

基于时空自编码网络的风电齿轮箱状态监测

刘长良¹, 徐 健², 王梓齐²

1. 新能源电力系统国家重点实验室(华北电力大学), 北京 昌平 102206;

2. 华北电力大学控制与计算机工程学院, 河北 保定 071000

基金项目: 北京市自然科学基金(4182061); 中央高校基本科研业务费(2020JG006, 2020MS117)

通信作者: 徐健, ncepu_zdsxj@163.com 收稿/录用/修回: 2022-02-28/2022-04-22/2022-06-08

摘要

在风电机组状态监测问题中, 常规自编码网络通常仅使用截面 SCADA(supervisory control and data acquisition)数据, 使得网络对数据时间特征的学习不足。因此, 提出一种基于时空自编码网络的风电齿轮箱状态监测方法: 使用 1 维卷积网络(1DCNN)级联双向长短时记忆网络(Bi-LSTM)作为编码层, 序贯提取面板数据的空间及时间特征, 以输入的重构误差作为预警指标实现在线状态监测。使用河北省某风电场实际数据验证, 结果表明: 相比故障记录时刻, 时空自编码网络能提前 20 d 发出报警信号, 且故障检出率和误报警次数均优于常规方法; 通过分析重构误差各分量的贡献率, 可知该齿轮箱故障中主要异常参数为油路压力和油池温度。

关键词

风电机组
状态监测
SCADA 数据
时空特征
自编码网络
中图法分类号: TM315
文献标识码: A

Condition Monitoring of Wind Turbine Gearbox Based on Spatial-temporal Autoencoder Network

LIU Changliang¹, XU Jian², WANG Ziqi²

1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China;

2. School of Control and Computer Engineering, North China Electric Power University, Baoding 071000, China

Abstract

Conventional autoencoder networks only use cross-section supervisory control and data acquisition data when monitoring wind turbine conditions, providing insufficient data to the network to learn about the temporal data characteristics. Therefore, a method of monitoring wind turbine gearbox conditions is proposed using a spatiotemporal autoencoder network. First, we use a one-dimensional convolutional neural network cascade bidirectional-long short-term memory network as the encoder layer to abstract the spatiotemporal characteristics of panel data sequentially. Second, input reconstruction errors are used as the warning index to realize online state monitoring. Finally, the results are verified using the actual data of a wind farm in Hebei province. The results demonstrate that, compared with the fault recording time, the spatiotemporal autoencoder network can send the alarm signals 20 days earlier, and the fault detection rate and false alarm times are better than the conventional methods. By analyzing the contribution rate of each component of the reconstruction error, it is observed that the main abnormal parameters of the gearbox fault are oil pressure and oil pool temperature.

Keywords

wind turbine;
condition monitoring;
SCADA (supervisory control and data acquisition) data;
spatial-temporal features;
autoencoder network

0 引言

在明确的碳达峰、碳中和目标之下, 风力发电将成为清洁能源增长的主力之一, 发展势头强劲^[1]。但风电机组通常安装在偏远的郊区或沿海地区, 受环境条件影响大^[2], 故障停机时有发生, 运行维护成本相对较高。齿轮箱是风电机组传动系统的关键设备, 平均停机维护时间最长, 维修成本也偏高^[3], 对其进行状态监测能够尽早发现潜在故障, 降低故障频率, 从而节约运维成本。

风电机组齿轮箱状态监测方法包括油液分析^[4]、声发射分析^[5]、振动信号分析^[6]以及使用数据采集和监督控制(supervisory control and data acquisition, SCADA)系统的离散数据进行分析等方法。因具备无需额外加装传感器、数据记录丰富、成本低廉等优势, 近期, 基于 SCADA 数据的风电机组状态监测研究颇多。部分学者通过建立回归模型的方法实现风电机组状态监测。王梓齐等^[7]运用正常行为建模方法计算出齿轮箱油池温度残差, 通过观察残差相对熵概率分布的差异, 对齿轮箱的运行状态进行实时监测; YANG 等^[8]基于多输入群体智能和支持矢量回归方法, 预测得到齿轮箱油温, 根据残差的趋势线, 实现齿轮箱状态监测。上述研究在状态监测问题中取得了一定效果。但是, 状态监测结果受选取目标变量的影响较大, 且难以综合描述设备运行状态。

自编码网络具有无监督学习的优势, 可同时计算多个变量的重构误差, 形成综合状态监测指标。近年来, 部分研究使用自编码网络进行风电机组状态监测: ZHAO 等^[9]通过将网络重构误差与自适应阈值进行比较, 完成对齿轮箱的状态监测任务, 但自编码网络的输入仅为截面 SCADA 数据, 未能融合数据的时间信息; 苏连成等^[10]将数据时间特征引入自编码网络, 有效提升了风电机组发电机和齿轮箱的状态监测精度, 但在使用面板 SCADA 数据时, 仅将多个时刻的数据首尾相连形成一列, 未考虑输入数据的时序性, 状态监测精度不高。

循环神经网络及其变体可以有效提取数据的时间特征, 在时间序列预测问题中得到了应用。黄婷婷^[11]等使用 LSTM 神经网络预测模型, 对股价指数数据进行预测; 黄荣舟^[12]等基于 LSTM 网络的时序数据预测模型, 对风电机组齿轮箱油温进行预测, 通过观察预测值与实际值残差的动态变化进行状态监测。

综上所述, 受网络结构限制, 常规自编码网络未能充分使用面板 SCADA 数据的时间特征, 而循环神经网络能够提取数据的时间特征。结合上述研究的优势, 对常规自编码网络进行了改进, 在编码层使用一维卷积网络级联双向长短时记忆网络, 序贯提取 SCADA 数据的空间及时间特征, 提升了状态监测效果; 通过观察重构误差各分量的贡献率, 可以实现故障分析。

1 时空自编码网络

本文提出的时空自编码(spatial-temporal autoencoder, STAE)网络是以自编码网络为整体框架, 使用 1 维卷积网络级联双向长短时记忆网络作为编码层的神经网络。下面对 STAE 网络所用的基础算法和框架结构进行介绍。

1.1 基础算法

自编码(autoencoder, AE)网络具有无监督学习的特性, 可以从无标注的数据中学习特征, 常用在目标识别^[13]、入侵检测^[14]、故障诊断^[15]等方面。AE 网络的基础结构如图 1 所示。

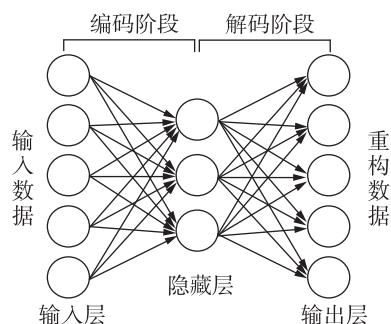


图 1 自编码网络结构图

Fig.1 AutoEncoder network structure

如图 1 所示, AE 网络一般由输入层、隐藏层及输出层三部分构成, 它以输入数据作为学习目标, 通过编码和解码两个阶段, 重构输入数据。在模型训练时, AE 通过梯度下降算法迭代运算, 使重构误差取得最小值。

AE 网络的重构误差是指网络输出层重构数据与实际 SCADA 数据之间的欧氏距离, 计算公式如下:

$$RE = \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \quad (1)$$

式中, $\mathbf{x}$ 为输入 SCADA 数据, $\hat{\mathbf{x}}$ 为输出重构数据。

卷积神经网络(convolutional neural network, CNN)内部通常包含多个卷积核, 每个卷积核由可学习的权重值和偏差组成^[16]。训练时, 初始权重值通过随机初始化给出, 偏差置 0, 均使用反向传

播算法自适应优化。

1 卷积神经网络(1-dimension convolutional neural network, 1DCNN)是一种形式简单的卷积神经网络,适用于1维数据的特征提取,运算原理如下:

$$\mathbf{y}_i^k = f_a(\mathbf{w}_k \otimes \mathbf{x}_i + b_k) \quad (2)$$

式中, $\mathbf{x}_i$ 为第 i 行数据向量, $\mathbf{w}_k$ 为第 k 个卷积核的权重向量,与 $\mathbf{x}_i$ 维度相同, $\otimes$ 为点积运算, b_k 为第 k 个卷积核的偏置, f_a 为激活函数, $\mathbf{y}_i^k$ 为第 k 个卷积核的第 i 个输出向量。

长短时记忆(long short term memory, LSTM)网络主要由输入门 i_t 、忘记门 f_t 、输出门 o_t 及记忆细胞 c_t 四部分组成,可以对时间序列中的关键信息进行传递与更新,充分使用数据的时间信息^[17]。

双向长短时记忆(bidirectional-long short term memory, Bi-LSTM)网络是 LSTM 网络的一种变体,包含前向 LSTM 和后向 LSTM 两部分,可以捕捉数据的双向联系,学习到更为丰富的时间特征^[18]。因此,在处理时序数据时, Bi-LSTM 网络往往能取得比 LSTM 网络更好的效果。Bi-LSTM 网络的结构如图 2 所示。

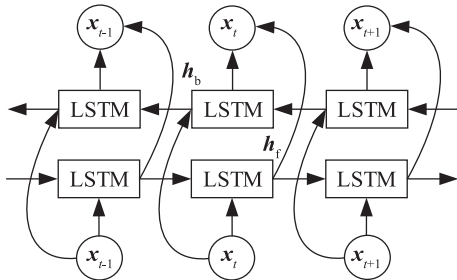


图 2 双向长短时记忆网络结构图

Fig.2 Bi-LSTM network structure

由图 2 可知, Bi-LSTM 网络的输出是由前向输出 h_f 和后向输出 h_b 组成,可以表示为

$$\mathbf{h}_f = \text{LSTM}(\mathbf{x}_t, i_{t-1}, o_{t-1}, f_{t-1}, c_{t-1}) \quad (3)$$

$$\mathbf{h}_b = \text{LSTM}(\mathbf{x}_t, i_{t+1}, o_{t+1}, f_{t+1}, c_{t+1}) \quad (4)$$

$$\mathbf{h}_t = [\mathbf{h}_f, \mathbf{h}_b] + b \quad (5)$$

式中, $\mathbf{x}_t$ 为 t 时刻 Bi-LSTM 网络的输入, $\mathbf{h}_t$ 为 t 时刻 Bi-LSTM 网络的输出, $[\]$ 为向量拼接符号, b 为 Bi-LSTM 网络的偏置。

1.2 提出的 STAE 网络

常规 AE 网络的编码层通常为全连接层,对数据特征的空间性、时间性均不敏感。结合 AE 网络无监督学习及数据重构的特性和 1DCNN、Bi-LSTM 网络在数据特征提取方面的优势,本文提出了 STAE 网络,框架结构如图 3 所示。其中,图 3(a)

为 STAE 与 AE 网络的对应关系,图 3(b)介绍了 STAE 网络特征提取过程。

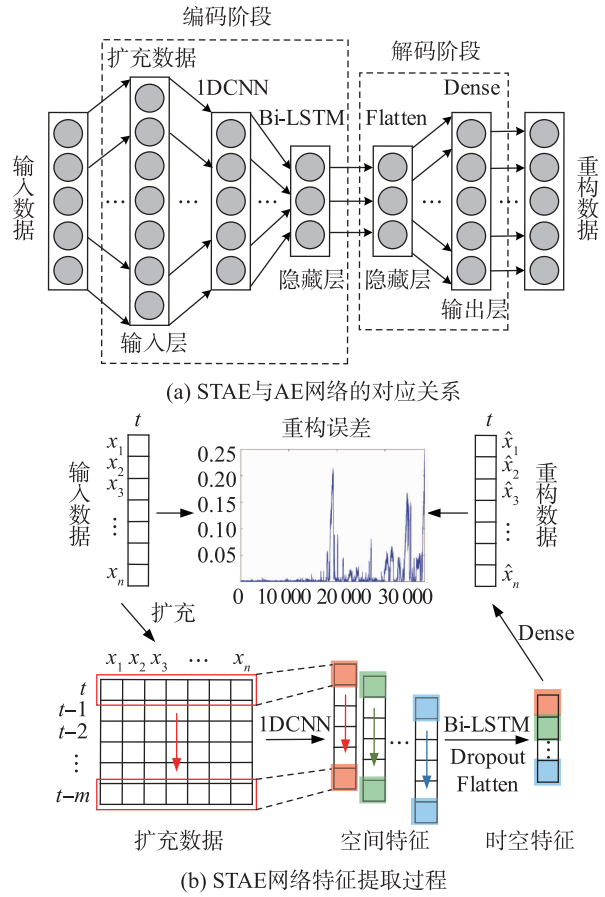


图 3 STAE 网络框架结构

Fig.3 The framework of STAE network

如图 3(a)所示,提出的 STAE 网络为非对称结构的自编码网络,使用 1DCNN 级联 Bi-LSTM 作为编码层;解码层为全连接 Dense 层,对提取特征进行解码重构。

图 3(b)为 STAE 网络特征提取过程和数据重构细节。数据输入后,首先进行数据扩充,以 t 时刻的实际数据作为主体,以 t 时刻前的一段历史数据作为辅助。

STAE 网络编码层中,使用 1DCNN 和 Bi-LSTM 序贯提取数据的空间及时间特征。其中,1DCNN 提取数据间的空间特征,将输入数据压缩为若干列时序数据;Bi-LSTM 提取数据中的时间特征。为提升网络的鲁棒性,避免出现过拟合现象, Bi-LSTM 后接 Dropout 层,随机去掉部分神经元。经过 Dropout 层及 Flatten 层对数据进行处理,形成包含单维 $1 \times M$ 个元素的时空特征数据。

STAE 网络解码层中,使用 Dense 层对时空特征进行解码,得到重构后的 t 时刻 SCADA 数据。

网络重构误差是指网络输出与 t 时刻实际数据之间的欧式距离。

STAE 网络通过综合考虑 SCADA 数据的时间与空间关系, 提升了常规自编码网络的特征提取能力, 可在无监督方式下对齿轮箱进行状态监测。

2 风电齿轮箱状态监测流程

本文通过正常行为建模方法^[19]完成对风电机组齿轮箱的实时状态监测。根据数据源的不同, 将状态监测分为离线阶段与在线阶段, 如图 4 所示。其中, ECDF(emperical cumulative density function)表示经验累积分布函数; a 为预警阈值; RE(reconstruction error)为重构误差。

本文采用的状态监测流程可分为 5 个步骤:

1) 数据预处理。首先, 将数据缺失段及异常采样点去除; 其次, 根据需要选择合适的变量作为研究对象; 随后, 将数据进行归一化处理。

2) STAE 模型训练。离线阶段, 将系统正常状态下 SCADA 历史数据经过数据预处理后, 送入搭建好的 STAE 网络模型进行训练, 并将模型的网络参数保存。

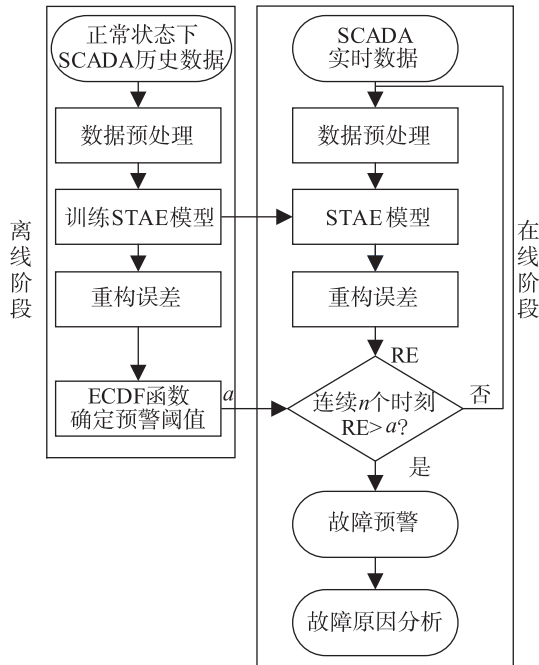


图4 状态监测流程图

Fig.4 Flow chart of condition monitoring

3) 确定预警阈值。累积分布频率曲线^[20]能够直观完整地描述重构误差的概率分布。离线阶段, 通过对齿轮箱正常状态下的重构误差统计分析, 选取累积分布频率 99.9% 时对应的重构误差作为预警

阈值。

4) 齿轮箱在线状态监测。在线阶段将实时 SCADA 数据经过预处理后, 送入离线阶段训练好的 STAE 网络。若连续 n 个时刻的重构误差大于设定阈值, 则表明风电机组齿轮箱存在潜在故障, 发出故障报警信号。

5) 故障原因分析。将故障信号对应的重构误差保存, 计算各分量的贡献率, 确定引起故障的主要参数; 观察各分量贡献率的变化情况, 对故障原因进行推理。

3 实例验证

本文实例为河北省某风场一台 1500 kW 额定功率的风电机组, 切入风速 3 m/s, 切出风速 25 m/s, SCADA 数据采样间隔 1 min。该机组于 2017 年 11 月 17 日 8:31 发生故障导致停运, 故障原因是齿轮箱油池温度高于上限值 70 ℃。

3.1 数据预处理

从 SCADA 系统数据库中导出 2017 年 8 月 16 日 0:00 至 11 月 17 日 8:31 (故障发生前 3 个月) 的风电机组运行数据。剔除数据中有功功率小于等于 0 kW、风速小于 3 m/s、风速大于 25 m/s 以及记录不完整的数据。

剔除记录异常数据后, 进行齿轮箱相关状态变量的选择。风电机组齿轮箱的运行状态容易受到多种因素影响。首先, 根据测点位置选取 5 个齿轮箱直接相关的变量, 包括: 齿轮箱滤网前油压、齿轮箱入口油压、齿轮箱驱动端轴承温度、齿轮箱非驱动端轴承温度、齿轮箱油池温度。随后, 依据皮尔逊相关系数, 在其余 60 个变量中选出 4 个与齿轮箱间接相关的变量, 包括: 发电机转速、主轴转速、实际扭矩、有功功率。

但是, 发电机转速和主轴转速、实际扭矩和有功功率相关系数接近 1, 变量间存在显著线性关系。为避免数据冗余, 将主轴转速与实际扭矩去掉。最终, 确定 7 个变量用于齿轮箱状态监测研究, 单位及取值范围如表 1 所示。

由表 1 可知, 各变量的取值范围差异较大, 为提升网络训练速度, 使用 Min-Max 归一化方法处理原始数据, 处理过程可表示为

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (6)$$

式中, x_i 为 x 的第 i 个原始数据, $x_{\max}$ 、 $x_{\min}$ 为 x 中的最大值和最小值, x'_i 为 Min-Max 归一化后的数据。

归一化后的数据样本整体情况如图 5 所示。

表 1 齿轮箱相关状态变量描述
Tab.1 Description of gearbox related state variables

标号	变量名称	取值范围
V1	齿轮箱入口油压 /Bar	[2.56, 3.43]
V2	齿轮箱滤网前油压 /Bar	[3.89, 6.03]
V3	齿轮箱驱动端轴承温度 /℃	[48.92, 78.8]
V4	齿轮箱非驱动端轴承温度 /℃	[48.28, 74.09]
V5	发电机转速 /(r/min)	[1 065.7, 1 979.1]
V6	有功功率 /kW	[0.03, 1 532.3]
V7	齿轮箱油池温度 /℃	[50.01, 70.23]

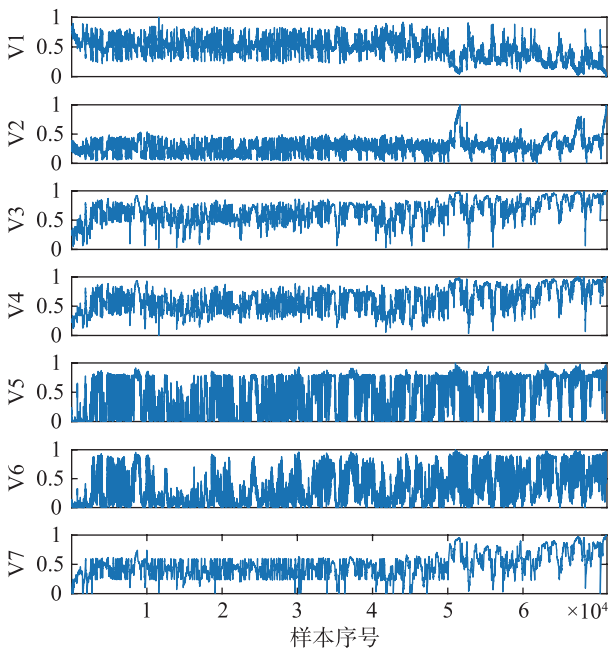


图 5 数据样本整体情况

Fig.5 The overall picture of the data sample

图 5 所示为 STAE 网络的输入数据，为便于观察齿轮箱从正常到故障的变化趋势，图中未画出记录异常数据。在进行归一化后，需要将数据切分生成数据样本，每个样本为包含连续 6 个时刻 7 个状态变量的面板 SCADA 数据。其中，最后一个时刻的数据为输入数据，其余时刻的数据为扩充数据。经过上述处理后，共得到 68 250 组有效数据，具体情况如表 2 所示。

表 2 数据划分表
Tab.2 Data partition table

数据描述	数据量 /组	数据范围	作用
训练集	21 000	1 ~ 21 000	模型训练
验证集	9 000	21 001 ~ 30 000	参数调整
测试集	38 250	30 001 ~ 68 250	模型测试

如表 2 所示，数据被划分为 3 个部分。训练集和验证集是正常状态下的历史数据；测试集数据用于模拟设备在线状态，验证提出方法的状态监测能力。

3.2 时空自编码网络方法验证

基于 TensorFlow 深度学习框架，使用 Keras 神经网络 API，搭建本文提出的 STAE 网络模型。根据训练样本的数据量，将 1DCNN 卷积核的大小设置为 1×7 ，数量设置为 12，使用 ReLU 激活函数；Bi-LSTM 的输入层时间步长设置为 6，输出设置为 6，使用 Sigmoid 激活函数。Bi-LSTM 后接 Dropout 层，防止模型过拟合。网络参数初始化方式为 TensorFlow 中默认的 Glorot uniform 参数随机初始化方法，随机种子设置为 42。

模型在训练时，学习率设置为 0.003，epochs 设置为 10，损失函数为平均绝对误差 MAE，采用 Adam Optimizer 更新网络权重。

离线阶段，为避免神经网络参数的随机性对训练结果造成影响，对 STAE 网络模型进行了 20 次重复训练，选用训练误差为中位数的模型参数。使用训练好的模型，生成训练样本重构误差，累积分布频率曲线如图 6 所示。

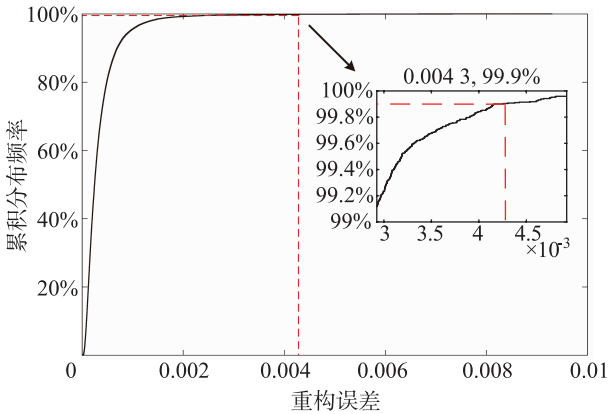


图 6 离线阶段重构误差累积分布频率曲线

Fig.6 Reconstructed error cumulative distribution frequency curve in off-line stage

如图 6 所示，累积分布频率为 99.9% 时，对应的重构误差为 0.004 3，设置其 STAE 网络预警阈值。在线阶段，将测试集数据送入离线阶段训练好的模型，得到测试集的重构误差。为增强报警信号的可信度，避免发生误报警现象，设定报警规则：当连续 5 个样本的重构误差大于预警阈值时，发出报警信号。图 7 为 STAE 网络在测试集上的状态监测结果，上半部分为测试集重构误差，下半部分为对应的报警信号。

从图 7 可知, STAE 网络在测试集上的首次报警序号为 17885, 对应的报警时间为 2017 年 10 月 28 日 6:06。相比故障记录时刻, 提前 20 d 轮箱运行状态异常。此后, 重构误差整体趋势先降低再升高, 报警信号先稀疏后密集, 表明风电机组的运行工况变化剧烈, 这也是状态监测的难点之一。

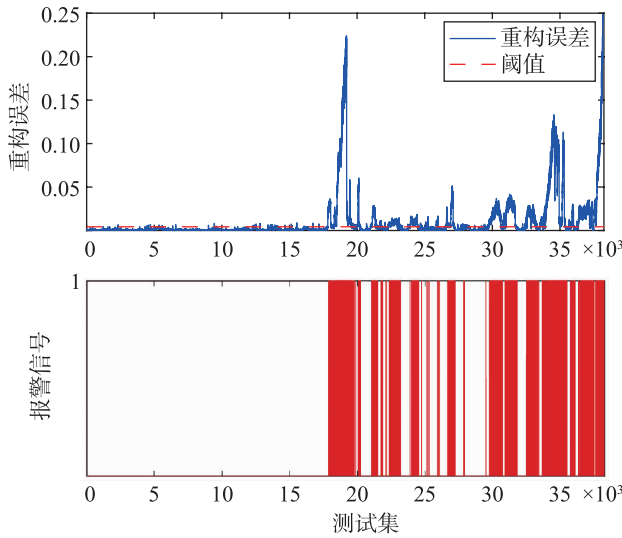


图 7 STAE 网络在测试集的重构误差及报警信号

Fig.7 Reconstruction error and alarm signal of STAE network in test set

此外, 测试集数据在 30 000 组之前报警不连续, 30 000 组以后有连续长时间的报警。这是因为齿轮箱从故障产生至停机, 其健康状态的恶化是一个渐变过程。对早期故障进行预警, 可以避免齿轮箱造成更严重的经济损失乃至安全事故。

3.3 对比实验

为验证提出 STAE 网络的状态监测性能, 分别采用常规自编码、时间自编码 (spatial autoencoder, SAE)、空间自编码 (temporal autoencoder, TAE)、先提取时间特征再提取空间特征的变体时空自编码 (variant spatial-temporal autoencoder, VSTAE) 四种网络重复上述实验, 并将实验结果进行定性和定量对比。

对比方法的网络结构和参数设置情况如下: 对照文[10], 把面板 SCADA 数据按时间顺序首尾相连形成一列, 将 AE 设置为 4 层网络, 各层神经元的个数分别为 42-30-15-7; SAE 网络在编码层仅使用 1DCNN 提取数据空间特征, 不使用 Bi-LSTM 网络; TAE 网络在编码层仅使用 Bi-LSTM 提取数据时间特征, 不使用 1DCNN 网络; VSTAE 网络在编码层先通过 Bi-LSTM 网络提取数据时间特征, 再通过

1DCNN 网络提取空间特征。为便于对比分析, 上述方法预警阈值的选取方法、故障报警方式及模型训练超参数均与 STAE 网络相同。

此外, 为避免网络参数的随机性对训练结果造成影响, 经离线阶段对模型进行训练后, 选用 20 次训练误差的中位数对应的模型参数。图 8 为 AE (autoencoder)、SAE (spatial autoencoder)、TAE (temporal autoencoder)、VSTAE (variant spatial-temporal autoencoder) 网络在测试集上的重构误差及报警信号。

如图 8 所示, 结合上述几种方法的报警情况, 认为齿轮箱故障发生在测试集样本 17 800 组之后。通过与提出 STAE 网络的预警结果对比可知:

- 1) AE 网络的预警阈值和重构误差整体偏高, 表明该方法对样本的学习能力弱, 未能充分提取面板数据中蕴藏的设备状态信息。
- 2) SAE 网络存在误报警现象, 且报警信号的连续性差, 仅在 20 000 组和 35 000 组数据处有连续报警, 表明该方法对故障状态不敏感。
- 3) TAE 网络的预警结果与提出方法较为接近, 但存在误报警现象, 降低了报警信号的可信度, 且在故障发生前期, 报警信号稀疏。
- 4) VSTAE 网络无误报警情况出现, 但报警信号过于稀疏, 仅能监测到设备显著异常, 说明该方法对设备正常状态样本学习不充分, 对异常状态不敏感。

此外, STAE 与 VSTAE 网络都对样本的时空特征进行提取, 但 VSTAE 网络的状态监测效果不佳。推测原因如下: 样本数据中同一时刻各变量的数据彼此差异很大, 而同一变量不同时刻的数据在短时间内差异不大, 先提取空间特征能够保留样本间更大的差异性, 学到更为丰富的样本特征。

为进一步分析状态监测的差异, 对上述方法的结果进行了定量对比。提出两个定量分析指标:

1) 将故障发生后的报警信号数量与样本数量之比定义为故障检出率。故障检出率越高, 表明该方法对故障的辨别能力越强。

2) 将重构误差的最大值 $RE_{\max}$ 与预警阈值 α 之比定义为故障灵敏度。故障灵敏度越高, 表明在该方法下, 正常与故障状态之间的差异更显著。表 3 为上述 5 种方法状态监测结果的定量对比情况。

通过观察上表可知: 从预警阈值来看, STAE 的预警阈值最小, 表明该方法对数据特征的学习能力最强; STAE 和 TAE 的故障检出率较高, 表明其

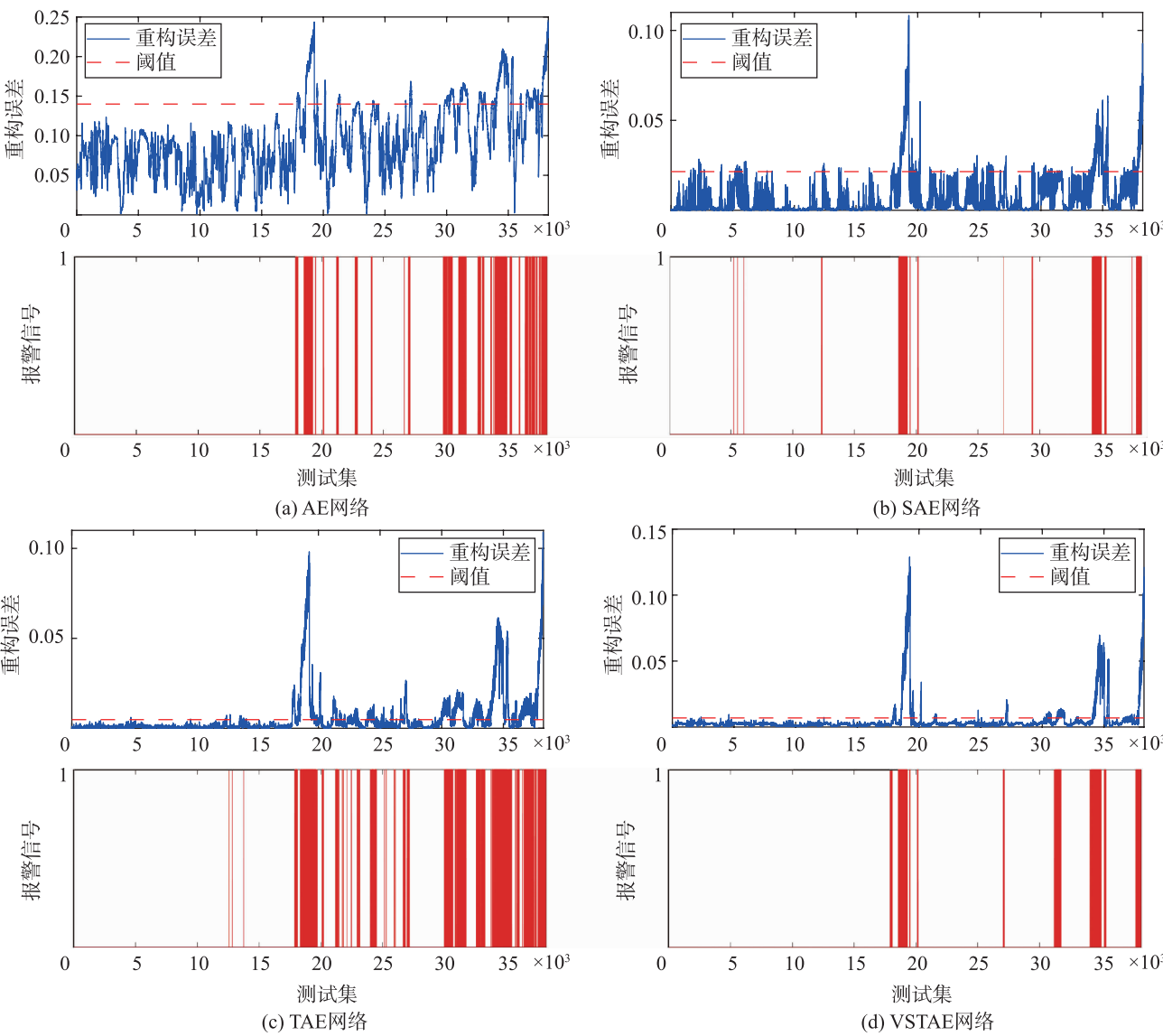


图 8 对比方法在测试集的重构误差及报警信号

Fig.8 Reconstruction error and alarm signal of comparative methods in test set

表 3 5 种方法的定量对比情况
Tab.3 Quantitative comparison of the five methods

方法	预警阈值	故障检出率 /%	故障灵敏度	误报警次数
STAE	0.004 3	46.21	56.09	0
AE	0.141 3	27.44	1.76	0
SAE	0.021 5	9.83	5.04	18
TAE	0.004 8	39.23	22.70	12
VSTAE	0.006 9	14.87	18.68	0

对故障的预警效果好，可以持续发出报警信号；STAE 的故障灵敏度远大于其他模型，表明齿轮箱发生异常后，故障趋势更显著。AE 预警阈值过高，对数据特征学习能力弱；SAE、TAE 存在误报警现象，不利于设备状态监测；VSTAE 故障检出率和灵敏度低，未能辨别设备非显著异常。综上所述，STAE 方法故障预警能力明显优于对比方法。

3.4 故障原因分析

STAE 网络的整体重构误差是由各分量重构误差的均值构成，可以计算得到各分量占比。如图 9

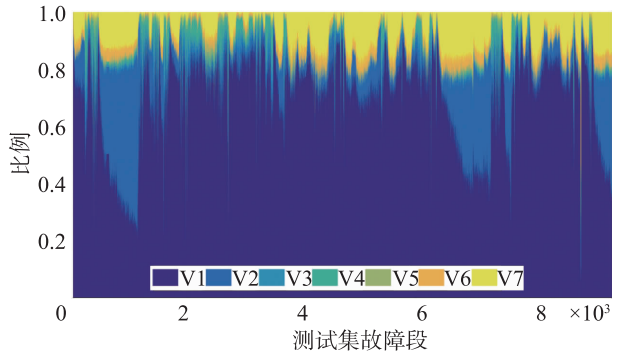


图 9 STAE 网络重构误差各分量占比
Fig.9 Ratio of components of STAE network reconstruction error

所示,为 STAE 网络测试集故障段的重构误差各分量占比。

如图 9 所示, V1、V2、V7 对应的齿轮箱入口油压、齿轮箱滤网前油压和齿轮箱油池温度的重构误差占比最高,可以确定本次齿轮箱故障中的主要异常参数为油路压力和油池温度。根据重构误差各分量占比变化趋势,推测齿轮箱故障起因是入口油压异常,随后影响到了滤网前油压,此时可能造成油液系统循环不畅,使得油温异常,最终导致油池温度高于上限值,风电机组停机。根据故障历史记录,仅知故障原因是齿轮箱油池温度高于上限值 70℃。由于缺乏设备维修记录和其他相关信息,实例中齿轮箱故障具体成因仍需结合实际情况进一步查验。

上述分析结果可以协助设备维护人员确定故障原因及维修策略,尽早解除齿轮箱故障造成的负面影响。

4 结论

为增强常规自编码网络针对面板 SCADA 数据

的特征提取能力,提出一种基于时空自编码网络的风电齿轮箱状态监测方法。使用河北省某风电场数据验证,得出结论如下:

1) 提出方法可有效提取面板数据中蕴藏的时空特征,实现对风电机组齿轮箱的实时状态监测。在实例中,提前 20 d 发现齿轮箱运行状态存在异常,提醒运维人员关注设备潜在故障。

2) 提出方法的状态监测效果明显优于常规方法。通过定性及定量分析可知,提出方法对数据特征的学习能力更强,故障检出率和故障灵敏度更高,且无误报警情况出现。

3) 通过对重构误差各分量贡献率及其变化趋势进行观察,确定实例中齿轮箱故障的主要异常参数为油路压力和油池温度,增强了状态监测结果的可解释性。

在现场中,由于传感器故障和数据传输干扰导致存在数据部分缺失的问题。本文是在完备数据集上进行的研究,如何使用欠完备数据进行状态监测是未来的研究方向。

参考文献

- [1] 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.
JIN X H, SUN Y, SHAN J H, et al. Fault diagnosis and prognosis for wind turbines: An overview[J]. Chinese Journal of Scientific Instrument, 2017, 38(5): 1041 - 1053.
- [2] REZAMAND M, KORDESTANI M, CARRIVEAU R, et al. Critical wind turbine components prognostics: A comprehensive review[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(12): 9306 - 9328.
- [3] 陈雪峰,郭艳婕,许才彬,等. 风电装备故障诊断与健康监测研究综述[J]. 中国机械工程, 2020, 31(2): 175 - 189.
CHEN X F, GUO Y J, XU C B, et al. Review of fault diagnosis and health monitoring for wind power equipment[J]. China Mechanical Engineering, 2020, 31(2): 175 - 189.
- [4] ZHU J, YOON J M, HE D, et al. Online particle-contaminated lubrication oil condition monitoring and remaining useful life prediction for wind turbines[J]. Wind Energy, 2015, 18(6): 1131 - 1149.
- [5] TANG J, SOUA S, MARES C, et al. An experimental study of acoustic emission methodology for in service condition monitoring of wind turbine blades[J]. Renewable Energy, 2016, 99: 170 - 179.
- [6] SALAMEH J P, CAUET S, ETIEN E, et al. Gearbox condition monitoring in wind turbines: A review[J]. Mechanical Systems and Signal Processing, 2018, 111: 251 - 264.
- [7] 王梓齐,刘长良. 基于 Box-Cox 变换和相对熵残差分析的风电机组齿轮箱状态监测[J]. 中国电机工程学报, 2020, 40(13): 4210 - 4219.
WANG Z Q, LIU C L. Wind turbine gearbox condition monitoring based on Box-Cox transformation and relative entropy residual analysis[J]. Proceedings of the CSEE, 2020, 40(13): 4210 - 4219.
- [8] YANG Y J, LIU A M, XIN H W, et al. Fault early warning of wind turbine gearbox based on multi-input support vector regression and improved ant lion optimization[J]. Wind Energy, 2021, 24(8): 812 - 832.
- [9] ZHAO H S, LIU H H, HU W J, et al. Anomaly detection and fault analysis of wind turbine components based on deep learning network[J]. Renewable Energy, 2018, 127: 825 - 834.
- [10] 苏连成,郭高鑫. 基于信息融合的风电机组关键部件状态识别[J]. 信息与控制, 2021, 50(3): 337 - 342, 349.

- SU L C, GUO G X. State identification of key components of wind turbine based on information fusion[J]. Information and Control, 2021, 50(3): 337–342, 349.
- [11] 黄婷婷, 余磊. SDAE-LSTM 模型在金融时间序列预测中的应用[J]. 计算机工程与应用, 2019, 55(1): 142–148.
- HUANG T T, YU L. Application of SDAE-LSTM model on financial time series forecasting[J]. Computer Engineering and Applications, 2019, 55(1): 142–148.
- [12] 黄荣舟, 汤宝平, 杨燕妮, 等. 基于长短时记忆网络融合 SCADA 数据的风电齿轮箱状态监测[J]. 太阳能学报, 2021, 42(1): 235–239.
- HUANG R Z, TANG B P, YANG Y N, et al. Condition monitoring of wind turbine gearbox based on LSTM neural network fusing SCADA data[J]. Acta Energaie Solaris Sinica, 2021, 42(1): 235–239.
- [13] FENG B, CHEN B, LIU H W. Radar HRRP target recognition with deep networks[J]. Pattern Recognition, 2017, 61: 379–393.
- [14] CAO V L, NICOLAU M, MCDERMOTT J. Learning neural representations for network anomaly detection[J]. IEEE Transactions on Cybernetics, 2019, 49(8): 3074–3087.
- [15] 王浙超, 曾九孙, 谢磊, 等. 基于去噪自编码器的故障隔离与识别方法[J]. 信息与控制, 2021, 50(6): 641–650.
- WANG Z C, ZENG J S, XIE L, et al. Fault isolation and identification method based on denoising autoencoder[J]. Information and Control, 2021, 50(6): 641–650.
- [16] JIANG G, HE H, YAN J, et al. Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox[J]. IEEE Transactions on Industrial Electronics, 2018, 66(4): 3196–3207.
- [17] HUA Y, ZHAO Z, LI R, et al. Deep learning with long short-term memory for time series prediction[J]. IEEE Communications Magazine, 2019, 57(6): 114–119.
- [18] 孟先艳, 崔荣一, 赵亚慧, 等. 基于双向长短时记忆单元和卷积神经网络的多语种文本分类方法[J]. 计算机应用研究, 2020, 37(9): 2669–2673.
- MENG X Y, CUI R Y, ZHAO Y H, et al. Multilingual text classification method based on bi-directional long short-term memory and convolutional neural network[J]. Application Research of Computers, 2020, 37(9): 2669–2673.
- [19] YANG L, ZHANG Z. A conditional convolutional autoencoder-based method for monitoring wind turbine blade breakages[J]. IEEE Transactions on Industrial Informatics, 2020, 17(9): 6390–6398.
- [20] RISULEO R S, BOTTEGAL G, HJALMARSSON H. Identification of linear models from quantized data: A midpoint-projection approach[J]. IEEE Transactions on Automatic Control, 2019, 65(7): 2801–2813.

作者简介

刘长良(1965–), 男, 博士, 教授。研究领域为风电机组故障预警与诊断, 热力系统建模与仿真等。

徐 健(1997–), 男, 硕士。研究领域为风电机组状态监测, 大数据分析。

王梓齐(1995–), 男, 博士。研究领域为风电机组的状态监测与故障预警。