

基于小样本异常特征挖掘算法的抽水蓄能电站 主设备故障声纹检测技术

于潇[†], 李世昌, 卢彬

(河北张河湾蓄能发电有限责任公司, 河北 石家庄 050001)

摘 要: 由于抽水蓄能电站中部分设备难以接近, 导致采集有效信号的难度较大, 为了降低设备故障检测中的漏检率与误检率, 将小样本异常特征挖掘算法与声纹识别技术结合, 提出一种新的抽水蓄能电站主设备故障检测方法。选取压电式加速度传感器作为声纹采集传感器设备, 设定相关参数, 并通过 Pearson 相关系数确定传感器布设位置, 借助声纹数据采集模块采集运行状态数据。将 EMD 与小波变换结合对采集得到的主设备运行状态数据实施去噪, 降低故障检测误差。采用小样本异常特征挖掘算法针对去噪数据集进行异常数据挖掘, 将挖掘结果与声纹数据处理结合实现主设备故障检测。实验结果表明, 所提方法不仅降低了故障检测的漏检率与误检率, 而且能够克服干扰信号的影响, 实用价值较高。

关键词: 小样本异常特征挖掘算法; 声纹识别技术; 抽水蓄能电站; 故障检测; 声纹采集传感器

中图分类号: TV735

文献标识码: A

Voiceprint Detection Technology for Main Equipment Faults in Pumped Storage Power Plants Based on Small Sample Anomaly Feature Mining Algorithm

YU Xiao[†], LI Shichang, LU Bin

(Hebei Zhanghewan Energy Storage Power Generation Co., Ltd., Shijiazhuang, Hebei 050001, China)

Abstract: Due to the difficulty in accessing some equipment in pumped storage power plants, it is difficult to collect effective signals. In order to reduce the missed and false detection rates in equipment fault detection, a new fault detection method for main equipment in pumped storage power plants is proposed by combining small sample anomaly feature mining algorithm with voiceprint recognition technology. Select a piezoelectric accelerometer as the voiceprint acquisition sensor device, set relevant parameters, and determine the sensor placement position through Pearson correlation coefficient. Use the voiceprint data acquisition module to collect operating status data. Combining EMD with wavelet transform to denoise the collected operational status data of the main equipment and reduce fault detection errors. Using a small sample anomaly feature mining algorithm for anomaly data mining on denoised datasets, combining the mining results with voiceprint data processing to achieve main device fault detection. The experimental results show that the proposed method not only reduces the missed detection rate and false detection rate of fault detection, but also overcomes the influence of interference signals, with high practical value.

Key words: small sample anomaly feature mining algorithm; voiceprint recognition technology; pumped storage power plants; fault detection; voiceprint acquisition sensor

当今社会,电力需求不断增长,抽水蓄能电站作为一种储能能力较强、灵活性较高的电站类型,具有重要的战略意义^[1,2]。但是,由于抽水蓄能电站的机电设备大多集中在水下,检测故障非常困难,研究抽水蓄能电站主设备故障检测技术,对于提升电站设备的稳定性、降低维护成本以及保障电力安全有着至关重要的作用,因此该研究具有非常重要的现实和社会意义。

现阶段,常用的抽水蓄能电站主设备故障监测技术主要有:振动信号分析技术通过对水轮机、发电机等设备的振动信号进行采集和处理,分析其频谱和特征,诊断设备的工作状态和发现潜在的故障;声学信号分析技术通过采集设备发出的声音信号,分析声学特征,识别可能存在的故障;热红外成像技术通过对设备表面的热分布进行监测,识别设备中潜在的故障;除此之外,还有一些辅助诊断技术,如机械间谐波分析、隔离检测、模式识别等技术^[3,4]。虽然上述技术已经取得了一定的成功,但仍然存在不足之处:部分设备难以采集到有效的信号,对于水下设备或难以接近的设备,采集有效信号难度较大,降低了故障的检测效果;由于设备噪声、复杂的信号特征等原因,信号处理方法存在局限性,对精度和抗干扰性的要求较高;部分方法需要人工干预,例如数据处理、特征提取等,增加了人力成本和时间成本^[5]。

分析当前检测技术的研究现状可知,如何进一步提高抽水蓄能电站主设备故障检测的精度、降低故障检测误检率和漏检率,是未来研究的重点方向。以此为目标,将小样本异常特征挖掘算法与声纹识别技术结合,对抽水蓄能电站主设备故障进行检测研究。

1 主设备运行状态数据采集及其降噪处理

1.1 主设备运行状态数据采集

为了实现对抽水蓄能电站主设备故障的有效检测,利用声纹采集传感器对主设备运行状态数据进行采集,获取运行数据与主设备声音信息,提取设备细微特征,为故障检测提供基础。

1.1.1 布设位置确定

声纹监测是一种能够通过少量传声器感知整体振动状态的技术^[6,7],在进行声纹监测时,需要选择恰当的监测位置,准确地描述主设备振动的特征。为此,需要指定一个能够描述监测位置与主设备振动特征相关性的指标。斯皮尔曼和肯德尔相关系数是秩相关系数,通过对数据重新排序来计算相关性,因此不适合用于描述声纹频谱与振动特征

之间的相关性。而皮尔森(Pearson)相关系数^[8,9]在声纹监测中可以用于描述监测位置处声纹频谱与主设备各处振动特征之间的相关性。

根据上述分析结果,测得主设备表面上距离基座 1.5 m、位置靠中的区域 Pearson 相关系数值最优,因此选择该位置作为声纹采集传感器布点位置。图 1 为声纹采集传感器布点位置示意图。



图 1 声纹采集传感器布点位置示意图

1.1.2 声纹数据采集模块

声纹采集主要包含三大部分,分别是传感器、数据处理、后台系统,其中:传感器用于采集主设备运行时的原始信号;数据处理包括数据采集、存储和转发,用于原始声纹信号数字化转换,内部进行循环存储和对外转发;后台系统用于显示和配置装置参数。图 2 为声纹数据采集模块架构图。

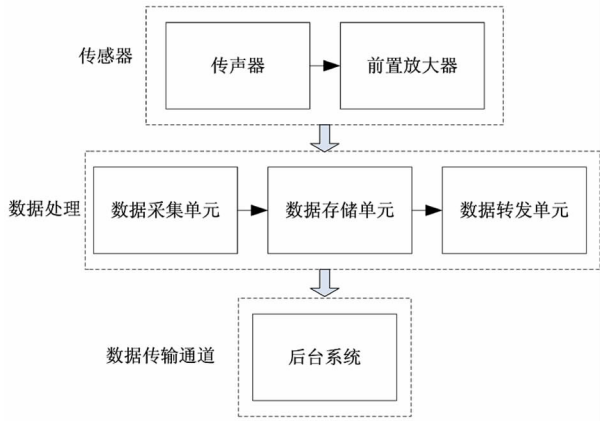


图 2 声纹数据采集模块架构图

(1)采集模块设计

声纹采集模块的设计对于声学传感器尤为重要,通过声音传感器将模拟信号经过放大电路和滤波电路,将稳定的声纹信号传输至 WM8960 芯片。该芯片集成了 D 类喇叭功放和 1 个喇叭,1 个声学传感器,3 路输入 2 路输出,是一款非常优秀的音

频处理芯片,能够实现更高质量的音频处理。通过 I2C 引脚将采集到的音频数据传输给主控制器,这为音频数据的处理和存储提供了便利。此外,WM8960 芯片的 DAC 信噪比为 98 dB,ADC 信噪比为 94 dB,完全符合音频采集性能要求。

(2) 传输模块设计

传输模块是声学传感器主体部分,主要功能包括了数字声纹信号的采集、处理、转发等,采用嵌入式 Linux 操作系统,极大地方便嵌入式软件的开发与应用,传输模块集成了 Cortex-A7 芯片,主频 800 MHz,512 M 内存空间,在性能上完全满足声学传感器对于声纹信号的处理要求。

1.2 主设备运行状态数据去噪

通过上文设置的声纹采集传感器采集得到大量的主设备运行状态数据,其中既包含对故障检测有帮助的有效数据,又包括部分含噪数据,含噪数据会对故障检测结果造成负面影响,因此,对其进行去噪处理至关重要。经验模态分解方法^[10](Empirical Mode Decomposition, EMD)能够有效地从原始信号中提取出包含不同尺度信息的本征模态函数(Intrinsic Mode Function, IMF),实现信号的降维和去噪。此外,小波变换则具有很好的去噪效果,能够通过去除高频噪声来保留原始信号中的重要信息。因此,将 EMD 与小波变换结合进行主设备运行状态数据处理可以有效地提高数据处理效果,得到更为准确和可靠的运行状态数据。

向原始信号中加入 m 组辅助白噪声可以有效地提高信号的随机性,进而提高数据处理的鲁棒性,并且可以在一定程度上避免信号在某些频率段上发生过拟合现象。加入白噪声后生成两个本征模态函数 IMF 集合:

$$\begin{bmatrix} S_1 \\ S_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} D \\ M \end{bmatrix} \quad (1)$$

式中: D 用于描述设备原始信号; M 用于描述加入的辅助噪声; S_1 和 S_2 用于描述含噪信号。

对式(1)集合中的信号进行分解得到:

$$F = \frac{1}{m} \sum_{i,j=1}^m (f_{ij})^M \quad (2)$$

式中: F 用于描述分解后的 IMF 分量; f_{ij} 用于描述第 i 个信号的第 j 个 IMF 分量。

基于信号分解结果,通过小波阈值去噪方法实现降噪。主要借助软硬阈值实现信号阈值处理,其中,硬阈值函数表达式为:

$$E_{k,l} = \begin{cases} e_{k,l} & |e_{k,l}| > \gamma \\ 0 & |e_{k,l}| \leq \gamma \end{cases} \quad (3)$$

软阈值函数的表达式为:

$$E_{k,l} = \begin{cases} \text{sgn}(e_{k,l})(|e_{k,l}| - \gamma) & |e_{k,l}| > \gamma \\ 0 & |e_{k,l}| \leq \gamma \end{cases} \quad (4)$$

式中: $\text{sgn}(\cdot)$ 表示符号函数; γ 表示阈值。

经过上述分析,将 EMD 分解方法和小波阈值去噪方法结合,用于实现主设备运行状态数据去噪,EMD 分解会得到含有较大噪声干扰的 IMF 分量,为了去除这些噪声干扰,可以采用小波阈值去噪方法进行信号降噪。去噪后的信号为:

$$D = \sum_{i=1}^m \sum_{j=1}^m D[f_{ij} \times E_{k,l}]^2 \quad (5)$$

通过将 EMD 分解与小波阈值去噪相结合的方式对主设备运行状态数据去噪,可以更好地抑制随机噪声,减少有效信号受到的干扰和损失。

2 抽水蓄能电站主设备故障声纹检测

上文中数据噪声已经通过 EMD 分解方法和小波阈值去噪方法得到了有效缓解,因此,接下来对去噪后的小样本进行异常特征挖掘,根据特征挖掘结果,展开故障检测。

2.1 小样本异常特征挖掘

小样本异常特征挖掘算法主要是针对数据集中样本数量较少的情况进行异常数据挖掘,传统的异常检测方法通常是基于正常数据建立模型,再使用该模型检测新数据是否异常^[11,12]。但是,当数据集中的异常点较少时,可能产生过拟合问题,因此,使用小样本异常特征挖掘算法对数据进行处理,检测出存在异常的数据样本。

由于抽水蓄能电站主设备实际运行状态下,每种工况条件下的噪声样本数据过少,不足以充分训练模型,而利用声纹采集传感器收集不同工况下的数据样本是非常费时、费力的任务。因此,使用参数迁移思想,对噪声数据特征进行训练和学习。图 3 为参数迁移训练示意图。

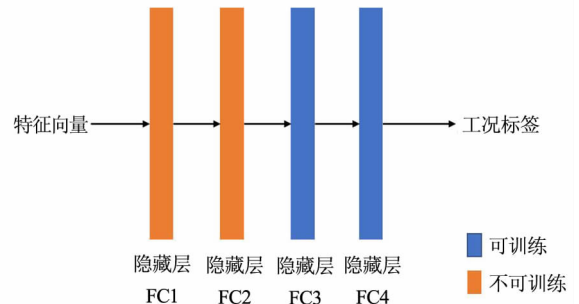


图 3 参数迁移训练示意图

由图 3 可知,参数迁移共有 4 个隐藏层,将训

练完成的模型保留第 1~2 层的参数,第 3~4 层用于训练。将隐藏层的最后一层输出即为小样本异常特征挖掘结果。

2.2 故障检测实现

在小样本异常特征挖掘的基础上,融合声纹采集结果和特征挖掘结果,建立抽水蓄能电站主设备故障检测模型。将实时采集的数据输入模型中进

行打分,根据分值判断当前状态和正常状态的偏离程度。设定一个阈值作为判别准则,打分值低于该阈值就被判别为异常状态,并发出报警^[13,14]。在早期可能存在一定比例的误报或漏报,由人工确认真实状态。将人工确认结果反馈输入至检测模型,模型通过参数调整更新进化输出最终检测结果^[15]。故障检测模型架构如图 4 所示。

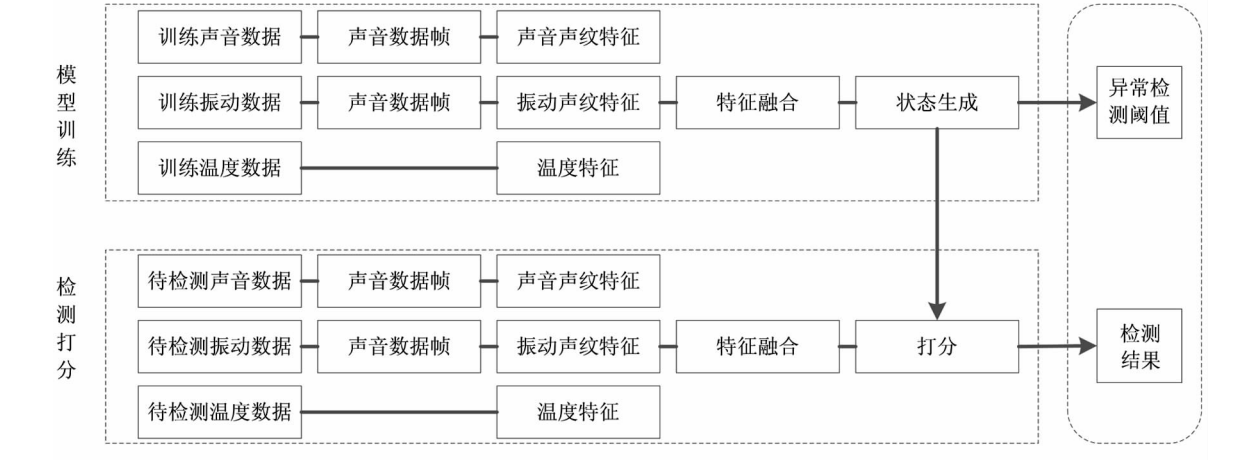


图 4 抽水蓄能电站主设备故障检测模型示意图

由图 4 可知,故障检测整个过程可分成 4 个阶段:

- (1)模型训练阶段利用声纹特征采集结果和小样本异常特征挖掘结果建立状态生成模型;
- (2)阈值确定阶段用训练得到的模型对数据进行打分,对这些打分值统计分析,确定异常检测阈值,通常阈值可选择为低于均值 3 倍或 6 倍的标准差;
- (3)检测打分阶段将从待测状态数据中提取的声纹特征输入模型中,得到检测打分值;
- (4)判别阶段将检测打分值和阈值比对给出异常检测结果,当检测分值低于异常检测阈值判定为异常状态,否则为正常状态,由此实现故障检测。

3 实验与结果分析

为了验证基于小样本异常特征挖掘算法的抽水蓄能电站主设备故障声纹检测技术的应用效果,展开实验测试分析。以某抽水蓄能电站为研究对象,该电站的主设备为一个带有滚动轴承的电机,其具体参数如表 1 所示。该电机在运行中可能会产生包括内环故障、外环故障和滚珠故障等在内的不同故障类型。使用声纹传感器采集设备振动信号,采样率为 20 kHz,采集时间为 5 min。为了保

证实验结果的准确性,将实验室温度保持在 20℃左右,湿度在 50%左右,确保实验环境稳定。

表 1 滚动轴承电机参数

参数	规格
电机型号	Y90s-4
功率/kW	1.5
轴承型号	6205
轴承内径/mm	20
额定载荷/kN	2.7
极限转速/(r·min ⁻¹)	28000

实验中,将振动信号分析技术和声学信号分析技术作为对比参照,与声纹检测技术进行对比分析,测试重点为故障误检率(False Positive Rate)、漏检率(False Negative Rate)以及故障信号分解效果,其中:

- (1)误检率是指将负类样本(没有故障)误判为正类样本(有故障)的比例。
- (2)漏检率是指将正类样本(有故障)误判为负类样本(没有故障)的比例。

不同技术的故障误检率和漏检率检测结果如图 5 和图 6 所示。

根据图 5 可知,在多次迭代中,振动信号分析技术、声学信号分析技术和所提技术的误检率均呈

现出逐渐提升的趋势,但是由于所提技术将 EMD 与小波变换方法结合进行主设备运行状态数据去噪,去除了干扰数据,避免其对故障检测结果造成负面影响,因此,所提技术的故障误检率更低,始终低于 3%。而振动信号分析技术和声学信号分析技术在故障信号采集中没有有效解决数据噪声,导致故障误检率较高。

根据图 6 可知,在多次迭代中,所提技术的故障漏检率均低于振动信号分析技术和声学信号分析技术,其漏检率始终低于 2%。由此可知,所提方法不仅能够有效提高设备故障检测结果准确性,而且提升检测结果的全面性,为设备运行安全、故障维修提供可靠的数据基础。

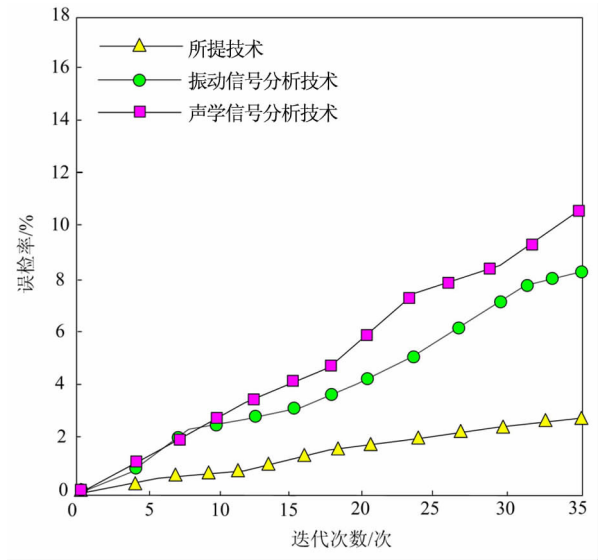


图 5 不同技术故障误检率对比结果

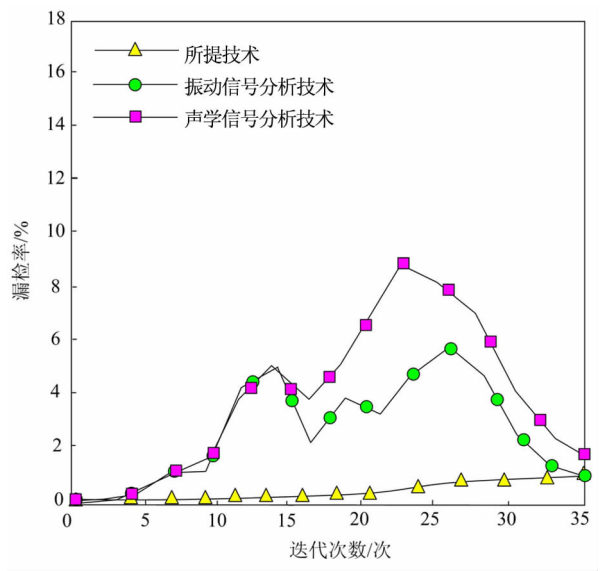


图 6 不同技术故障漏检率对比结果

如果抽水蓄能电站主设备在运行中发生故障,会产生故障信号,与正常信号混叠,这些信号会影响故障检测结果,因此,采用上述技术处理故障信号,判断不同技术能否获取准确的故障检测结果。结果如图 7 所示。

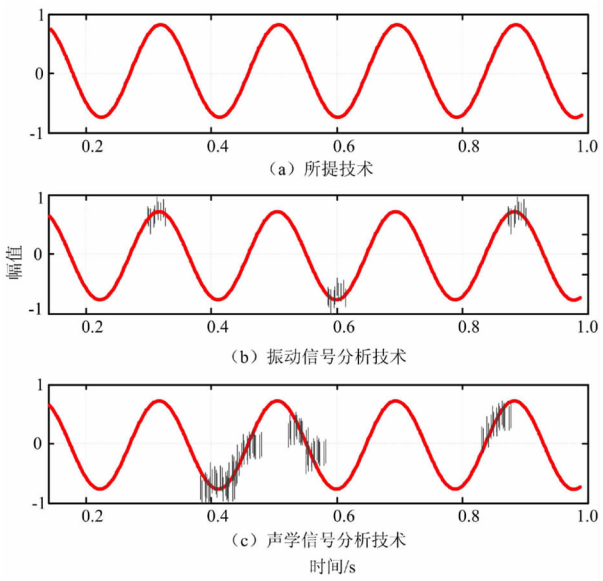


图 7 信号处理效果

由图 7 可知,采用振动信号分析技术和声学信号分析技术对信号进行处理,可以看到明显的模态混叠现象,而经过所提技术处理后,有效避免了这一现象。说明所提技术能够有效处理设备故障信号中的信号混叠问题,避免干扰信号的影响,提升故障检测效果。

5 结 论

抽水蓄能电站是一种能够存储和调度大量电力的重要设备,而主设备故障的发生会直接影响到电站的运行和稳定性。因此,对于抽水蓄能电站主设备故障的检测和预测具有重要意义。提出了一种基于小样本异常特征挖掘算法的抽水蓄能电站主设备故障声纹检测技术,其可以通过分析设备运行时所发出的声音信号,识别出潜在的故障特征。实验结果表明,与传统的故障检测技术相比,所提技术具有准确、全面等优点,可以为抽水蓄能电站的故障诊断提供有效的手段。但是该技术在检测中没有考虑设备关停期间的故障问题,接下来将重点研究该问题,以进一步提升所提技术的应用效果,为抽水蓄能电站主设备运行安全提供保障。