样本集  $\mathbf{X}_{\text{fault}} \in \mathbb{R}^{L \times p}$  ( $p$  为故障特征个数), 首先将每个时序样本  $\mathbf{x}_i (1 \leq i \leq L)$  看作一个数据点, 并得到其最近邻的信息, 根据下式求得邻接矩阵  $\mathbf{A}(i, j)$ :

$$\mathbf{A}(i, j) = \begin{cases} 1 & \text{if } j = k_i^1 \text{ or } k_j^1 = i \text{ or } k_j^1 = k_i^1 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

式中:  $k_i^1$  为第  $i$  个样本的最近邻索引;  $k_j^1$  为第  $j$  个

样本的最近邻索引。邻接矩阵  $\mathbf{A}$  通过  $j = k_i^1$  将每个数据点  $i$  链接到其第 1 个邻居, 通过  $i = k_j^1$  实施对称, 并通过  $k_i^1 = k_j^1$  链接具有相同邻居的数据点  $(i, j)$ 。式(8)得到对称稀疏的邻接矩阵, 根据此邻接矩阵可以得到一个有向图, 其中连接在一起的样本点即为同一簇。

FINCH 算法 (图 2) 首先依据快速最近邻方法计算故障数据集中每一样本的第 1 邻居, 得到表征所有样本间距离信息的第 1 邻居整数向量  $\mathbf{k}^1$ , 基于该向量求得邻接矩阵  $\mathbf{A}(i, j)$ , 进而得到一个有向图, 该有向图将链接在一起的若干样本点视为同一簇, 即同一故障类型, 至次完成第 1 次聚类; 然后 FINCH 算法通过依次更新第 1 邻居整数向量与邻接矩阵, 迭代聚类, 直至所有样本的故障标签同一。可以看出, FINCH 聚类的结果为多层树  $\mathcal{L} = \{\Gamma_1, \Gamma_2, \dots, \Gamma_L\}$ , 第  $i$  层  $\Gamma_i$  即为第  $i$  次迭代聚类的结果, 最顶层的所有样本故障类型同一。

给定磨煤机故障类型数, 在多层树的基础上, 继续应用集群算法来获得最终的分类结果。

集群算法	
输入:	故障样本集 $\mathbf{X}_{\text{fault}}$ , 分区 $\Gamma_i = \{C_1, C_2, \dots, C_{r_i}\}$ , 磨煤机故障类型数 $C_{r_{\text{req\_num}}}$ ( $C_{r_{\text{req\_num}}} < C_{r_i}$ ) 过程:
1.	for iters = $C_{r_i} - C_{r_{\text{req\_num}}}$ do
2.	计算当前分区 $\Gamma_i$ 每一聚类簇 $C_j$ 的簇平均向量 $\mu_j$ , 形成簇平均向量矩阵 $\mathbf{M} = \{\mu_1, \mu_2, \dots, \mu_{r_i}\}$ , $\mathbf{M} \in \mathbb{R}^{C_i \times m}$
3.	由簇平均向量矩阵 $\mathbf{M}$ , 计算第一邻居整数向量 $\mathbf{k}^1 \in \mathbb{R}^{C_i \times 1}$
4.	根据 $\mathbf{k}^1$ 得到式(8)所表达的邻接矩阵 $\mathbf{A}$
5.	对 $\forall \mathbf{A}(j, k) = 1$ , 保留一个具有最小距离 distance( $j, k$ ) 的对称链路 $\mathbf{A}(j, k)$ , 并将其他链路设置为 0
6.	更新簇的故障标签: 合并对应的 $(C_j, C_k)$ 簇
7.	end for
输出:	$\Gamma_{\text{req\_num}} = \{C_1, C_2, \dots, C_{r_{\text{req\_num}}}\}$ , $\Gamma_{\text{req\_num}}$ 包含磨煤机全部故障样本的故障标签

图 2 给定簇数量的 FINCH 聚类  
Fig.2 FINCH clustering with given number of clusters

### 1.3 构建 PCA-FINCH 磨煤机故障诊断模型

基于 PCA-FINCH 磨煤机故障诊断模型构建流程如图 3 所示。

基于 PCA-FINCH 的磨煤机故障诊断模型具体构建步骤如下:

1) 首先对已获得的磨煤机历史数据集进行标准化处理;

2) 建立磨煤机状态监测 PCA 模型: 对正常样本计算协方差矩阵并进行协方差分解, 提取主元, 进而训练模型得到控制限  $T^2$  统计量与  $Q$  统计量;

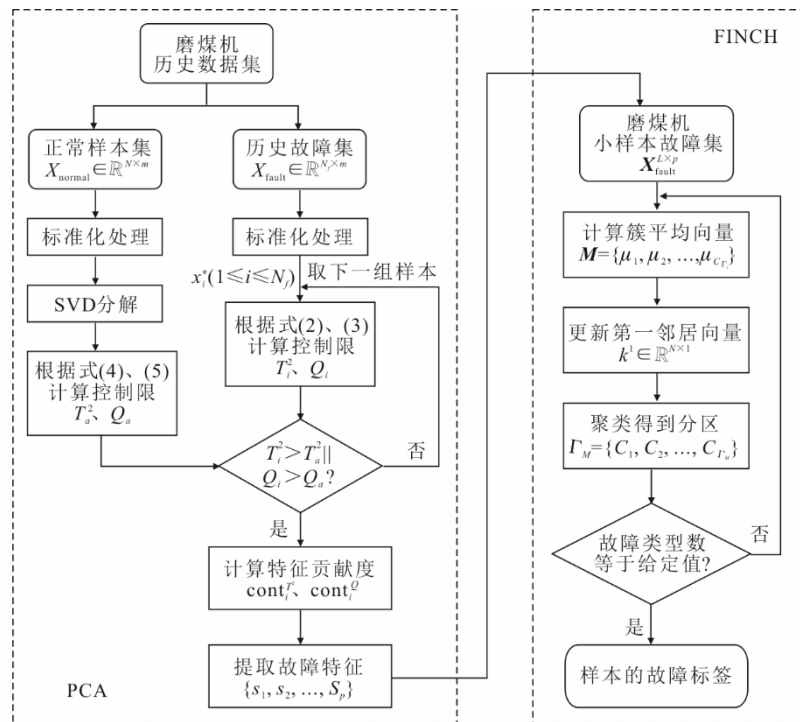


图3 基于 PCA-FINCH 磨煤机故障诊断模型构建流程

Fig.3 The construction process of coal mill fault diagnosis model based on PCA-FINCH

3) 检测故障: 若 PCA 检测到当前时序样本的  $T_i^2$  值或者  $Q_i$  值超出控制限, 则判定磨煤机设备异常或发生故障;

4) 故障特征提取: 当检测到磨煤机发生故障, PCA 对所有的监测变量进行分析, 通过计算每一状态变量对当前故障的故障贡献度  $\text{cont}_i^T$  与  $\text{cont}_i^Q$ , 形成故障贡献图, 通过故障贡献图可以准确提取磨煤机当前故障下的故障特征;

5) 获取小样本故障集: 基于 PCA, 对包含磨煤机不同故障在内的历史故障集进行处理, 分别提取不同故障类型下的故障特征, 得到故障特征集  $\{s_1, s_2, \dots, s_p\}$ , 剔除非特征变量, 提取故障时段内以  $s_1, s_2, \dots, s_p$  为特征的时序样本, 形成小样本故障集  $X_{\text{fault}}^{L \times p}$ ;

6) FINCH 对小样本故障集进行故障分类: 依次计算样本簇的平均向量矩阵  $M$  与第 1 邻居整数向量  $k^1$ , 更新邻接矩阵  $A(i, j)$ , 迭代聚类, 直至当前分区内所有故障样本的故障标签同一, 聚类停止;

7) 根据磨煤机实际发生的故障类型数, 对某一分区 (该分区内的样本故障类别数需大于给定值) 的分类结果继续应用集群算法, 最终得到所有样本的故障标签, 即故障类型。

## 2 实例验证

### 2.1 数据预处理

对数据样本进行预处理可以更准确提取数据集信息, 过程主要包含标准化处理和去除异常值 2 部分。

1) 数据标准化 数据集内各个变量具有不同的量纲且数值的量级相差很大, 如果直接将原始数据传入模型计算, 将会削弱数值较低的特征在综合分析中的作用, 使得 PCA 算法无法有效学习客观全面的数据信息, 最终导致建模失败。对于变量  $x_i \in \mathbb{R}^{N \times m}$ , 按照式(9)进行标准化计算。

$$x_{i,j}^* = \frac{x_{i,j} - \mu_i}{\sigma_i} \quad (1 \leq i \leq m, 1 \leq j \leq N) \quad (9)$$

2) 剔除异常值 异常值是错误的孤立测量值, 预处理时要将这些异常值去除, 避免对数据集的正常分布造成影响。

### 2.2 PCA 故障检测

本文以浙江华能某电厂 1 000 MW 机组的 HP1163 型中速磨煤机为研究对象, 针对堵磨故障、爆燃故障、皮带打滑故障共进行了 3 次数据采集。具体为: 1) 采集时间为 2019 年 11 月 5 日 12:00 至 15:20, 包含堵磨故障在内的 2 500 组时序数据;

2) 采集时间为 2020 年 6 月 8 日 08:00 至 11:20, 包含爆燃故障在内的 2 500 组时序数据; 3) 采集时间为 2018 年 7 月 16 日 06:40 至 08:20, 包含皮带打滑故障在内的 1 200 组时序数据。

为了对磨煤机设备进行故障检测, 首先使用表征设备正常运行状态的样本集训练 PCA, 得到观测阈值  $T^2$  统计量与  $Q$  统计量; 然后利用已训练的 PCA 对包含设备故障状态的历史数据集  $X_{mill\_blocking}^{2500 \times 6}$ 、 $X_{deflagration}^{2500 \times 6}$ 、 $X_{slipping}^{1200 \times 6}$  分别进行故障检测, 验证 PCA 故障检测的有效性; 最后提取故障特征。

1) 堵磨故障检测与故障特征提取 对包含堵磨故障的历史数据集  $X_{mill\_blocking}^{2500 \times 6}$ , 共采集磨煤机电流 (S1)、磨煤机进口一次风温 (S2)、进口一次风量 (S3)、出口风粉温度 (S4)、磨碗压差 (S5) 以及进口一次风压 (S6) 6 个相关变量。PCA 检测结果如图 4、图 5 所示。

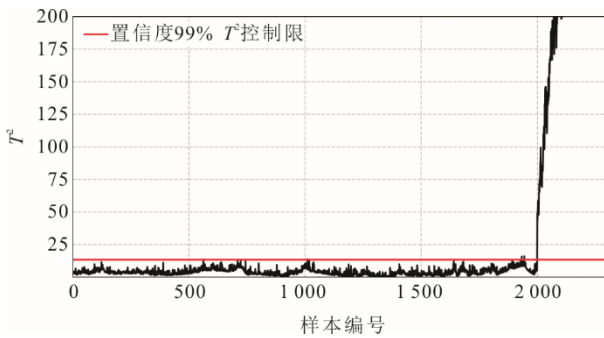


图 4 堵磨故障时置信度 99% 下的  $T^2$  检测图  
Fig.4  $T^2$  test results with 99% confidence under the blockage fault

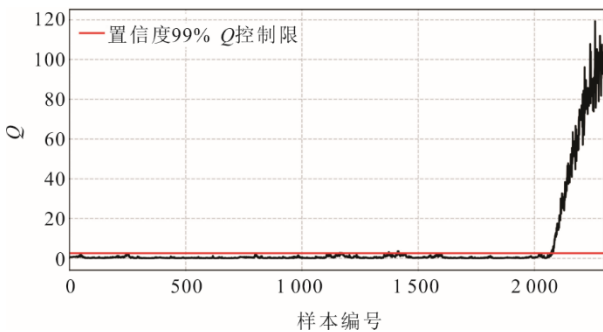


图 5 堵磨故障时置信度 99% 下的  $Q$  检测图  
Fig.5  $Q$  test results with 99% confidence under the blockage fault

由图 4 及图 5 可知, 第 2 000 个样本超  $T^2$  控制限, 随后超出  $Q$  控制限, 可以判定设备状态异常、发生故障。进一步, 通过 PCA 故障贡献图对各个变量进行故障贡献度分析, 提取堵磨故障下的故障特征, 如图 6 所示。由图 6 可以得出, 第 1 变量、第 2 变量与第 4 变量对堵磨故障的故障贡献度较大, 因

此提取磨煤机电流、磨煤机进口一次风温与磨煤机出口风粉温度作为堵磨故障的故障特征。

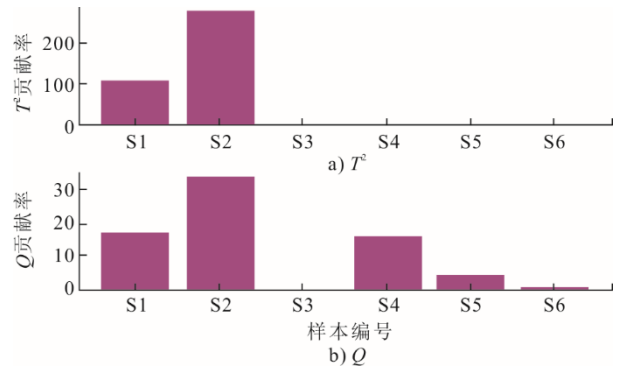


图 6 堵磨故障下各变量故障贡献度分析  
Fig.6 Analysis of the contribution degree of each variable to the fault under the blockage fault

2) 爆燃故障检测与故障特征提取 对包含爆燃故障的历史数据集  $X_{deflagration}^{2500 \times 5}$ , 共采集磨煤机进口一次风量 (S3)、磨煤机出口风粉温度 (S4)、磨碗压差 (S5)、进口一次风压 (S6)、以及出口压力 (S7) 5 个相关变量。PCA 检测结果如图 7、图 8 所示。

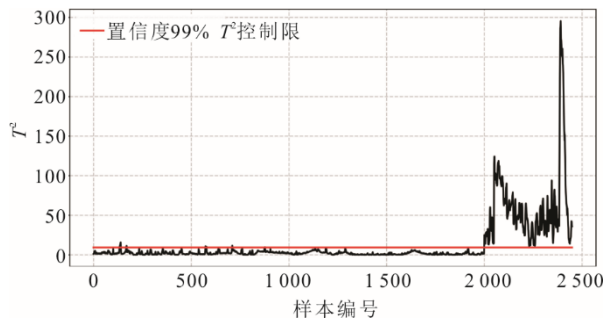


图 7 爆燃故障时置信度 99% 下的  $T^2$  检测图  
Fig.7  $T^2$  test results with 99% confidence under the deflagration fault

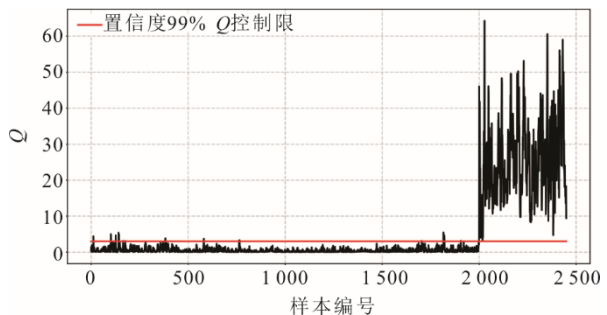


图 8 爆燃故障时置信度 99% 下的  $Q$  检测图  
Fig.8  $Q$  test results with 99% confidence under the deflagration fault

由图 7、图 8 可知, 第 2 000 个样本超出  $T^2$  控制限, 同时超出  $Q$  控制限, 可以判定设备状态异常、发生故障。然后对各个变量进行故障贡献度分析, 提取爆燃故障下的故障特征, 如图 9 所示。由图 9

可知,第 1 变量、第 2 变量与第 4 变量对爆燃故障的故障贡献度较大,因此提取磨煤机进口一次风量、磨煤机出口风粉温度与进口一次风压作为爆燃故障的故障特征。

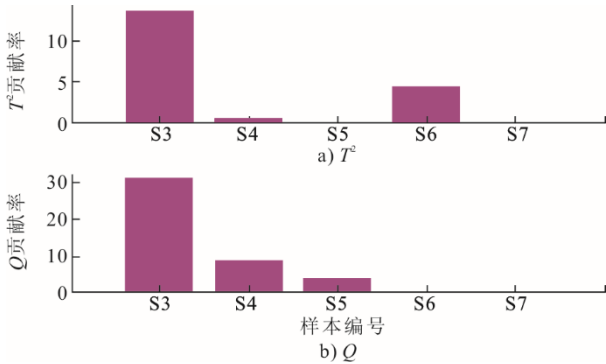


图 9 爆燃故障下各变量故障贡献度分析  
Fig.9 Analysis of the contribution degree of each variable to the fault under the deflagration fault

3) 皮带打滑故障检测与故障特征提取 对包含皮带打滑故障的历史数据集  $X_{slipping}^{1200 \times 6}$ ,共采集磨煤机电流(S1)、磨煤机进口一次风量(S3)、出口风粉温度(S4)、主蒸汽温度(S8)、再热蒸汽温度(S9)以及主蒸汽压力(S10)6个相关变量。PCA 状态检测结果如图 10、图 11 所示。

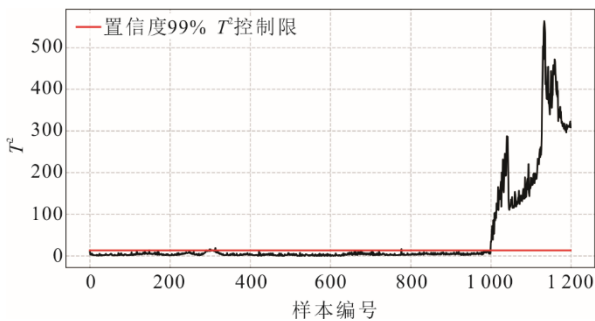


图 10 打滑故障时置信度 99% 下的  $T^2$  检测图  
Fig.10  $T^2$  test results with 99% confidence under the slipping fault

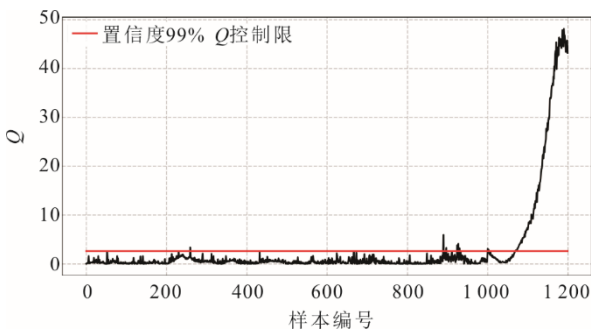


图 11 打滑故障时置信度 99% 下的  $Q$  检测图  
Fig.11  $Q$  test results with 99% confidence under the slipping fault

由图 10、图 11 可知,第 1 000 组样本超出  $T^2$

控制限,随后超出  $Q$  控制限,可以判定设备状态异常、发生故障。进一步,通过 PCA 故障贡献图对各个变量进行故障贡献度分析,提取打滑故障下的故障特征,如图 12 所示。由图 12 可以得出,第 1 变量、第 2 变量、第 3 变量与第 6 变量对皮带打滑故障的故障贡献度较大,因此提取磨煤机电流、出口风粉温度、磨煤机进口一次风量与主蒸汽压力作为皮带打滑故障的故障特征。

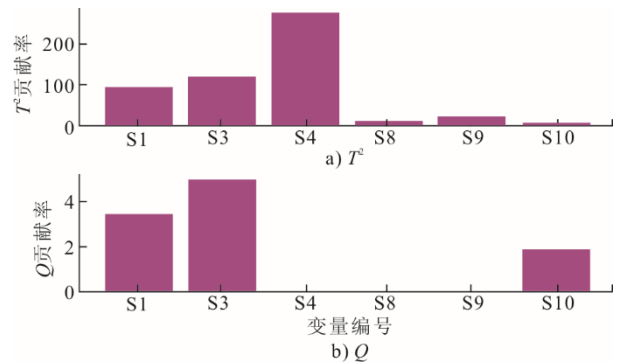


图 12 打滑故障下各变量故障贡献度分析  
Fig.12 Analysis of the contribution degree of each Variable to the fault under the slipping fault

以上,通过 PCA 方法对磨煤机设备进行了故障检测,检测图表明 PCA 能够利用控制限  $T^2_\alpha$  与  $Q_\alpha$  准确识别各个故障的发生,同时利用故障贡献图对设备的堵磨故障、爆燃故障及皮带打滑故障一一提取了故障特征,在此基础上获取了小样本故障集  $X_{fault}^{1500 \times 10}$ ,将其作为 FINCH 分类器的输入用以验证模型的分类精度。

### 2.3 FINCH 分类结果

本文引入准确率、召回率、精确度、 $F_1$  得分共 4 项评价指标来评价 FINCH 分类器对于小样本故障集的分类性能。利用 PCA 方法提取的小样本故障集对 FINCH 分类模型进行分类准确率验证,小样本故障集  $X_{fault}^{1500 \times 10}$  共有 1 500 组时序样本,其中堵磨故障样本 100 组、爆燃故障样本 200 组、皮带打滑故障样本 200 组,正常样本 1 000 组,分类结果如图 13 所示。

为突出 FINCH 模型的有效性,将 FINCH 分类器与决策树 CART 分类器、随机森林 RF 分类器及支持向量机 SVM 分类器进行对比。将故障样本集按 6:4 的比例随机划分为训练集与测试集。决策树 CART 利用学习曲线获得最大深度最优取值为 5。随机森林 RF 使用网格搜索法进行超参数寻优,得到最佳决策树棵数为 15。支持向量机 SVM 采用

RBF 核函数进行训练, 由于核参数  $C$  和惩罚系数对训练结果影响较大, 使用网格搜索法对 2 个参数进行寻优, 优化后的核参数与惩罚系数的取值分别为 1.731、0.097。各分类器收敛速度对比如图 14 所示。

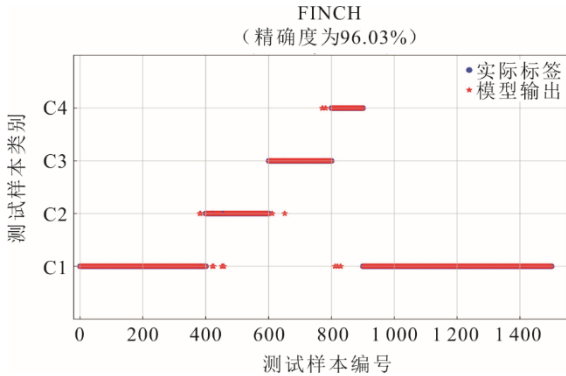


图 13 FINCH 分类结果  
Fig.13 FINCH classification results

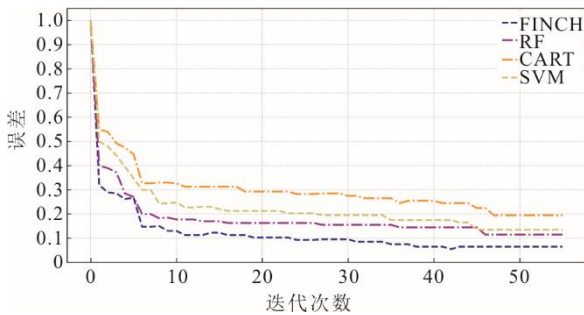


图 14 各分类器收敛速度对比  
Fig.14 Comparison of the convergence speed of each classifier

由图 14 可以看出, 在初始迭代阶段, FINCH 算法的迭代误差曲线以最大的梯度快速下降, 然后以最快的速度收敛, 其他 3 种方法的收敛速度与收敛精度要差一些。因此, FINCH 方法在实际应用中的收敛速度更快, 效率更高。各分类器对每一故障的分类准确率及性能对比分别如图 15、图 16 所示。

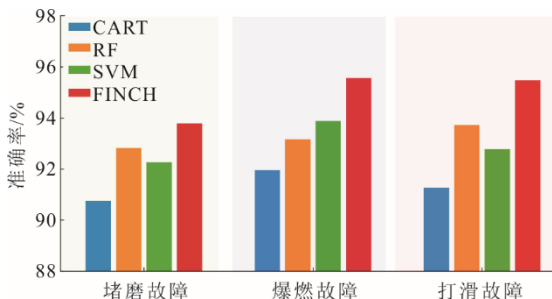


图 15 各分类器对每一故障的分类准确率  
Fig.15 Classification accuracy of different classifier for each fault

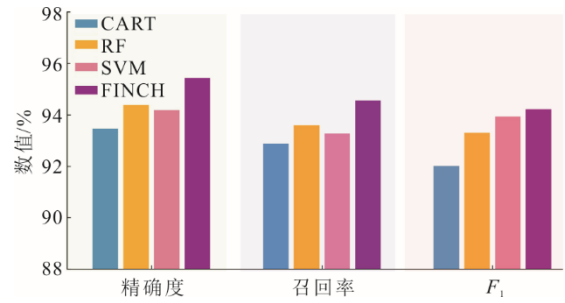


图 16 各分类器性能对比  
Fig.16 Comparison of the performance of each classifier

对比图 15、图 16 的结果可知, 无论是在召回率、精确度还是在  $F_1$  得分指标方面, FINCH 分类器的表现均优于决策树 CART、随机森林 RF、与支持向量机 SVM 的机器学习分类模型, 其在精确度上, 分别提高了 2.61 百分点、1.74 百分点、1.85 百分点, 从而验证了 FINCH 分类器在小样本故障分类上的有效性。

### 3 结 论

本文采用 PCA 方法通过  $T^2$  控制限与  $Q$  控制限对故障样本的进行甄别, 从表征设备状态的海量历史数据样本中提取小概率发生的故障样本, 采用的 FINCH 分类器针对故障小样本, 通过邻接矩阵与有向图的自然聚类, 实现了故障的高精度分类。总体来说, 本文提出的 PCA 与 FINCH 相结合的故障诊断方法为生产过程中出现的小概率故障事件, 如磨煤机故障等的精确诊断提供了一种有效的方法借鉴。

#### [参 考 文 献]

- [1] AGRAWAL V, PANIGRAHI B K, SUBBARAO P M V. Intelligent decision support system for detection and root cause analysis of faults in coal mills[J]. IEEE Transactions on Fuzzy Systems, 2017, 25(4): 934-944.
- [2] 樊帅, 肖军. 锅炉制粉系统故障诊断方法[J]. 热力发电, 2015, 44(2): 13-17.  
FAN Shuai, XIAO Jun. Fault diagnosis method of pulverizing system in utility boilers[J]. Thermal Power Generation, 2015, 44(2): 13-17.
- [3] 任梦玮, 焦嵩鸣. 改进型非线性状态估计的制粉系统故障诊断[J]. 热力发电, 2015, 44(12): 87-92.  
REN Mengyi, JIAO Songming. Coal milling system fault diagnosis based on improved NSET[J]. Thermal Power Generation, 2015, 44(12): 87-92.
- [4] 秦刚华, 雷丽君, 郭鼎, 等. 燃煤电站高压加热器系统故障仿真与诊断[J]. 热力发电, 2019, 48(6): 108-114.  
QIN Ganghua, LEI Lijun, GUO Ding, et al. Fault simulation and diagnosis for high pressure heater of coal-fired power plants[J]. Thermal Power Generation, 2019, 48(6): 108-114.
- [5] 杨淑红. 基于小波神经网络火电厂锅炉故障诊断研究[J]. 煤炭技术, 2011, 30(10): 46-48.

- YANG Shuhong. Research on fault diagnosis of power plant boiler based on WNN[J]. Coal technology, 2011, 30(10): 46-48.
- [6] 梁华. 基于模糊神经网络的凝汽器真空故障诊断方法[J]. 仪器仪表用户, 2015, 22(4): 72-76.  
LIANG Hua. Condenser vacuum fault diagnosis method based on fuzzy neural network[J]. Instrumentation, 2015, 22(4): 72-76.
- [7] 梁娜, 丁常富, 丁振宇. 基于概率神经网络的高压加热器故障诊断及仿真[J]. 发电设备, 2010, 24(2): 97-100.  
LIANG Na, DING Changfu, DING Zhenyu. Fault diagnosis and simulation for high-pressure heaters based on probabilistic neural network[J]. Power Equipment, 2010, 24(2): 97-100.
- [8] 费树岷, 李延红, 柴琳. 基于 RSPNN 的制粉系统故障诊断[J]. 控制工程, 2012, 19(3): 412-415.  
FEI Shumin, LI Yanhong, CHAI Lin. Fault diagnosis based on RSPNN for pulverizing systems[J]. Control Engineering of China, 2012, 19(3): 412-415.
- [9] 沈继忱, 刘玲, 李健. 火电厂凝汽器泄漏故障诊断研究[J]. 黑龙江电力, 2012, 34(5): 327-329.  
SHEN Jichen, LIU Ling, LI Jian. Study on the fault diagnosis of condenser leakage for thermal power plant[J]. Heilongjiang Electric Power, 2012, 34(5): 327-329.
- [10] 米江, 纪国宜. 基于改进 BP 神经网络的风机故障诊断研究[J]. 发电设备, 2010, 24(6): 430-434.  
MI Jiang, JI Guoyi. Fault diagnosis of fans based on improved BP neural network[J]. Power Equipment, 2010, 24(6): 430-434.
- [11] 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.  
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.
- [12] 刘一鸥. 基于支持向量机的设备故障诊断研究[J]. 电子设计工程, 2016, 24(22): 67-70.  
LIU Yiou. Research on equipment fault diagnosis based on support vector machine[J]. Electronic Design Engineering, 2016, 24(22): 67-70.
- [13] 翟嘉琪, 杨希祥, 程玉强, 等. 机器学习在故障检测与诊断领域应用综述[J]. 计算机测量与控制, 2021, 29(3): 1-9.  
ZHAI Jiaqi, YANG Xixiang, CHENG Yuqiang, et al. Overview of application of fault detection and diagnosis based on machine learning[J]. Computer Measurement & Control, 2021, 29(3): 1-9.
- [14] 田松峰, 胥佳瑞, 王美俊, 等. 基于 EEMD 云模型与 SVM 的汽轮机转子故障诊断方法[J]. 热力发电, 2017, 46(4): 111-114.  
TIAN Songfeng, XU Jiarui, WANG Meijun, et al. A rotor fault diagnosis method based on EEMD cloud model and SVM[J]. Thermal Power Generation, 2017, 46(4): 111-114.
- [15] 刘冲. 基于优化支持向量机的采煤机故障诊断技术[J]. 山东农业大学学报(自然科学版), 2015, 46(1): 132-135.  
LIU Chong. Breakdown diagnosis of coal excavator based on optimized support vector machine[J]. Journal of Shandong Agricultural University(Natural Science Edition), 2015, 46(1): 132-135.
- [16] 张凡, 孙文磊, 王宏伟, 等. 改进蝙蝠算法优化支持向量机的故障诊断方法[J]. 机械科学与技术, 2023, 42(3): 446-452.  
ZHANG Fan, SUN Wenlei, WANG Hongwei, et al. A fault diagnosis method with improved bat algorithm optimization support vector machine[J]. Mechanical Science and Technology for Aerospace Engineering, 2023, 42(3): 446-452.
- [17] 李锋, 汤宝平, 刘文艺. 遗传算法优化最小二乘支持向量机的故障诊断[J]. 重庆大学学报, 2010, 33(12): 14-20.  
LI Feng, TANG Baoping, LIU Yiping. Fault diagnosis based on least square support vector machine optimized by genetic algorithm[J]. Journal of Chongqing University, 2010, 33(12): 14-20.
- [18] 张烨, 黄伟. 基于天牛群算法优化 SVM 的磨煤机故障诊断[J]. 机电工程, 2022, 39(3): 411-418.  
ZHANG Ye, HUANG Wei. Fault diagnosis of coal mill based on PSO-SVM[J]. Journal of Mechanical & Electronic Engineering, 2022, 39(3): 411-418.
- [19] 孙栓柱, 江叶峰, 董顺, 等. 基于灰箱模型的中速磨煤机故障诊断方法[J]. 动力工程学报, 2018, 38(3): 211-220.  
SUN Shuanzhu, JIANG Yefeng, DONG Shun, et al. Fault diagnosis of a medium speed coal mill based on grey box mode[J]. Journal of Chinese Society of Power Engineering, 2018, 38(3): 211-220.
- [20] 张文涛, 马永光, 董子健, 等. 基于粒子群算法优化核极限学习机的磨煤机故障诊断[J]. 电力科学与工程, 2018, 34(9): 54-58.  
ZHANG Wentao, MA Yongguang, DONG Zijian, et al. Fault diagnosis of coal mill based on PSO-KELM[J]. Electric Power Science and Engineering, 2018, 34(9): 54-58.
- [21] QIAN H, SUN B, GUO Y, et al. A parallel deep learning algorithm with applications in process monitoring and fault prediction[J]. Computers and Electrical Engineering, 2022, 99: 107724.
- [22] GUO Y, LI K, LAVERTY D M, et al. Synchrophasor-based islanding detection for distributed generation systems using systematic principal component analysis approaches[J]. IEEE Transactions on Power Delivery, 2015, 30(6): 2544-2552.
- [23] AHMED A M H A A, FENG Z G, WASIM I, et al. Novel approach of principal component analysis method to assess the national energy performance via energy trilemma index[J]. Energy Reports, 2019, 5: 704-713.
- [24] SARFRAZ S, SHARMA V, STIEFELHAGEN R. Efficient parameter-free clustering using first neighbor relations[C]. IEEE/CVF: Conference on Computer Vision and Pattern Recognition(CVPR), 2019.
- [25] YAN T, HUANG X, ZHAO Q. Hierarchical superpixel segmentation by parallel CATrees labeling[J]. IEEE Transactions on Image Processing, 2022, 31: 4719-4732.
- [26] REDDY C K, VINZAMURI B. A survey of partitional and hierarchical clustering algorithms[M]. Data Clustering: Algorithms and Applications, 2018: 87-110.

(责任编辑 邓玲惠)