

文章编号:1671-6833(2019)05-0044-07

基于 Bi-RNN 的风电机组主轴承温度预警方法研究

尹诗^{1,2},侯国莲¹,于晓东¹,李宁¹,王其乐²,弓林娟¹

(1.华北电力大学 控制与计算机工程学院,北京 102206; 2.中能电力科技开发有限公司,北京 100034)

摘要: 主轴承是风电机组能量传递的关键设备,本文以双馈风力发电机组主轴承为研究对象,首先采用高斯混合模型(gaussian mixture model, GMM)对机组工况进行辨识;其次在各个子工况空间内建立基于双向循环神经网络(bi-directional recurrent neural network, Bi-RNN)的风电机组主轴承温度模型;然后,采用随机森林算法对主轴承温度模型残差进行建模与预测,从而实现机组主轴承故障预警;最后以某大型风电场机组为对象建模并开展仿真研究。结果表明,基于工况辨识的 Bi-RNN 神经网络算法结合随机森林算法对主轴承故障预警具有较强的实用性和较高的准确率。

关键词: 风电机组; 主轴承; 工况辨识; Bi-RNN; 随机森林

中图分类号: TK83 文献标志码: A doi:10.13705/j.issn.1671-6833.2019.05.008

0 引言

近年来,风能在世界能源结构中的地位越来越突出,风电将逐步成为火电、水电之后的第三大常规能源^[1]。由于风电机组运行环境恶劣,受气象、设备老化等多种不确定因素的影响,容易出现性能劣化与运行状态恶化问题,导致运维费用居高不下。统计资料表明,陆上和海上风电机组的运维费用分别占各自风电场收入的 10%~15% 和 20%~35%^[2]。在风电机组故障统计中,风电机组中发生电气和控制系统故障较为频繁,但是平均故障修复时间(mean time to repair, MTTR)较短,传动系统中的主轴承、齿轮箱、发电机等设备 MTTR 较长^[3-4]。因此,笔者通过风电机组主轴承温度预警来了解设备健康情况,这对合理安排检修具有重要的实践意义。

文献[5-8]分别采用频域分析法和时域分析法对轴承或齿轮箱进行故障诊断分析,但没有考虑故障劣化的渐变过程。文献[9]基于主成分分析法建立了正常运行状态下轴承的温度预测模型,并通过滑动窗口残差统计方法对故障进行有效预测,该方法中阈值的设定增加了人为参与度。文献[10]利用改进的集合经验模态分解方法构建轴承健康劣化度模型,但模型未考虑时序数据故障

劣化渐变过程对故障预警模型的影响。文献[11]利用非线性状态估计(nonlinear state estimate technology, NSET)方法建立齿轮箱正常工作状态下的温度模型并用其进行温度预测,采用滑动窗口方法计算残差的统计分布特性,当残差的均值或标准差的置信区间超出预先设定的阈值时,认为齿轮箱工作状态异常,该方法同样需要人为设定阈值判断故障情况。文献[12]通过采集风电机组主轴转动频率、滚动体频率和保持架频率等参数,利用深度置信网络模型对风电机组进行主轴承故障诊断,输入训练样本,逐层无监督训练达到局部参数最优,反向微调使整体性能最优,并用测试集数据进行验证,该方法未考虑风电机组运行过程中数据会随季节发生长周期变化,造成模型精度不高。

以上这些研究方法虽然对风电机组主轴承故障预警研究起到了促进作用,但仍存在研究盲点:

第一,算法模型多基于某段时间的采集与监视控制系统(supervisory control and data acquisition, SCADA)数据,但风电机组运行过程中数据通常会随季节发生长周期变化,这些变化有可能将故障早期形成过程中产生的数据变化淹没,无法在真正意义上实现故障预警。

第二,在建立时序数据模型过程中,通常只考

收稿日期:2018-12-27;修订日期:2019-02-21

基金项目:中央高校基本科研业务费专项资金资助项目(2019JG004);国家自然科学基金资助项目(61603136)

作者简介:尹诗(1988—),男,山东菏泽人,华北电力大学博士生,主要从事风电数据挖掘, E-mail:yinshi502@

虑模型输入输出的非线性关系,很少考虑时序数据故障劣化的渐变过程对故障预警模型的影响。

第三,算法多采用人为设定阈值来评价风电机组故障状态,阈值设定不合理会导致模型预测精度不高、泛化性差的问题。

因此,笔者将 GMM、Bi-RNN 神经网络算法、随机森林算法相结合,提出一种风电机组主轴承温度预警方法,使用 SCADA 数据和设备更换台账就能准确预警主轴承故障。

1 预警方法设计

笔者以风电机组主轴承为研究对象,将 GMM、Bi-RNN 神经网络算法、随机森林算法有机结合,开展主轴承温度预警研究。

考虑到风电机组运行的工况不同,利用 GMM 对机组工况进行辨识,克服了气候周期性变化及气象条件瞬时变化对预警结果的影响,扩大了算法的时间尺度和适用范围。

考虑到机组设备故障劣化的渐变过程对模型精度的影响,利用 Bi-RNN 神经网络算法建立主轴承温度模型,使之能够更加精准地预测以长时间序列数据为特征的故障劣化问题。

采用随机森林算法建立主轴承故障预警决策模型,降低人为设定阈值引起的复杂性和不确定性,使得该预警方法更加有利于工程实际应用,具体方法框图如图 1 所示。

首先,将多台风电机组 SCADA 数据预处理,从中选取有效数据,确定 Bi-RNN 神经网络算法的输入数据和输出数据;再采用 GMM 对机组运行工况进行辨识,根据机组运行建立 3 个子工况空间;然后在每个子工况空间内建立基于 Bi-RNN 神经网络的主轴承温度模型,并计算模型输出与 SCADA 数据之间的残差;最后将计算的残差按天提取残差特征向量,根据残差分布建立基于随机森林的主轴承故障预警决策模型。将测试机组数据同样按照以上步骤进行计算,根据随机森林算法输出结果判断风电机组主轴承异常情况。

2 算法原理及实现

2.1 基于 GMM 的风电机组工况辨识

GMM 是一种聚类算法,该算法利用高斯概率密度函数(正态分布曲线)精确量化事物,是将一个事物分解为若干基于高斯概率密度的函数模型。某一高斯模型的概率密度函数为:

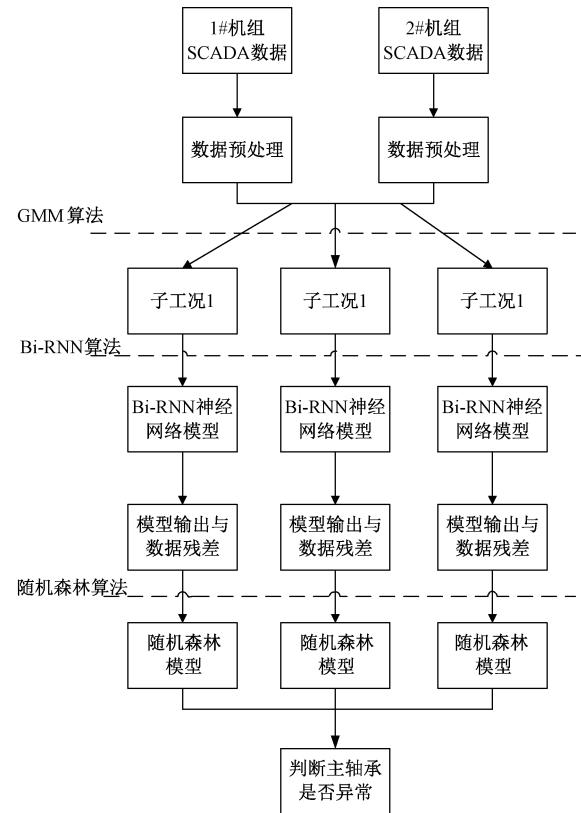


图 1 风电机组主轴承温度预警建模方法结构图
(2 台风电机组示例)

Fig.1 Structure diagram of temperature prediction model for main bearing of wind turbine
(Examples of 2 wind turbines)

$$N(x | \mu, \Sigma) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right], \quad (1)$$

其中, x 为 D 维向量; μ 为 D 维均值; Σ 为 $D \times D$ 维协方差矩阵。

GMM 的概率密度函数为:

$$p(x) = \sum_{k=1}^K \pi_k N(x | \mu_k, \Sigma_k). \quad (2)$$

式中: K 为高斯混合模型中高斯模型数量; $\sum_{k=1}^K \pi_k = 1$ 且 $0 \leq \pi_k \leq 1$ 。

应用 GMM 对风电机组运行工况进行辨识,建立 3 个子工况空间,将生成的每个高斯模型中的数据期望值作为多维特征点,代表隶属于本高斯模型全部有效数据点的数据特征。

2.2 基于 Bi-RNN 神经网络的风电机组主轴承温度模型

深度学习模型是一种拥有多个非线性映射层级的深度神经网络模型,能够对输入信号逐层抽象并提取特征,挖掘出更深层次的潜在规律^[13]。在众

多深度学习模型中,循环神经网络(recurrent neural network, RNN)将历史数据的时序特征引入到网络结构设计中,具有较强的适应性,如图2所示。

令第 t 时刻的输入为 $x_t \in R_{x\text{dim}}$,隐层节点的输出为 $h_t \in R_{h\text{dim}}$,输出层 $o_t \in R_{o\text{dim}}$,输入到隐层的权重矩阵 $U \in R_{x\text{dim} \times h\text{dim}}$,隐层自循环的权重矩阵 $W \in R_{h\text{dim} \times h\text{dim}}$,隐层到输出层的权重 $V \in R_{h\text{dim} \times o\text{dim}}$,对应偏执向量分别为 $b_h \in R_{h\text{dim}}, b_o \in R_{o\text{dim}}$,那么正向计算:

$$h_t = f(net_{ht}), \quad (3)$$

$$net_{ht} = x_t U + h_{t-1} W + b_h, \quad (4)$$

$$o_{th} = g(net_{ot}), \quad (5)$$

$$net_{ot} = h_t V + b_o, \quad (6)$$

其中, net_{st}, net_{ot} 分别表示激活前对应的加权和; $f(\cdot), g(\cdot)$ 表示激活函数。

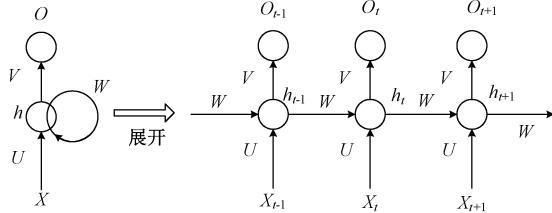


图2 RNN神经网络前向传播展开图

Fig.2 Expansion diagram of RNN neural network forward propagation

Bi-RNN神经网络如图3所示。输出由两个RNN的状态共同决定,算法既考虑时序数据的过去也考虑时序数据的将来,因此具有较强的时间特征。Bi-RNN神经网络在语言建模、视频行为分析、医学分析、图像识别等方面具有良好的应用^[14-18],在设备可靠性领域,劣化渐变过程对建立故障预警模型更为重要,因此,笔者将Bi-RNN神经网络算法引入到主轴承温度预警建模过程中。

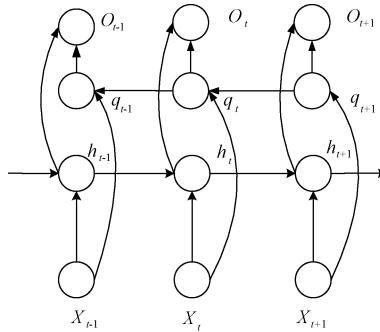


图3 Bi-RNN神经网络前向传播图

Fig.3 Diagram of Bi-RNN neural network forward propagation

Bi-RNN可以看作两个RNN的连接:一个是最前向传播RNN,另一个是反向传播RNN。反向传

播RNN的计算公式为(7)~(10)。Bi-RNN输出由两个RNN的状态求和共同决定。

$$q_t = f(net_{qt}), \quad (7)$$

$$net_{qt} = x_t U + q_{t+1} W + b_h, \quad (8)$$

$$o_{tq} = g(net_{ot}), \quad (9)$$

$$net_{ot} = q_t V + b_o. \quad (10)$$

根据公式(5)得出Bi-RNN的前向输出为两个RNN的状态和:

$$o_t = o_{tq} + o_{th}. \quad (11)$$

Bi-RNN神经网络输入参变量选择与主轴承温度密切相关的7个参数。

①风速:传动系统的转速与风速正相关,相应的主轴承温度会发生变化。

②发电机转速:与主轴承转速正相关。

③有功功率:与主轴承温度正相关。

④风向角:风向角影响风电机组机舱对风的准确性,影响机组有功功率。

⑤环境温度:由于风电机组所处的环境温度差异,即使机组功率和风速相同时,环境温度的不同也会导致主轴温度有很大差异。

⑥发电机绕组温度:主轴承转速越高,发电机转速越高,发电机绕组温度相对较高。

⑦齿轮箱油温:主轴承转速越高,齿轮箱传动系统转速越高,齿轮箱油温和主轴承温度相对越高。

输入和输出数据将做归一化处理,归一化计算公式为:

$$x_t = \frac{x_t - x_{\min}}{x_{\max} - x_{\min}}. \quad (12)$$

模型输入的参变量为数据的10分钟最大值、最小值、均值,输出数据选取10分钟均值。

模型参数选取时,设定输入层21个神经元,隐层设置4层记忆细胞,其中第一层100个神经元、第二层60个神经元、第三层35个神经元、第四层15个神经元,每一层的激活函数选择为线性整流函数(rectified linear unit, ReLU),输出层共1个神经元,激活函数为线性激活函数。

2.3 基于随机森林的主轴承故障预警决策

随机森林是一个包含多个决策树的分类器,并且输出的类别是由决策树的投票决定。随机森林算法框图如图4所示。将要分类的样本输入多个决策树,以少数服从多数的原则表决出这个样本的最终分类类型。

将风电机组主轴承温度模型输出与SCADA数据残差按天计算,得到残差向量的6个特征,分别是:模型残差向量最小值、模型残差向量最大

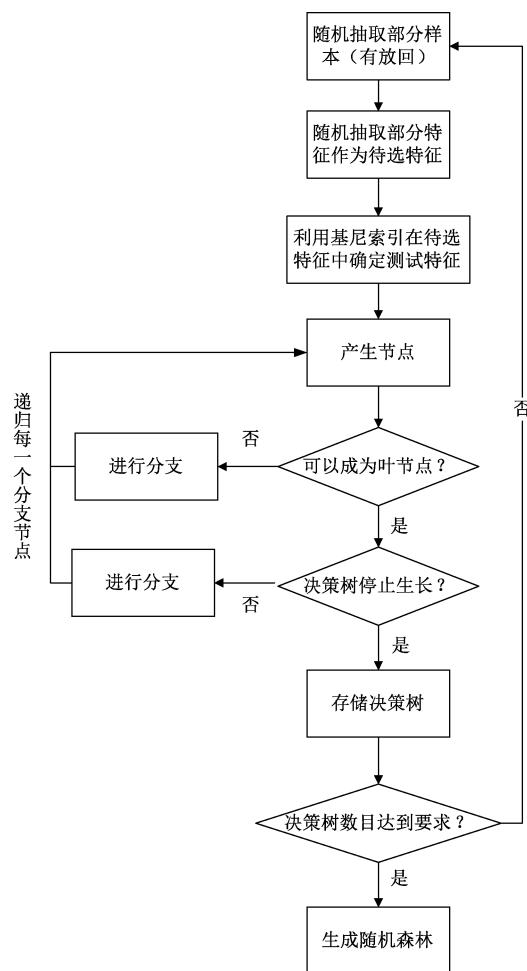


图 4 随机森林算法框图

Fig.4 Block diagram of random forest

值、模型残差向量平均值、模型残差向量中位数、模型残差向量标准差、模型残差向量欧氏距离.将该特征向量作为随机森林算法的输入,以是否为异常时间段标记值为随机森林算法的输出(1代表正常,0代表异常),样本设置为等权重,训练主轴承故障预警决策模型.

主轴承故障预警决策判定标准为:将故障警戒线设置为0.6,如果在一个月内随机森林算法输出有大量的数据点位于故障警戒线下且呈现出无反弹的特征,可以认为主轴承已发生磨损或已发生和主轴承温度相关的其他一些类似异常,反之判断主轴承正常.

3 实例分析及验证

笔者选用内蒙古某风电场33台1.5 MW双馈风力发电机组,切入风速为3.5 m/s,切出风速为25 m/s,在风速为13 m/s时进入恒功率阶段.选取25台主轴承正常运行的机组数据作为模型训练数据,每1台机组选取1个月数据,覆盖时间为全年,机组工况覆盖为切入风速到切

出风速的全工况.

为保证算法模型精度,将参与训练和测试的风电机组SCADA数据预处理,包括:剔除错位数据、噪声数据、停机数据,保留正常工况下的有效数据点,选取训练数据30 103条.

选取数据过程中,将风电机组升功率阶段中发电机恒转速阶段的SCADA数据去除,保留变速最大功率追踪阶段数据,其优点在于:①最大程度保留了风电机组动态参数丰富阶段的有效数据;②能避免因风力资源时间分布不均衡,尤其在风电机组长周期运行时,恒转速阶段风速差别较大的问题.

测试数据共选取8台风电机组数据,其中4台风电机组主轴承异常,4台风电机组主轴承正常.A002号机组和A005号机组由于主轴承磨损严重分别在2017年12月3号和12月15号更换;A007号机组由于主轴开裂于2017年1月13号更换主轴承;A018号机组由于主轴滚子表面剥落于2017年5月1日更换,选取每台机组更换主轴承前1个月的运行数据为测试数据.

应用GMM对风电机组运行工况划分为3个子工况:工况1是风速较低、有功功率相对较低的工况;工况3是风速较大、有功功率相对较高的工况;工况2介于两种工况之间.GMM建模结果如图5所示,每个工况的多维特征点如表1所示.

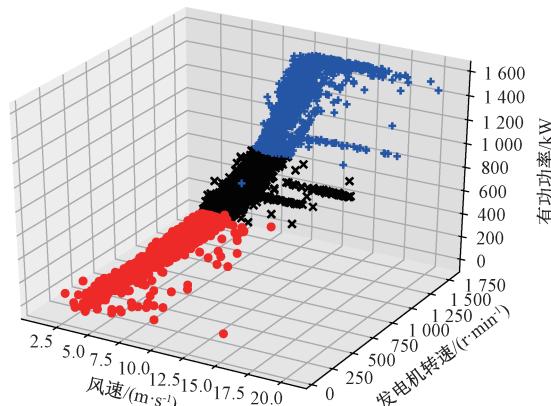


图 5 GMM 建模结果

Fig.5 GMM modeling results

表 1 某 GMM 生成的多维特征点

Tab.1 Multidimensional feature points generated

by a GMM

多维特征点工况	风速/(m·s⁻¹)	发电机转速/(r·min⁻¹)	有功功率/(kW)
工况1	4.53	1 060.4	134.9
工况2	6.95	1 443.0	458.3
工况3	9.74	1 723.0	1 135.7

在3个子工况下,分别使用RNN和Bi-RNN神经网络建立风电机组主轴承温度模型。RNN与Bi-RNN训练误差对比见表2,误差计算采用均方误差(mean-square error, MSE)。

在模型训练过程中,尽管RNN神经网络在前几次训练中误差和收敛速率高于Bi-RNN,但是很快Bi-RNN神经网络的收敛速率和训练误差均好于RNN,这表明在同等条件下,Bi-RNN神经网络算法训练精度优于RNN模型。

表2 3种工况下RNN和Bi-RNN训练误差
Tab.2 Training errors of RNN and Bi-RNN of three conditions

工况训练误差	RNN 迭代次数 (100次)	Bi-RNN 迭代次数 (100次)	RNN 迭代次数 (200次)	Bi-RNN 迭代次数 (200次)	RNN 迭代次数 (500次)	Bi-RNN 迭代次数 (500次)	RNN 迭代次数 (1 000次)	Bi-RNN 迭代次数 (1 000次)
工况1	0.339 9	0.378 7	0.034 6	0.021 5	0.026 0	0.016 0	0.021 7	0.013 1
工况2	0.277 7	0.357 5	0.041 8	0.023 8	0.027 2	0.017 6	0.021 8	0.013 0
工况3	0.217 1	0.291 1	0.041 7	0.023 7	0.025 5	0.015 9	0.021 7	0.012 9

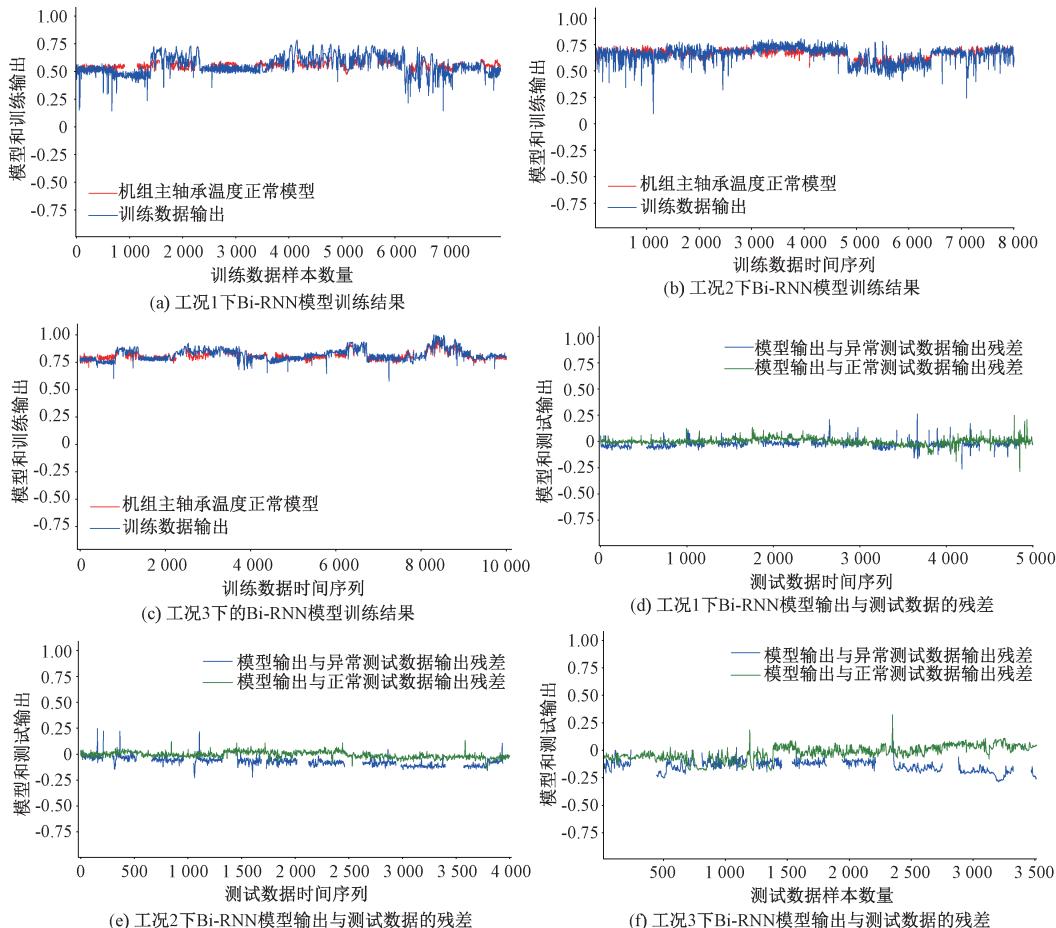


图6 基于Bi-RNN神经网络的风电机组主轴承温度模型训练和测试

Fig.6 Temperature model training and testing of wind turbine main bearing based on Bi-RNN

将测试机组SCADA数据输入风电机组主轴承温度模型并计算与SCADA数据之间的残差。按照2.3中提出的特征向量提取方法,将特征向量输入主轴承故障预警决策模型获取随机森林算法

模型训练和测试结果如图6所示。图中红色虚线为机组主轴承温度模型输出,蓝色实线为训练数据的输出。绿色曲线为机组主轴承正常测试样本数据与模型输出之间的残差,蓝色曲线为机组主轴承异常测试样本数据与模型输出之间的残差。从图6中可以看出,当主轴异常时,以主轴承温度为表征的模型输出和数据之间的残差增大,尤其是在工况3的时间序列后期这种残差更大。

的输出结果。测试的8台机组的故障预警情况如图7所示。

从图7(a)中可以看出,30#机组、32#机组、33#机组一个月每天的正常率都在警戒线以上,按

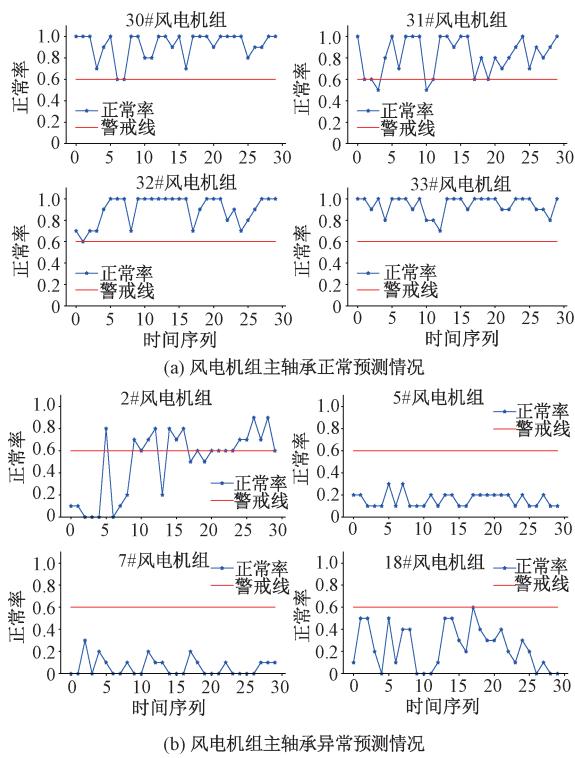


图 7 风电机组主轴承运行状态预测情况
Fig.7 Prediction of main bearing operation state of wind turbine

照主轴承故障预警决策判定标准,这 3 台机组判定为主轴承正常。31#机组有两个时刻机组主轴承的正常率在警戒线以下,但又快速反弹至警戒线以上,按照决策标准同样可判断该机组主轴承正常,后续可重点关注该台机组主轴状态。4 台主轴承正常的机组预测全部正确。

从图 7(b)中可以看出,5#机组、7#机组、18#机组一个月每天的正常率都在警戒线以下并且无反弹的趋势,根据主轴承故障预警决策判定标准,这 3 台机组判定为主轴承异常。2#机组在时间序列前期正常率在警戒线震荡,在后期正常率又上升至警戒线以上,后续应持续关注。4 台主轴承异常的机组中有 3 台机组预测正确。

4 结论

提出了一种风电机组主轴承温度预警方法,将 GMM、Bi-RNN 神经网络算法、随机森林算法有机结合预警潜在的故障。

相比于传统的故障预警方法,所用的算法对原始数据无特定要求,基本参数调校快。在方法设计中利用 GMM 对机组工况进行辨识,克服了气候周期性变化及气象条件瞬时变化对后续建模的影响,扩大了算法的时间尺度适用范围;同时,基于 Bi-RNN 神经网络的风电机组主轴承

温度模型,增加了故障劣化渐变过程的时序特征,使得故障预警模型更接近于实际状态;此外,基于随机森林的主轴承故障预警决策减低了人为设定阈值的复杂性和不确定性,提升了故障预测精度。文章最后通过实例验证了该方法的有效性与实用性。

参考文献:

- [1] 刘振亚.中国电力与能源[M].北京:中国电力出版社,2012.
- [2] TAVNER P. Offshore wind turbines: reliability, availability and maintenance [M]. London: The Institution of Engineering and Technology, 2012.
- [3] RIBRANT J, BERTLING L M. Survey of failures in wind power systems with focus on Swedish wind power plants during 1997 – 2005 [C]. IEEE Transactions Energy Conversion, 2007, 22(1):167–173.
- [4] 金晓航,孙毅,单继宏,等.风力发电机组故障诊断与预测技术研究综述[J].仪器仪表学报,2017,38(5):1041–1053.
- [5] 李浪.基于振动信号的风电机组轴承故障诊断研究[D].北京:华北电力大学电气与电子工程学院,2017.
- [6] 罗毅,甄立敬.基于小波包与倒频谱分析的风电机组齿轮箱齿轮裂纹诊断方法[J].振动与冲击,2015,34(3):210–214.
- [7] 龙泉,刘永前,杨勇平.基于粒子群优化 BP 神经网络的风电机组齿轮箱故障诊断方法[J].太阳能学报,2012,33(1):120–125.
- [8] 黄元维.基于支持向量机的风电机组主轴轴承故障诊断[J].仪器仪表用户,2016,23(11):88–92.
- [9] 丁佳煜,许昌,葛立超,等.基于轴承温度模型的风电机组故障预测研究[J].可再生能源,2018,36(2):276–282.
- [10] 董兴辉,张光,程友星,等.一种风电机组轴承健康劣化趋势预测方法[J].动力工程学报,2018,38(5):374–379.
- [11] 刘长良,闫萧.基于工况辨识的风电机组齿轮箱状态监测[J].热能动力工程,2016,31(7):41–46.
- [12] 王春梅.基于深度置信网络的风电机组主轴承故障诊断方法研究[J].自动化仪表,2018,39(5):33–37.
- [13] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015(521):436–444.
- [14] 喻晓东.基于 BRNN 和注意力机制的视频行人再识别方法研究[D].济南:山东大学控制科学与工程学院,2018.
- [15] 张克君,李伟男,钱榕,等.基于深度学习的文本自动摘要方案[J].计算机应用,2019,39(2):311–315.

- [16] 胡婕,张俊驰.双向循环网络中文分词模型[J].小型微型计算机系统,2017,38(3):522-526.
- [17] 杨文柱,刘晴,王思乐,等.基于深度卷积神经网络的羽绒图像识别[J].郑州大学学报(工学版),2018,39(2):11-17.
- [18] 毛晓波,张群,梁静,等.基于 PSO-RBF 神经网络的雾霾车牌识别算法研究[J].郑州大学学报(工学版),2017,38(4):46-50.

Research on Temperature Prediction Method for Main Bearing of Wind Turbine Based on Bi-RNN

YIN Shi^{1,2}, HOU Guolian¹, YU Xiaodong¹, LI Ning¹, WANG Qile², GONG Linjuan¹

(1.School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China; 2.ZhongNeng Power-Tech Development Co,LTD, Beijing 100034, China)

Abstract: The main bearing was the key equipment for energy transmission of wind turbines. In this paper, the main bearing of doubly fed wind turbines were examined. Firstly, the working conditions of wind turbines were identified by using Gauss mixture method. Secondly, the temperature model of main bearing of wind turbines based on Bi-directional Recurrent Neural Network (Bi-RNN) was established in each sub-working condition space. Thirdly, random forest was used to establish and predict residual temperature model of main bearing, so as to realize the fault warning of main bearing of wind turbines. Finally, algorithm model were used to carry out and simulate in a large wind farm. The results showed that the Bi-RNN neural network based on conditions identification which combined with random forest algorithm had strong practicability and high accuracy for main bearing early warning.

Key words: wind turbine; main bearing; condition identification; Bi-RNN; random forest

(上接第 37 页)

Study on the Influence of Battery Energy Storage Operation Control on Active Distribution Network

LI Xue, SONG Yanlong

(School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China)

Abstract: With the massive integration of distributed generation and electric vehicles, the problems of power and voltage quality in active distribution network were increasingly shown. Aiming at this issue, the factors affecting the interval control were analyzed in terms of energy storage capacity and power, daily load curve characteristics and unit time firstly. An improved interval control method for energy storage output model was then proposed to solve the problem of multiple charging/discharging operation in one cycle. Considering the correlation of random variables, probabilistic load flow using point estimate method was analyzed to state the influence of distributed generation, electric vehicles and energy storage station on voltage level. Finally, simulation analysis was operated on the improved IEEE-33 node active distribution network system with battery energy storage station. The results showed that the integration of energy storage station could effectively reduce the fluctuation of system power and voltage.

Key words: battery energy storage station; interval control; correlation; probabilistic load flow; point estimation method