

基于ROC分析及样本熵的轴承故障检测

朱鸿斌, 徐维超*, 吴海平

广东工业大学自动化学院, 广东 广州

收稿日期: 2024年3月9日; 录用日期: 2024年4月3日; 发布日期: 2024年4月10日

摘要

轴承故障检测对旋转机械的维护至关重要。尽管已有的故障诊断方法取得了长足的进步, 但仍然面临着一些挑战, 如缺乏足够的故障数据用于训练, 对复杂的分布式数据的有效性不高, 对早期故障的敏感性低, 以及噪声和离群值的干扰。因此, 本文提出了一种基于样本熵和ROC分析的故障检测方法, 通过提取振动信号的样本熵指标, 然后使用ROC分析对其进行检测。实验结果表明, 本文所提方法能够以较高的准确率检测轴测故障。

关键词

故障检测, 样本熵, ROC分析

Bearing Fault Detection Based on ROC Analysis and Sample Entropy

Hongbin Zhu, Weichao Xu*, Haiping Wu

School of Automation, Guangdong University of Technology, Guangzhou Guangdong

Received: Mar. 9th, 2024; accepted: Apr. 3rd, 2024; published: Apr. 10th, 2024

Abstract

Bearing fault detection is crucial for the maintenance of rotating machinery. Although existing fault diagnosis methods have made significant progress, they still face some challenges, such as a lack of sufficient fault data for training, low effectiveness against complex distributed data, low sensitivity to early faults, and interference from noise and outliers. Therefore, this article proposes a fault detection method based on sample entropy and ROC analysis, by extracting the sample entropy index of vibration signals and then using ROC analysis to detect them. The experimental results show that the method proposed in this article can detect axial faults with high accuracy.

*通讯作者。

Keywords

Fault Detection, Sample Entropy, ROC Analysis

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

旋转机械设备在现代工业中有着广泛的应用,涉及石化、发电、冶金和化学工程等关键行业。滚动轴承是旋转机械中的一个重要且容易损坏的零件之一。因此,轴承故障诊断技术已成为一个突出的研究领域[1]。据统计,大多数旋转机械中约30%的故障源于轴承问题。并且轴承的运行状况直接影响轴承的性能机器的可靠性。这是因为轴承故障会在旋转机械设备中引发严重振动,导致设备损坏。因此,对滚动轴承进行高精度、及时的故障检测方法具有重要的意义和必要性。一般来说,滚动轴承发生故障时会产生异常振动和噪音,异常振动和噪音的信息通过信号幅度、频域等在时域中反映在振动信号中[2]。

基于振动信号的概率密度,学者们推导出量纲指标(均方根值、平均值等)和无量纲指标(波形指标、脉冲指标、裕度指标、峰度指标等)[3]。不幸的是,无量纲指标容易受到外界干扰信号的影响。此外,无量纲指标虽然不容易受到外界干扰,但仍存在严重的重叠问题。据此,Xiong等人提出了互量纲指标,并惊奇地发现互量纲指标1)对故障敏感;2)不受工况影响;3)能够缩短无量纲指标内部结构的距离,从而减少不同故障类别的无量纲指标的重叠[4]。基于互无量纲指标,朱等人利用AUC以及相关的方差的性质提出了一种有效的滚动轴承的故障检测方法,其中互无量纲和主成分分析用于特征提取,AUC以及相关的方差用于进一步的故障检测。实验结果表明该方法能够以较高的准确率和较低的误检率检测故障发生时间[5]。

值得注意的是,与正常轴承的振动信号相比,故障轴承的振动信号的复杂度是有差异的。信息熵是一种可用于评估信号复杂性的度量指标。利用信息熵,可以对振动信号进行量化,通过捕捉复杂的信息来促进特征提取[6]。在列车轴承故障诊断领域,样本熵和排列熵这两种信息熵形式已被用作特征提取的主要手段。在文献[7]中,有效地利用样本熵提取了隐藏在转向架振动信号中的故障特征;而在文献[8]中,利用排列熵从列车滚动轴承的振动信号中提取故障特征。结合轴承振动信号的差异性以及样本熵的特性,本文将使用样本熵对信号进行特征提取。

在检测阶段,我们利用两个滑动窗口沿样本熵来计算受试者工作特征曲线下的面积(AUC)。随后,这些度量被用作测试统计数据,以判断故障发生点。值得注意的是,由于Xu等人 and Waegeman等人分别基于动态规划和图形技术开发的快速算法,AUC的计算可以在线性时间内实现[6][7]。因此,在检测过程中,无需担忧本文算法的时间复杂性。接着,基于AUC零假设下近似正态分布的特性,根据置信区间设置相应的阈值,并认为超过阈值的AUC值是潜在故障发生时刻。最后,利用关键参数K,即K值方法,消除潜在故障点中的虚警点,进而实现高准确率故障检测。

2. 相关理论

2.1. 曼-惠特尼统计量

令独立同分布的样本序列 X_1, X_2, \dots, X_m 和 Y_1, Y_2, \dots, Y_n 分别服从于累积概率密度函数 F_X 、 F_Y 。从文献

[9]可知, 计算 AUC 的经验无偏估计量为:

$$\hat{\theta} \triangleq \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m (I(Y_i > X_j)) \quad (1)$$

其中

$$I(t) = \begin{cases} 1 & t > 0 \\ \frac{1}{2} & t = 0 \\ 0 & t < 0 \end{cases}$$

由上式可知, 当 X_1, X_2, \dots, X_m 与 Y_1, Y_2, \dots, Y_n 完全分离时, AUC 的估计值为 0 或者 1; 当 X 与 Y 完全重叠时, AUC 的估计值为 0.5。值得注意的是, 在零假设情况下, 当样本量 m 和 n 较大的情况时, 式中的 $\hat{\theta}$ 近似高斯分布[10], 即

$$\hat{\theta} \sim \mathcal{N}\left(\frac{1}{2}, \frac{m+n+1}{12mn}\right) \quad (2)$$

2.2. 样本熵理论

样本熵(SampEn)的物理意义是度量数据点序列中的不确定性和混沌性。时间序列的复杂性和不确定性可以通过计算在信号序列中产生新模式的概率来衡量。新模式出现的可能性越大, 序列的表现就越复杂和不确定。与近似熵相比, SampEn 具有两个主要优点: 1) SampEn 的计算不依赖于数据序列的长度; 2) SampEn 具有更好的一致性。一致性越高, 说明嵌入维数和相似容差的变化对样本熵测度的影响是相同的[6]。

假设待分析的信号为 $x = \{x_1, \dots, x_l\}$, 其样本熵的估计值, 可以描述如下[11]:

(a) 按信号序号组成一组维数为 m 的向量序列, 即 $X_m(1), \dots, X_m(l-m+1)$, 其中 $X_m(i) = x_i, \dots, x_{i+m-1}$;

(b) 计算(a)中 $X_m(i)$ 与 $X_m(j)$ 的距离, 并 $d[X_m(i), X_m(j)]$ 为 $X_m(i)$ 与 $X_m(j)$ 两者对应的元素中最大差值的绝对值, 即

$$d[X_m(i), X_m(j)] = \max(|x_{i+v} - x_{j+v}|)$$

其中 $v=1, \dots, m$ 。

(c) 统计(b)中 $d[X_m(i), X_m(j)]$ 小于相似性容限 r 的数目 B , 并进一步计算其占比, 即

$$B_i^m(r) = \frac{B_i}{N-m-1}$$

(d) 计算所有的平均值

$$B^m(r) = \frac{\sum_{i=1}^{N-m} B_i^m(r)}{N-m}$$

(e) 将维数增加至 $m+1$ 维, 并重复上述(a)~(d)步骤, 可得

$$B^{m+1}(r) = \frac{\sum_{i=1}^{N-m} B_i^{m+1}(r)}{N-m}$$

(f) 基于以上步骤, 信号 $x = \{x_1, \dots, x_l\}$ 的样本熵最终可以表示为:

$$SampEn = -\ln \frac{B^{m+1}(r)}{B^m(r)}$$

式中, l 为时间序列的长度, m 表示嵌入维数, r 为相似容差的大小。值得注意的是, 嵌入维度 m 的值理论上为 1 或 2。类似的容差 r 很大程度上取决于实际的应用环境, 通常设置为 $0.1 \times \text{std}$ 到 $0.25 \times \text{std}$, 其中 “std” 表示原始时间序列的标准偏差[6]。

3. 实验

3.1. 算法流程

(a) 首先, 在收集到的信号中提取健康的振动信号样本, 并将健康信号与未知信号合并, 组成待测信号序列;

(b) 提取原始振动信号的样本熵, 其中 m 表示嵌入维数, r 为相似容差的大小分别设为 500, $0.2 \times \text{std}$;

(c) 使用两个滑动窗口(均设置为 30)沿着(b)所获得的一维样本熵特征, 然后根据式(1)和计算相应的 AU。由此得到了一个新的一维 AUC 统计序列;

(d) 将 AUC 与相应的决策阈值进行比较, 估计出潜在的故障发生时刻。其中决策阈值的设置根据式(2)并将显著性水平 α 设为 0.05;

(e) 对连续势潜在的故障发生时刻进行统计, 并令连续势潜在的故障发生时刻的个数大于 K (本文中 $K = 20$)中的极大值为预测的故障发生时刻。

3.2. 数据来源

实验是基于 Case Western 进行的储备大学(CWRU)轴承数据中心[12], 其主要包括电机、扭矩传感器、测功机, 如图 1 所示。

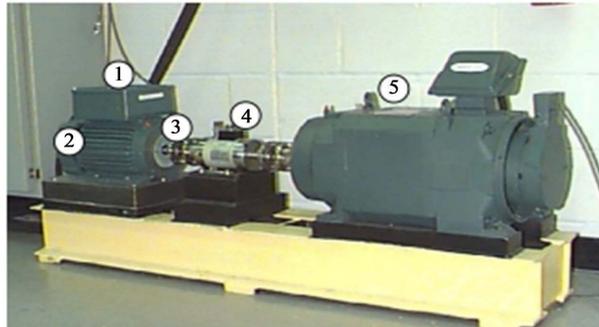


Figure 1. CWRU bearing test rig, which mainly consists of (1) Electric motor; (2) Fan end bearing; (3) Drive end bearing; (4) Torque transducer & encoder and (5) Dynamometer

图 1. CWRU 轴承试验台, 主要由(1) 电动机; (2) 风机端轴承; (3) 驱动端轴承; (4) 扭矩传感器和编码器; (5) 测功机

本文采用采样频率为 12,000 Hz 的加速度计采集原始振动数据(如健康、内圈故障和外圈故障)。为了进一步验证我们的方法的有效性, 采用故障直径为 0.18 mm (数据集中最小的故障直径)来测试我们的方法。

3.3. 性能评价指标

为了评估我们提出的方法的性能, 采用准确度指数(用 ACC 表示)作为优点值, 其定义如下:

$$S_{correct} = \sum_t I \left\{ \left| \hat{p}_t - p_t \right| < \frac{\text{Window}}{2} \right\}$$

$$\text{Accuracy} = \frac{S_{\text{correct}}}{N}$$

其中, p_i 和 \hat{p}_i 分别表示真实的故障发生时刻和预测的故障发生时刻。

3.4. 实验仿真及结果

本小节将验证本文所提方法的有效性和实用性。为了保证重复性和准确性, 本节的试验次数设置为 1000 次。此外, 为了进一步验证本文算法在微小故障(早期故障)方面的检测性能, 本实验所使用的故障为小的故障直径。

Table 1. The accuracy of fault detection

表 1. 故障检测准确率

Load	FaultType		
	Ball	Inner race	Outer race
0	1	0.998	1
1	1	1	1
2	1	1	1
3	1	1	0.999

表 1 列出本文算法在故障检测方面的准确性, 故障类型包括轴承滚珠磨损、外圈磨损和内圈磨损, 其中没有负载下的正常轴承的振动信号为参考信号。由表 1 不难看出, 本文所提算法在故障检测方面具有较好的性能, 对于滚珠故障、内圈磨损、外圈磨损等轴承故障准确率均在 99% 以上。

4. 结论

本文出了一种基于 AUC 分析以及样本熵的故障检测。通用利用振动信号在不同状态下的振动信号复杂度存在差异的特性, 提取原始振动信号的样本熵特征。然后使用 ROC 分析对其进行检测。实验结果表明, 本文提出的方法能够以较高的准确率实现故障检测。此外, 滑动窗口的大小, K 值设置以及样本熵方法的相关参数的设置都会影响本文所提算法的性能。因此, 未来的工作可进一步研究这些参数对算法检测性能的影响。

基金项目

本文研究工作由国家自然科学基金项目(62171141)赞助。

参考文献

- [1] Ye, R., Wang, W., Ren, Y., *et al.* (2020) Bearing Fault Detection Based on Convolutional Self-Attention Mechanism. 2020 *IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*, Weihai, 14-16 October 2020, 869-873. <https://doi.org/10.1109/ICCASIT50869.2020.9368683>
- [2] Li, X., Zhu, Y., Zhou, Z., *et al.* (2013) Fault Diagnosis Based on Nonlinear Dynamic Modeling in Rolling Element Bearing Systems. 2013 *IEEE International Symposium on Assembly and Manufacturing (ISAM)*, Xi'an, 30 July-2 August 2013, 12-15.
- [3] Zhang, Q.H., Shao, L., Li, H., *et al.* (2009) Research of Concurrent Faults Diagnosis Technology of Rotating Machinery Based on the Non-Dimensional Parameter. *Journal of Huazhong Univ. of Science and Technology (Natural Science Edition)*, **37**, 156-159.
- [4] Xiong, J., Zhang, Q., Wan, J., *et al.* (2017) Data Fusion Method Based on Mutual Dimensionless. *IEEE/ASME Trans-*

-
- actions on Mechatronics*, **23**, 506-517. <https://doi.org/10.1109/TMECH.2017.2759791>
- [5] Zhu, H., Xu, W., Delpha, C., *et al.* (2022) Mutual Dimensionless Indices and ROC Analysis in Bearing Fault Occurrence Detection. *IECON 2022-48th Annual Conference of the IEEE Industrial Electronics Society*, Brussels, 17-20 October 2022, 1-6. <https://doi.org/10.1109/IECON49645.2022.9968491>
- [6] 赵志宏, 杨绍普. 一种基于样本熵的轴承故障诊断方法[J]. 振动与冲击, 2012(6): 136-140, 154.
- [7] 秦娜, 金炜东, 黄进, 等. 基于 EEMD 样本熵的高速列车转向架故障特征提取[J]. 西南交通大学学报, 2014, 49(1): 27-32.
- [8] 冯波. 基于 SVD-PE 的高速列车滚动轴承故障诊断模型[J]. 组合机床与自动化加工技术, 2018(7): 108-110+114. <https://doi.org/10.13462/j.cnki.mmtamt.2018.07.028>
- [9] Xu, W., Liu, S., Sun, X., *et al.* (2016) A Fast Algorithm for Unbiased Estimation of Variance of AUC Based on Dynamic Programming. *IEEE Access*, **4**, 9553-9560. <https://doi.org/10.1109/ACCESS.2016.2628102>
- [10] Wang, Y., Huang, G., Yang, J., *et al.* (2020) Change Point Detection with Mean Shift Based on AUC from Symmetric Sliding Windows. *Symmetry*, **12**, 599. <https://doi.org/10.3390/sym12040599>
- [11] 但长林, 李三雁, 张彬. 基于样本熵和 SVM 的滚动轴承故障诊断方法研究[J]. 中国测试, 2020, 46(11): 37-42.
- [12] Loparo, K. (2012) Case Western Reserve University Bearing Data Center. Bearings Vibration Data Sets, Case Western Reserve University, Cleveland, 22-28.