

基于地铁列车监控数据的牵引变流器滤网堵塞故障预警研究

梁师嵩

(中车南京浦镇车辆有限公司, 210031, 南京//高级工程师)

摘要 针对地铁列车牵引变流器冷却系统的滤网脏堵问题,提出了一种基于机器学习的变流器滤网脏堵预警模型,以指导滤网的预测维修决策。首先,为确定强分类的特征参数,结合列车牵引系统的检修记录,利用随机森林分类模型对列车历史数据进行特征筛选;其次,基于选定特征构建了孤立森林异常检测模型,并对模型参数进行优化;最后,分别在历史健康数据、故障数据集上进行测试,以确定模型能够识别滤网脏堵的准确度。结果表明:所构建的孤立森林模型分数能够直接反映牵引变流器滤网的脏堵程度,可有效指导滤网脏堵的维修决策。

关键词 地铁列车;牵引变流器;滤网脏堵;故障预警;机器学习

中图分类号 U224.2⁺2

DOI:10.16037/j.1007-869x.2022.03.016

Research on Fault Early Warning of Traction Converter Filter Fouling Based on Metro Train Monitoring Data

LIANG Shisong

Abstract Aiming at the problem of filter dirt and fouling of train traction converter cooling system, a converter filter fouling early warning model is proposed to guide the predictive maintenance decision of the filter screen. First, in order to determine the characteristic parameters of strong classification, combined with the maintenance records of the train traction system, the random forest classification model is used to complete the feature screening of the train historical data. Second, the isolated forest anomaly detection model is constructed based on selected features, and the model parameters are optimized. Finally, tests are performed on historical health data and fault data sets to make sure that the model can accurately identify filter fouling. Results show that the scoring of the constructed isolated forest model can directly reflect the degree of traction converter filter fouling, effectively guiding the maintenance decision for filter fouling.

Key words metro train; traction converter; filter fouling;

fault early warning; machine learning

Author's address CRRC Nanjing Puzhen Co., Ltd., 210031, Nanjing, China

0 引言

地铁列车牵引系统的冷却装置故障关系到列车的运行安全。列车牵引系统的冷却装置风机滤网若发生脏堵,将对牵引系统的正常运行产生重要影响^[1]。列车牵引部件(如变流器、变压器等)的滤网若未及时清洗、除尘,容易降低列车运行可靠性,造成安全隐患^[2]。因此,如何建立准确、有效的滤网脏堵预警模型,用以指导滤网脏堵的预测维修,是亟需解决的问题。

为保证列车牵引变流器的工作可靠性,应确定合理的滤网维修方式与维修周期。目前,对于滤网的运用维修仍以计划预防修为主,并结合换件修、集中修,以确保滤网的维护质量。但经验表明,列车运行的外界环境变化对计划修的影响非常大,对一级修的影响尤为显著^[1]。因此,列车的运用维修应结合外部运行环境,适当缩短或延长各部件的维修周期,以避免因不及时维修造成的系统可靠性隐患或频繁维修增加的额外维修成本。为此,利用数据挖掘技术,在列车现有数据上构建系统预测模型,以推动由计划修到状态修、再到预测修的进程,是近年的发展热点;利用预测维修理论,能够有效降低维修成本、缩短维修周期,且能够更有针对性地进行维修处理^[3-4]。

根据故障清单的样本条件,可将列车牵引变流器滤网部件的故障预测方法分为两类:一是依靠异常检测的方法^[5-6]。该类方法完全基于健康状态历史数据进行预测模型的正常域标定,如文献[3]构建了油冷却器的油温、温升、温差等加权特征指标,

并通过无脏堵健康数据确定指标正常边界。与性能退化模型相比,基于阈值参数判别的检测方法的性能相对更依赖历史数据集的大小,故异常识别结果受阈值参数取值变化的影响较大。二是利用历史故障数据拟合性能指标的退化轨迹,该类方法需建立在足够的故障试验数据基础上。文献[1]基于不同脏堵程度的模拟数据标定风机电流值,采用设定多级电机电流控制门限的方式,以进一步减少故障误判;文献[7]利用多任务深度学习技术,拟合出变流器性能退化的曲线,实现脏堵程度的量化。

事实上,列车的外界环境多变(如受沙尘天气等影响),且滤网脏堵程度量化指标的定义依赖主观经验,缺乏确切的定义标准,因此,如何基于现阶段的列车牵引系统实时采集数据来实现滤网脏堵预警,仍需进一步探讨。

本文在不增加传感设备的前提下,利用列车历史数据,采用聚类方法进行滤网脏堵的异常检测,并对脏堵程度完成了量化。首先,根据变流器滤网的计划修清洗记录对列车采集的历史数据标签进行处理,并利用随机森林模型进行特征权重分析,以筛选出模型特征;其次,利用孤立森林模型构建牵引变流器滤网脏堵预警模型,并对模型的参数进行设定。本文所构建的孤立森林模型的模型分数能够直接反映变流器滤网的脏堵程度,可用于滤网脏堵的维修决策。

1 牵引变流器冷却系统原理分析

列车牵引变流器主要包括逆变整流模块、热交换器、风机等部件,其散热的原理是:冷却水在流经热交换器依靠风机换风降温后,再次回流至冷却水管路,完成循环降温。显然,相同的外界环境条件下,如果变流器滤网发生严重脏堵,必然导致定功率、或风机的有效换风面积减小,进而可能引发柜体温度上升、柜体内外温差增大等问题。

在建立预警模型时,一方面需要考虑外部环境变化、滤网清洗标准控制等差异,低信噪比数据并不利于复杂回归模型的拟合效果;另一方面,还要认识到故障数据标签不足的现状。本文利用冷却水进口和出口的温度差、压力差等特征向量,构建孤立森林模型,以确定滤网脏堵的正常域边界。

2 牵引变流器滤网脏堵预警模型的建立

2.1 样本分析与筛选

如表1所示,模型的输入数据至少应包括冷却风机状态、列车参考速度、环境温度、冷却水进出口温度、冷却水进出口压力等方面。

表1 滤网脏堵预警模型的输入数据类型
Tab.1 Type of input data of filter fouling early warning model

数据类型	具体描述
冷却风机状态	判别风机的工作状态
列车参考速度/(km/h)	根据各轴车速计算得到的速度值
环境温度/℃	外界环境温度
冷却水出口压力/Pa	冷却器出水口的水压
冷却水入口压力/Pa	冷却器入水口的水压
冷却水出口温度/℃	冷却器出水口的水温
冷却水入口温度/℃	冷却器入水口的水温

为降低外部环境、列车运行状态等因素对模型的干扰,本文将列车以大于50 km/h的速度运行时的数据作为历史数据集,并设定数据样本筛选的条件为:① 列车参考速度大于50 km/h;② 冷却风机为运行状态(即状态信号为“1”);③ 冷却器出水口、入水口的压力无明显异常(基于设定参考阈值);④ 冷却器出水口、入水口的温度无明显异常(基于设定参考阈值)。

2.2 特征构建与特征选择

将随机森林分类器的分类准确率作为可分性判据进行特征的重要性度量,这是机器学习中经典的特征评价策略^[8]。因此,为明确特征变量对于区分滤网脏堵与洁净状态的贡献度大小,应确定构建模型使用的输入特征。本文利用随机森林分类模型来量化特征变量的重要性。

2.2.1 数据准备

每次进行滤网清洗维护时,滤网在脏堵程度、清洗标准均不可避免地存在一定的差异性^[9],但滤网物理脏堵程度的主观量化值差异性水平并不会对分类的准确性产生根本性影响,因此,本文认为清洗前1 d采集的数据可以作为滤网脏堵的故障数据,清洗维修后次日采集的数据可以作为滤网的健康数据。根据该原则,本文随机选取了2次滤网清洗过程,记录清洗前1 d和清洗后1 d的数据,依次对数据进行筛选和标签处理。

2.2.2 特征生成

所构建的数据特征包括原始物理特征、再生物理特征的统计特征。其中:原始物理特征参数包括柜体温度、环境温度、进水口温度、出水口温度、进水口压力、出水口压力和列车参考速度;再生物理特征参数包括柜体温度和环境温度的差值、进水口温度和出水口温度的差值、进水口压力和出水口压力的差值。再生物理特征指基于原始物理特征形成的数学特征,包括最大值、中位值、平均值、差分值等。

2.2.3 特征筛选

上述数据经标签处理、特征提取后得到特征变

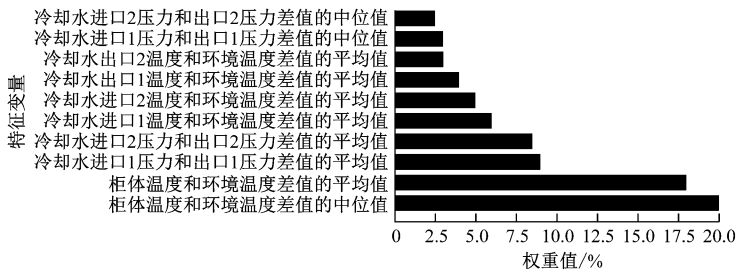


图1 滤网脏堵预警模型特征变量的权重分布

Fig.1 Weight distribution of characteristic variables of filter fouling early warning model

本文选择某列车的A、B两节车厢作为研究对象进行对比分析。如图2所示,以测试当日作为时间参照,A车在测试日的前1 d未进行滤网清洗,B车在测试日的前1 d进行了滤网清洗。从测试当日柜体温度和环境温度差值的变化曲线可以看出:A车的温度差值明显大于B车的温度差值。

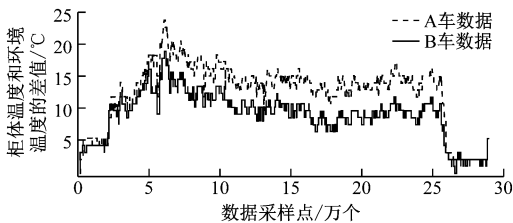


图2 滤网清洗(B车)和滤网未清洗(A车)的车辆柜体温度和环境温度差值的对比

Fig.2 Comparison of vehicle cabinet and environment temperature difference between conditions of cleaned filter(vehicle B) and uncleaned filter(vehicle A)

如图3所示,另选某个检测日,在该检测日前累计7 d内A、B车均未进行过清洗,此时A、B车在检测当日柜体温度和环境温度的差值变化曲线非常接近。这说明柜体温度和环境温度的差值是关联

量数据。本文针对双路冷却水构成的冷却系统(存在进口1、进口2和出口1、出口2),利用随机森林分类模型进行强分类能力特征变量筛选,得到特征变量的权重分布如图1所示。由图1可知:区分滤网脏堵与清洁的主要特征变量重要性按照从大到小排序,前6位分别为:柜体温度和环境温度差值的中位值、柜体温度和环境温度差值的平均值、冷却水进口1压力和出口1压力差值的平均值、冷却水进口2压力和出口2压力差值的平均值、冷却水进口1温度和环境温度差值的平均值、冷却水进口2温度和环境温度差值的平均值。因此,本文将选定这6个特征变量作为模型特征。

滤网清洗与否的直接、关键的变量,这与本文特征筛选得到的结论相符。

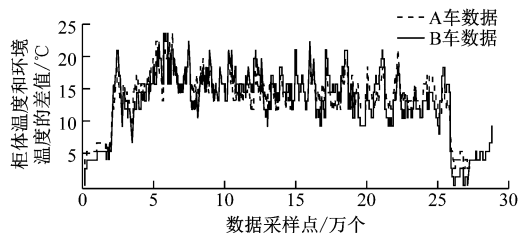


图3 累计7 d未进行滤网清洗的车辆柜体温度和环境温度差值的对比

Fig.3 Comparison of vehicle cabinet temperature and environment temperature difference after 7 days without filter cleaning

2.3 模型的建立与训练

由于故障标签的数据量不足,应结合所构建的特征建立孤立森林的异常检测模型,用以作为滤网脏堵预警模型。本文结合列车的实际维修清洗记录,默认仅有极少数日期发生过滤网严重脏堵的情况,即异常点占总样本量比例极小。根据该原则,选定2020年3月、4月列车白天的运行数据作为训练样本,在综合实际检修记录数据的基础上确定模

型孤立异常点的比例值。

滤网脏堵预警模型的整体输入为当前时刻的列车参考速度、柜体温度、环境温度、冷却水进出口温度、冷却水进出口压力等数据;模型的整体输出为当前时刻滤网脏堵的模型分数,其取值范围为 0~100,数值上等于异常样本数占总样本数的比例。因此,该分值越高,则表面滤网的脏堵情况越严重。图 4 为滤网脏堵预警模型的计算流程。

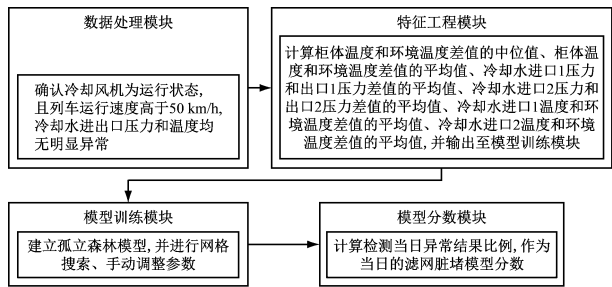


图 4 滤网脏堵预警模型的计算流程

Fig. 4 Calculation process of filter fouling early warning model

- 如图 4 所示,滤网脏堵预警模型的训练流程为:
- 1) 数据处理模块。筛选满足上文所述条件的列车数据,目的是通过数据清洗规则排除非运行状态的干扰噪声。
 - 2) 特征工程模块。针对处理后的数据,生成柜体温度和环境温度差值的中位值、柜体温度和环境温度差值的平均值、冷却水进口 1 压力和出口 1 压力差值的平均值、冷却水进口 2 压力和出口 2 压力差值的平均值、冷却水进口 1 温度和环境温度差值的平均值、冷却水进口 2 温度和环境温度差值的平均值。
 - 3) 模型训练模块。以上述 6 个特征变量作为孤立森林算法的输入,在训练集上进行网格搜索,并结合经验,对孤立森林模型超参数作进一步的优化。
 - 4) 模型分数模块。计算检测当日异常数据的

比例,并将异常样本比例作为滤网脏堵的预警模型分数结果输出,将结果进行记录,如表 2 所示。

表 2 训练数据集中清洗日期前后的模型分数对比
Tab. 2 Comparison model scores of training dataset before and after cleaning date

检测日期	模型分数
2020-03-17	95.86
2020-03-18	96.71
2020-03-19	61.66
2020-04-23	96.16
2020-04-24	94.65
2020-04-25	59.20

注:所有检测均在当日的白天进行。

根据检修记录,2020 年的 3 月 18 日和 4 月 24 日列车在完成载客任务回库后,在夜间进行了滤网清洗。表 2 的实际检修结果显示,滤网在 3 月 18 日、4 月 24 日白天处于脏堵状态。

表 2 数据表明,3 月 18 日(清洗前)的模型分数明显高于 3 月 19 日(清洗后)的模型分数,4 月 24 日(清洗前)的模型分数同样高于 4 月 25 日(清洗后)的模型分数。这说明所建立模型的输出分值能够真实地反映滤网的清洁状态。

2.4 模型预测结果分析

本文选定某 3 列车(T1、T2、T3)2020 年的 5 月 17 日至 5 月 22 日的历史数据,用以验证滤网脏堵预警模型的准确性。预警模型在测试数据上的模型分数如表 3 所示,其中:T1 的滤网清洗时间为 5 月 18 日夜间和 5 月 20 日夜间;T2 的滤网清洗时间为 5 月 17 日夜间和 5 月 19 日夜间;T3 的滤网清洗时间为 5 月 17 日夜间、5 月 19 日夜间和 5 月 21 日夜间。

由表 3 可以看出:

- 1) 以 T1 为例,滤网清洗当日(清洗前)所得的滤网脏堵预警模型分数显著高于次日(清洗后)的

表 3 滤网脏堵预警模型在测试集上的模型分数

列车	各检测日期下的模型分数					
	2020-05-17	2020-05-18	2020-05-19	2020-05-20	2020-05-21	2020-05-22
T1	77.70	78.09	71.17	68.71	61.43	65.56
T2	87.02	57.24	91.83	61.18	67.08	68.15
T3	83.08	69.36	86.60	76.23	57.31	79.20

注:所有检测均在当日的白天进行。

模型分数,这表明本文建立的孤立森林模型分数能够反映滤网脏堵程度的状态变化,即脏堵越严重,模型分数越高;

2) 以 T2、T3 为例,清洗前、后模型分数差值幅度差异的原因可能在于清洗前的滤网脏堵程度受当日环境因素影响,或清洗前滤网有效流通面积存在差距,但这些差异并不影响脏堵程度模型分数符合预期规律;

3) 对 T3 在 5 月 21 日滤网清洗维护后脏堵评分的情况进行原始物理特征观测,发现目标对象在滤网清洗后 1 d 内列车参考速度偏小,该原因或导致冷却牵引变流系统的散热效果变差。

综上所述,本文所建立的滤网脏堵预警模型,其在测试数据集上的模型分数变化能够表征滤网清洗的数据变化规律,可有效地反映滤网的脏堵程度。

3 结语

本文针对列车牵引变流器冷却系统的滤网脏堵问题,提出了基于机器学习的变流器滤网脏堵预警模型,用以指导滤网的预测维修决策。

1) 基于特征工程与特征权重分析手段,确定采用柜体温度和环境温度差值的中位值、柜体温度和环境温度差值的平均值、冷却水进口 1 和出口 1 压差的平均值、冷却水进口 2 和出口 2 压力差的平均值、冷却水进口 1 和出口 1 环境温差的平均值、冷却水进口 2 和出口 2 环境温差的平均值等 6 个变量作为判别滤网脏堵的强分类能力特征变量。采用这些特征变量进行预警分析,在提升模型效果的同时也降低了滤网脏堵异常预警模型的复杂度。

2) 本文建立了基于孤立森林的滤网脏堵预警模型,该模型能够在历史数据集上较好地反应当前列车牵引变流器的滤网脏堵状态,并能有效识别滤网清洗前后的故障状态。

3) 目前影响滤网脏堵预警模型所输出的脏堵程度量化指标准确性的因素包括训练集数据量的限制、外界环境变化的不可控性、计划清洗作业规范性的主观影响等。如何提高输出指标的准确性,是未来进一步研究的方向之一。

参考文献

[1] 尹鹏飞. 兰新铁路客运专线动车组运用维修研究[D]. 北京:

中国铁道科学研究院,2015.

YIN Pengfei. Research on the maintenance of electric multiple units for LanXin high-speed railway[D]. Beijing: China Academy of Railway Sciences, 2015.

[2] 杨彪,赵玮. 通过监控动车组散热风机电流变化判断滤网脏堵方法的深入研究[J]. 铁道车辆,2019(5):29.

YANG Biao, ZHAO Wei. Intensive research on judgement method of dirty or blocked filter net through monitoring change of current in radiating air fan on multiple units[J]. Rolling Stock, 2019(5):29.

[3] 王伟,顾佳,刘光俊. 动车组牵引变压器空气滤网预测性维修[J]. 集成电路应用,2019(3):72.

WANG Wei, GU Jia, LIU Guangjun. Predictive maintenance of air filter for traction transformer of EMU[J]. Applications of IC, 2019(3):72.

[4] 刘可安,戴计生,徐海龙. 城轨车辆关键部件状态修技术与系统解决方案[J]. 机车电传动,2020(4):5.

LIU Kean, DAI Jisheng, XU Hailong. Condition repair technology and system solutions for key components of urban rail vehicles[J]. Electric Drive for Locomotives, 2020(4):5.

[5] 周孝华,熊军,陈绍林,等. 基于电机负载响应的空调器过滤网脏堵检测[J]. 制冷与空调,2019(5):22.

ZHOU Xiaohua, XIONG Jun, CHEN Shaolin, et al. Dirty plugging detection of air conditioner's filter based on motor response to load changes[J]. Refrigeration and Air-Conditioning, 2019(5):22.

[6] PURARJOMANDLANGRUDI A, GHAPANCHI A H, ESMALI-FALAK M. A data mining approach for fault diagnosis: an application of anomaly detection algorithm[J]. Measurement, 2014, 55:343.

[7] 顾佳,黄明,关岳. 高速列车牵引变流器故障诊断研究[J]. 振动、测试与诊断,2020(5):997.

GU Jia, HUANG Ming, GUAN Yue. Fault diagnosis of traction converter for high-speed train[J]. Journal of Vibration, Measurement & Diagnosis, 2020(5):997.

[8] 姚登举,杨静,詹晓娟. 基于随机森林的特征选择算法[J]. 吉林大学学报(工学版),2014(1):142.

YAO Dengju, YANG Jing, ZHAN Xiaojuan. Feature selection algorithm based on random forest[J]. Journal of Jilin University (Engineering and Technology Edition), 2014(1):142.

[9] 张琦. 旅客列车空调系统蒸发器脏堵原因分析及对策[J]. 铁道运营技术,2016(3):1.

ZHNAG Qi. Analysis and countermeasures of evaporator fouling and blocking in air conditioning system of passenger train[J]. Railway Operation Technology, 2016(3):1.

(收稿日期:2021-05-29)