

DOI: 10.19650/j.cnki.cjsi.J2006043

基于生成对抗网络的风电机组在线状态监测*

金晓航^{1,2,3}, 许壮伟², 孙毅^{1,2,3}, 单继宏^{1,2}, 王欣⁴

(1. 浙江工业大学 特种装备制造与先进加工技术教育部重点实验室 杭州 310023; 2. 浙江工业大学机械工程学院 杭州 310023; 3. 宁海县浙工大科学技术研究院 宁海 315600; 4. 浙江运达风电股份有限公司 杭州 310012)

摘要:基于风电机组数据采集与监控系统(SCADA)的大量时序数据分析,提出一种基于生成对抗网络(GAN)的风电机组在线状态监测方法。首先,通过GAN模型中的生成器获得一组与SCADA数据相同维度的生成数据;其次,利用生成的和真实的SCADA数据对GAN模型进行优化训练进而得到用于区分机组健康状态的判别器模型。利用所提方法分别对一台故障风电机组和一台健康风电机组的SCADA数据分析后发现:GAN方法能够有效监测风电机组的在线运行状态,比SCADA系统提早5d发现故障机组的异常;当风电机组正常工作时,GAN方法比其他方法(如马氏距离、主成分分析、深度神经网络、支持向量机等)误报的次数更少;当机组发生故障后,GAN方法比上述其他方法能检测出更多的异常样本。

关键词:风电机组;数据采集与监控系统;生成对抗网络;状态监测

中图分类号: TH17 文献标识码: A 国家标准学科分类代码: 460.99

Online condition monitoring of wind turbine based on generative adversarial network

Jin Xiaohang^{1,2,3}, Xu Zhuangwei^{1,2}, Sun Yi^{1,2,3}, Shan Jihong^{1,2}, Wang Xin⁴

(1. Key Laboratory of E & M, MOE, Zhejiang University of Technology, Hangzhou 310023, China; 2. College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023, China; 3. Ninghai ZJUT Academy of Science and Technology, Ninghai 315600, China; 4. Zhejiang Windey Co., Ltd., Hangzhou 310012, China)

Abstract: Based on the analysis of large amounts of time series data collected by supervisory control and data acquisition system (SCADA) in wind turbines, a wind turbine online condition monitoring approach based on generative adversarial network (GAN) is proposed. Firstly, a data set that has the same dimension with the SCADA data is generated with the generative model. Secondly, the generated SCADA data and the real SCADA data are used to optimize and train the GAN model. After training, the obtained discriminative model in GAN is used for distinguishing the health condition of wind turbines. Finally, the proposed approach was used to analyze the SCADA data of a healthy and a faulty wind turbines. The result shows that the GAN-based approach can effectively monitor the online operation condition of the wind turbines, it can detect the anomalies of the faulty wind turbine 5 days earlier than the SCADA system. When the wind turbine works normally, the number of false alarms reported by the GAN approach is less than other approaches (such as Mahalanobis distance, principal component analysis, deep neural network and support vector machine). When the wind turbine fails, the GAN-based approach can detect more abnormal samples than other approaches.

Keywords: wind turbine; supervisory control and data acquisition system (SCADA); generative adversarial network; condition monitoring

收稿日期: 2020-01-16 Received Date: 2020-01-16

* 基金项目: 国家自然科学基金(51505424, 51675484)、浙江省自然科学基金(LY15E050019)、宁波市自然科学基金(2018A610045)、浙江省大学生科技创新活动计划(新苗人才计划)项目资助

0 引言

随着对清洁能源需求的提升,风电行业快速发展,风电机组的装机容量逐年稳步提升^[1]。然而风电行业中存在着一个不容忽视的问题,即高昂的运维成本。大约有20%的风场运营收入花费在风电机组的运维上^[2,3]。为了提高风电机组的运行可靠性,防止由机械、电气故障引发的重大安全事故,优化风电机组的维护策略,需要对风电机组进行在线状态监测,进而在故障发生的早期进行异常预警,减少机组因故障带来的巨额经济损失。当前虽然一些风电机组状态监测和故障识别方法相继被提出,但也存在一定的问题:如基于状态监测系统(condition monitoring system, CMS)的状态监测方法需要安装额外的振动传感器和数据采集装置^[2],基于声学分析的风电机组故障诊断方法除了需要安装传声器阵列外还容易受到外界环境噪声的影响^[3,4]等。另外,这些方法普遍存在加装额外的传感器进而导致成本增加的问题,与此同时,每台大型风电机组在设计建造之初就将各类传感器集成安装在变桨、偏航、主轴、齿轮箱、发电机、变流器、变压器等风电机组的子系统或主要部件上,通过数据采集与监控系统(supervisory control and data acquisition system, SCADA)实现风电机组及其各部件工作状态监测与控制、历史运行信息查询、故障报警等功能。虽然SCADA系统监测了风电机组上的100个多运行状态数据,但通常是在风电机组发生了较严重的故障后,才会由SCADA系统报警。通过SCADA系统本身难以检测出风电机组早期故障,容易错失维护风电机组的最佳时机。然而SCADA系统监测并积累了大量的状态数据,因此怎样从这些数据中发现其蕴含的价值,实现机组故障或异常状态的早期预警,这在近些年来受到了风电厂商、风场业主的密切关注^[5]。

受人工智能领域中大数据分析方法的启发,较多学者研究了由SCADA数据驱动的风电机组状态监测方法。如Yang等^[6]基于多变量状态估计技术对SCADA数据进行分析,识别了风电机组齿轮箱故障;Gill等^[7]基于Copula函数建立了风电机组功率曲线的概率模型,利用SCADA数据研究了对风电机组叶片退化、偏航系统早期故障征兆进行有效监测的可行性;Meik等^[8]利用自适应神经模糊推理系统对风电机组的SCADA数据进行了分析,对风电机组的运行状态进行了识别;Chen等^[9]利用高斯过程和主成分分析方法对24台风电机组SCADA数据进行了分析,分辨出了其中正常、异常的风电机组;刘师等^[10]将高斯混合模型、动态时间规整及熵权值算法融合后,提出了一种基于群体多维特征相似性的故障预警策略,提前判断出了具有潜在故障的风电机组;Kusiak

等^[11]通过多台风电机组SCADA数据的比较,采用统计学方法对机组有功功率运行数据进行了分析,从而识别出了有故障的风电机组。上述研究工作大都是在较多风电机组故障数据的基础上进行的。然而实际中虽然风电机组故障频发^[12],但机组发生故障后会马上就会进行停机维修,不再采集数据,导致机组故障数据的样本量比较少、故障数据的获取比较困难、成本也较高。这使得基于故障数据学习分析,进而实现风电机组运行状态识别的方法实现比较困难。同时,在风电机组的日常运行过程中,SCADA系统持续工作,获取并积累了大量机组健康状态下的数据,因此怎样通过对这些健康数据的学习来实现机组的在线状态监测逐渐进入人们的视线。当前,此方面的研究主要是利用模型预测机组的某一特征值后与该特征的真实值之间进行残差分析,从而实现状态监测。如郭鹏等^[13]建立了多变量功率曲线模型,分析模型预测功率与真实功率的残差值,发现了风电机组的异常;贾子文等^[14]通过滑动窗口和残差分析发现了风电机组齿轮箱的异常。然而残差分析的方法容易受到噪声的干扰,残差在机组健康状态下出现超标的情况,使得机组时常出现误报警。

生成对抗网络(generative adversarial network, GAN)作为一种无监督学习模型,能够自动为数据添加分类标签,且网络模型经过训练后输出结果相对平稳,受外界噪声影响而导致的输出波动较小。因此,本文基于GAN模型提出一种完全基于风电机组健康运行状态下的SCADA数据建模方法,在实现机组健康状态监测的同时,减少误报警的次数。最后,利用实际SCADA数据对所提方法的有效性和鲁棒性进行了讨论与分析。

1 生成对抗网络算法

生成对抗网络模型^[15]是一种无监督的深度学习模型,模型中包含生成器和判别器两个模块。生成器负责生成数据,判别器负责区分数据,如图1所示。

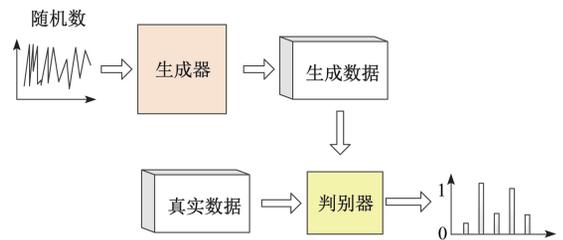


图1 生成对抗网络

Fig.1 The generative adversarial network

在生成对抗网络中,生成器和判别器一般都采用神经网络结构或卷积神经网络结构。其中卷积神经网络结

构大多用于处理图像数据^[16],而考虑到风电机组 SCADA 数据为时序数据,采用卷积神经网络结构将带来诸多不便,因此采用神经网络作为生成器和判别器的结构。假设经过神经网络第 k 层隐藏层的线性变化后输出为 $Z^{[k]}$,非线性变化后输出为 $A^{[k]}$,权重为 $W^{[k]}$,偏差为 $b^{[k]}$,则神经网络第 k 层隐藏层的传播公式如下:

$$Z^{[k]} = W^{[k]} \cdot A^{[k-1]} + b^{[k]} \quad (1)$$

$$A^{[k]} = g(Z^{[k]}) \quad (2)$$

式中: $g(\cdot)$ 为激活函数,选用 sigmoid 函数,如式(3)

所示:

$$g(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

假设真实数据样本数为 m ,每个样本存在 n 维的特征, $\{x_1, x_2, x_3, \dots, x_m\}$, $x_i \in R^n$,组成训练数据分布 $p_x(x)$,则输入相同维度的随机数样本 $\{z_1, z_2, z_3, \dots, z_m\}$, $z_i \in R^n$,组成随机数分布 $p_z(z)$ 。同时,生成器模型为 G ,判别器模型为 D ,则生成对抗网络的结构如图 2 所示。

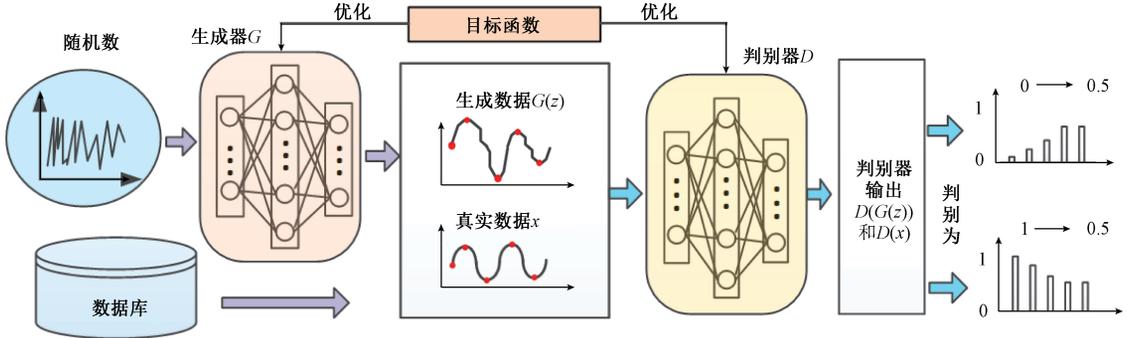


图2 生成对抗网络

Fig.2 Structure principle of the generative adversarial network

在生成对抗网络的训练过程中,生成器以生成与真实数据接近的生成数据,做到以假乱真为目标;判别器以区分出真实数据与生成数据,做到辨明真假为目标。训练时,相互交替训练生成器和判别器,使其在训练过程中的优化目标相互对抗,进而朝着各自的优化目标共同发展。在训练开始时生成器输出的生成数据将被判别器标记为0,真实数据将被标记为1。随着训练的深入,生成数据开始接近真实数据,判别器将逐渐无法分辨出生成数据和真实数据间的区别。当生成数据与真实数据输入判别器后,输出标记都收敛至0.5时,生成对抗网络训练完成,得到一个能模拟真实数据的生成器模型和一个能区分生成数据与真实数据的判别器模型。生成对抗网络定义的优化目标函数 $V(D, G)$ 如下:

$$\min_G \max_D V(D, G) = E_{x \sim p_x(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (4)$$

式中: E 表示均值; $x \sim p_x(x)$ 表示 x 来自真实数据; $z \sim p_z(z)$ 表示 z 来自随机数据。

训练过程中,将优化目标函数如下分别执行:

$$\min_G V(D, G) = E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (5)$$

$$\max_D V(D, G) = E_{z \sim p_z(z)} [\log(1 - D(G(z)))] + E_{x \sim p_x(x)} [\log D(x)] \quad (6)$$

同时生成对抗网络的具体训练步骤如下所示。

1) 设置超参数,如学习率,迭代次数,隐藏层神经元数等^[17],随机初始化生成器和判别器隐藏层神经元的权

重系数。

2) 将随机数样本 z 输入生成器,获得生成数据 $G(z)$ 。

3) 将真实数据 x 与生成数据 $G(z)$ 输入判别器模型,获得接近于1的输出 $D(x)$ 和接近于0的输出 $D(G(z))$ 。

4) 通过式(5)和梯度下降优化算法^[18]更新生成器模型参数,优化生成器模型。

5) 通过式(6)和反向传播算法更新判别器模型参数,优化判别器模型。

6) 循环交替执行步骤3)~4)直至判别器输出 $D(x)$ 和 $D(G(z))$ 均收敛至0.5,则训练完成。

2 基于 GAN 的风电机组健康状态监测

风电机组的运行情况一般可分为以下3个不同的阶段:1)当风速低于切入风速时,风电机组处于待机阶段,风电机组风轮自由旋转,机组不向电网输送功率;2)当风速处于切入风速和切出风速之间时,风电机组处于正常发电阶段,机组向电网输送功率,同时根据风速的大小又可以细分为最大风能捕获区和恒功率运行区;3)当风速大于切出风速时,为避免损坏,风电机组开始制动,进入停机阶段,如图3所示。基于GAN的健康状态监测方法是通过分析正常发电阶段风电机组运行数据实现机组异常状态的预警。

基于SCADA系统采集到的数据,提出了一种基于生

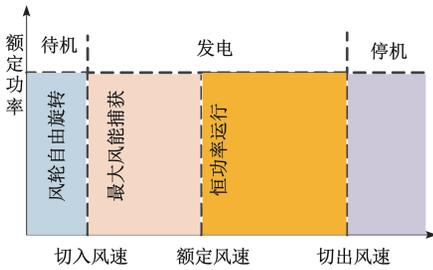


图 3 风电机组风速功率曲线

Fig.3 The curve of power vs. wind speed for the wind turbine

成对抗网络模型的在线监测方法,如图 4 所示。

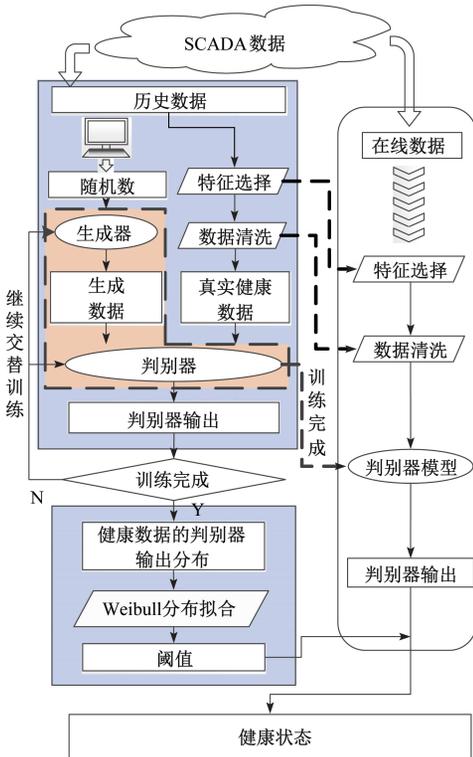


图 4 基于生成对抗网络的风电机组在线状态监测方法

Fig.4 Online condition monitoring approach of wind turbine based on generative adversarial network

首先,从 SCADA 数据库中获取风电机组健康状态下的历史数据后进行分析,根据监测的部位或子系统结合工程人员的经验选取对应的特征作为模型输入。然后根据输出功率,将图 3 中所述的待机阶段和停机阶段内的数据删除,从而筛选出运行阶段区域的数据。

其次,将预处理后的运行阶段的数据作为真实健康数据,用以后续模型训练所需。随后,将一组与真实健康数据维度相同且满足正态分布的随机数数据输入生成器模型,得到生成的健康数据。然后,将上述生成的健康数据和真实健康数据一起输入判别器模型,得到判别结果。

最后循环交替训练生成器和判别器模型,更新生成的健康数据与判别结果,直至训练完成。

训练完成后,判别器已经拥有判断输入是否来自健康数据分布的能力,此时,由上文所述生成对抗网络算法原理可知,将健康数据输入判别器模型时,其理论输出应该在 0.5 附近波动,同时越接近 1 代表输入数据是健康数据的概率越大,越接近于 0 代表输入数据是异常数据的概率越大。

利用 Weibull 分布确定阈值是一种较为常用的方法,同时有研究表明 Weibull 分布能够较好的描述现实中风速的分布^[19-21],考虑到正常状态下风电机组的工作状态和风速相关,因此借用 Weibull 分布^[22]拟合判别器对历史健康数据的输出概率分布,并通过极大似然估计法估计参数,得到最优的 Weibull 分布参数。Weibull 分布的概率密度函数为:

$$f(x; \lambda, k) = \begin{cases} \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (7)$$

对应的累积概率分布函数为:

$$F(x) = \begin{cases} 1 - e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (8)$$

式中: f 为概率密度函数; F 为累积概率分布函数; λ 和 k 为待估计参数。

随后设定 Weibull 分布中用于区分健康和异常数据的阈值,并转化为判别器模型输出数据对应的阈值,进而实现风电机组健康状态的监测。

3 案例分析

3.1 案例一:故障机组的在线状态监测

1) 风电机组 SCADA 数据分析和清洗

某风场一台风电机组的 SCADA 系统在 2017 年 2 月 5 日 11 时 20 分报“发电机轴承温度过高故障停机”故障。风场运维人员对该台风电机组进行了检修,发现该台机组的发电机在转动时内部有异常的摩擦声响,拆卸发电机后进一步发现其非驱动端轴承已经发生了严重的故障,如图 5 所示。从图 5 中可以看到轴承的滚动体表面附上了一层铜,这是由于轴承滚动体与内外圈之间原先的滚动摩擦变为滑动摩擦,导致了轴承温度的升高,并进一步触发了 SCADA 系统的报警。“发电机非驱动端轴承”的特征参量在 2017 年 2 月(2017 年 2 月 5 日~2017 年 2 月 20 日)出现大幅度异常变化,如图 6 所示。

取该台风电机组 2016 年 1 月到 2017 年 4 月的 SCADA 数据进行分析,将 2016 年 1~3 月的数据作为历史健康数据训练模型,将 2016 年 4 月~2017 年 4 月的数据作为测试数据,验证 GAN 方法对风电机组状态监测的

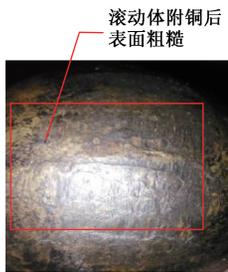


图5 轴承滚动体故障
Fig.5 Bearing ball fault

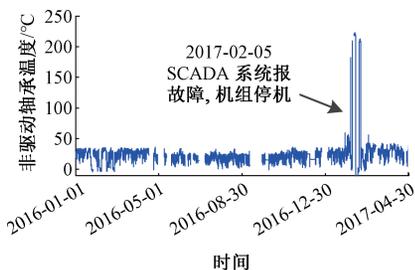


图6 风电机组发电机非驱动端轴承温度时序图
Fig.6 Bearing temperature time sequence diagram of the wind turbine generator at the non-driven end

效果。首先对 SCADA 数据进行特征选择。鉴于发电机发生了故障,在咨询了风电机组厂商后,选取 SCADA 数据中与发电机健康状态相关的特征数据进行分析。具体的特征数据包括:有功功率、风速、风向、发电机转速、发电机绕组温度、发电机驱动端轴承温度和发电机非驱动端轴承温度。观察数据后发现在一些时间段上出现了数据缺失,如图6中2016年8月30日后出现的空白区域,这是由于风速过高等原因导致风电机组停机,SCADA 系统没有采集到相关时间段上的数据。因此在模型训练的时候,需要对风电机组停机时段的 SCADA 数据进行预处理工作。

将2016年1~3月的 SCADA 数据作为训练数据,并进行数据清理工作。首先,去除数据样本中的机组停机时的空白数据。其次,根据风速和发电功率间的关系,将如图7所示风电机组在风速不为0且无功率输出的时间点或时间段上的数据删除,整理出机组正常发电阶段的健康数据,如图8所示。图8中有少量的离群点,查询这些离群点的原始 SCADA 数据发现,由于传感器采集出错或噪声过大等原因产生了上述的离群数据点,与风电机组工作状态无关,因此在建模分析的时候,未考虑上述离群点数据。

2) 基于 GAN 的风电机组状态监测

经过数据清洗后,留下的2016年1~3月的样本数为19463个,将其作为训练集。同时生成相同样本数的同

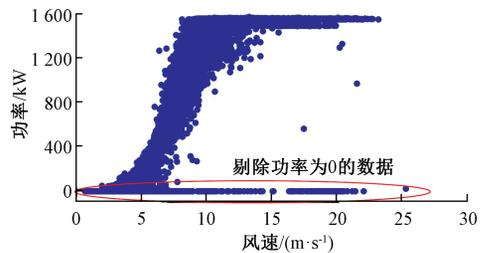


图7 数据清洗前风电机组风速与发电功率散点图
Fig.7 Scatter plot of power vs. wind speed for the wind turbine before data cleaning

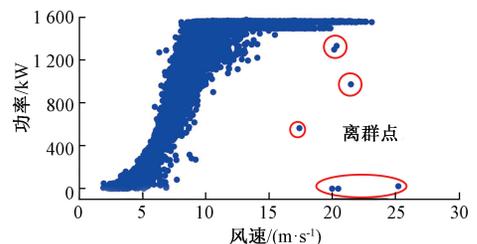


图8 数据清洗后风电机组风速与发电功率散点图
Fig.8 Scatter plot of power vs. wind speed for the wind turbine after data cleaning

维度随机数数据,将生成的随机数据和训练集数据输入 GAN 进行模型训练。

训练过程中利用 Adam 方法^[23-24]对生成对抗网络模型参数进行优化计算,设置生成器和判别器的隐藏层数均为1,隐藏层神经元数均为16,优化的初始学习率为0.0001,训练次数为8000。训练完成后,生成对抗网络中的判别器将输出训练集中各样本的标签,如图9所示,绘制其分布直方图,并利用 Weibull 分布拟合该分布,得到 Weibull 分布的参数 $k=107.8, \lambda=0.505$ 。由生成对抗网络原理可知,判别器输出都在0.5附近波动,同时判别器输出越接近于1,输入数据越接近健康数据,输出越接近0,输入数据越区别于健康数据,因此异常数据都偏向于0,阈值只需设置下限。设置允许的异常数据比例 x ,利用 Weibull 分布累计概率密度分布函数计算其所对应的阈值 T ,使得区间 $[0, T]$ 内包含的数据比例为 x ,此处取 $x=0.3\%$,求得 $T=0.478$,如图10所示。

利用训练完成的判别器模型,监测风电机组2016年4月份以后的测试数据,直至风电机组发生故障进行停机维修的2017年2月,判别器模型输出如图11(a)所示。图11(a)中的绿色虚线为风电机组 SCADA 系统进行报警的时间:2017年2月5日,对红色虚线圈内的数据点即报警时间点附近的数据进行放大,如图11(b)所示。从图11(b)中可以看到判别器输出在“2017-02-01 00:10:00”首次超出阈值,并于随后几天内都出现超出阈

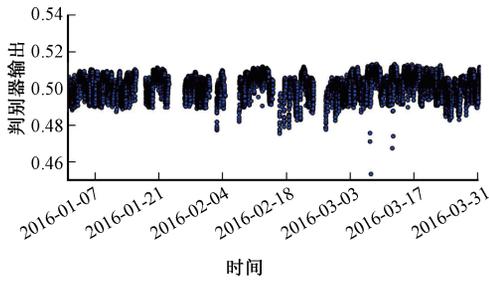


图 9 GAN 判别器对训练集(健康风电机组数据集)的输出
Fig.9 GAN discriminator output for training data set (healthy wind turbine data set)

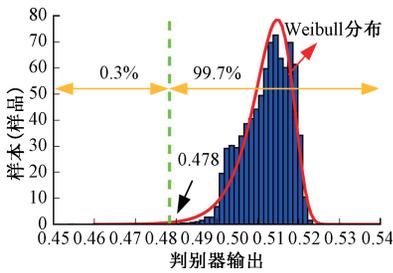


图 10 判别器对训练集(健康风电机组数据集)的输出直方图
Fig.10 Histogram of GAN discriminator output for training data set (healthy wind turbine data set)

发出警报,由此可知,本文基于生成对抗网络的在线监测方法提早 5 d 检测出了风电机组发电机异常。

该台风电机组在经过一段时间的维修后,继续投入使用,利用上述方法对风电机组的健康状态进行持续监测后发现:更换了故障轴承,刚刚结束维修后的发电机一开始工作状态并没有恢复到原先理想的工作状态,具体表现为利用 GAN 方法监测到发电机维修后的一段时间内(即 2017 年 2 月 20 日~2017 年 2 月 27 日)仍然存在判别器输出超出阈值的情况,其他时间点上判别器输出均未超出阈值,如图 12 所示。考虑到新轴承可能存在磨合期、安装误差等情况,因此认为风电机组发电机在维修后运行状态良好。

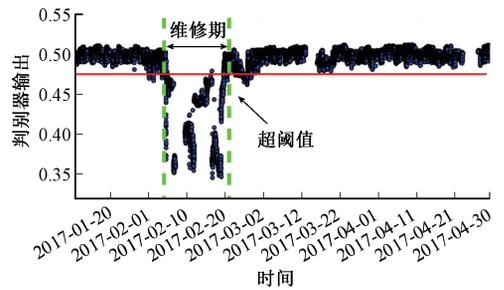
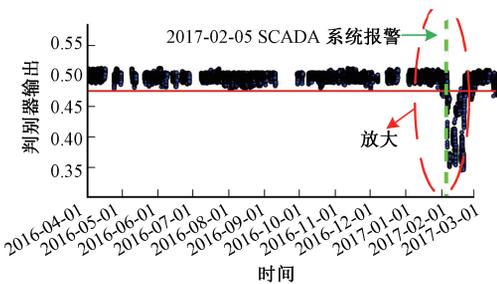
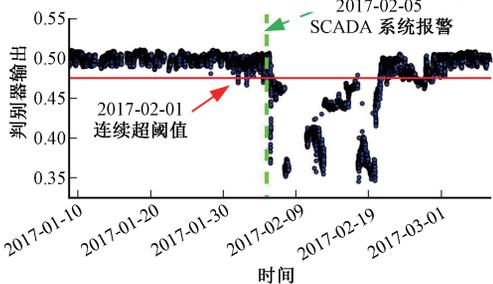


图 12 风电机组维修后在线状态监测结果
Fig.12 Online condition monitoring result of wind turbine after repairing



(a) 2016年4月至2017年3月的判别器输出
(a) The discriminator output from April 2016 to March 2017



(b) 2017年2月5日前后1个月的判别器输出局部放大图
(b) The partial enlarged discriminator output in the period one month before and after 5th, February 2017

图 11 不同时间段内风电机组在线状态监测结果
Fig.11 Online condition monitoring results of wind turbine in different time periods

3.2 案例二:健康机组的在线状态监测

利用本文方法对另外一台风电机组进行健康状态监测。该台风电机组在 2016 年 1 月~2016 年 10 月期间内断断续续工作,未发生过故障。将该台风电机组的 2016 年 1 月~2016 年 3 月的数据作为训练集,2016 年 4 月~2016 年 10 月的数据作为测试集。

参照前述章节对训练集数据进行清洗工作后,用于训练生成对抗网络模型。训练完成后,通过 Weibull 分布拟合判别器的输出结果,进而得到该台风电机组的监测阈值为 0.454。随后利用训练完成的判别器模型分析该机组测试集(2016 年 4 月~2016 年 10 月)数据,结果如图 13 所示。从图 13 中可以看到判别器的输出基本在红色阈值线以上,表明机组健康运行,符合实际情况。但在 2016-06-27 13:40:00 出现一个超阈值、突变的离群点,进一步分析机组的原始数据后发现,此离群点对应的风速值为 22.247 m/s,该风速远大于 2016 年 6 月内其他时刻的风速且接近机组的切出风速,机组在该时刻对应的其他特征(如发电功率、温度等)未见异常,因此推测由于阵风等原因使得机组周围环境风速短时间内过大,但机组工作正常。

值范围的情况,而 SCADA 系统则在 2017 年 2 月 5 日才

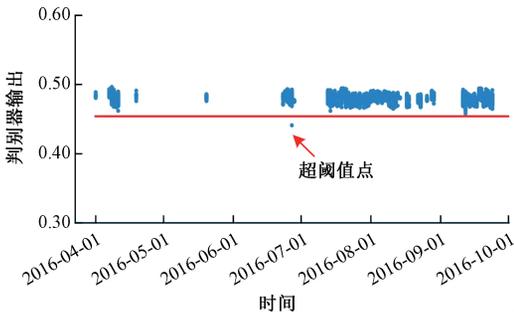


图 13 另一台正常工作风电机组的在线状态监测结果

Fig.13 Online condition monitoring result of another normal working wind turbine

3.3 与其他方法的对比分析

由于 SCADA 数据中健康数据和故障数据样本量相差很大,数据存在严重的不平衡问题,导致按准确率和误警率等指标来衡量各方法的性能相对困难,例如:在统计所有监控样本的情况下,误警率为误警次数与总样本量的比值,然而在对比过程中由于误警次数少、总样本量过多等原因各方法的误警率都在 0.01% 以下,数值较小且较为接近,对比相对困难。然而风电机组实际运行过程中每一次误警都会给风电厂商带来较大的损失和困扰,因此需要其他指标来衡量方法的性能。在实际工业场景下,系统报警的次数与方法监测结果中超阈值的样本量和报警策略有关,因此本文在对比中采用健康月份平均超出阈值样本量 a 和故障月份超出阈值的样本量 b 作为评判方法的指标。随后分别利用马氏距离 (Mahalanobis distance, MD)^[25]、主成分分析 (principal component analysis, PCA)^[25-26]、单分类支持向量机 (one-class support vector machine, OC-SVM)^[26] 和深度神经网络 (deep neural network, DNN)^[26] 等方法,构建模型同样对案例一中的风电机组进行了健康状态的监测研究,并通过评价指标对结果进行了对比,如表 1 所示。从表 1 中可以发现,基于 GAN 的方法敏感度一般,其报警时间为 2017 年 2 月 1 日,相较 PCA 和 OC-SVM 在预警时间上轻微落后。但是基于 GAN 的方法在鲁棒性上有较好的性能,能减少样本的误判,可以发现 GAN 方法的 a 值仅为 1,远小于其他方法,代表其在健康月份中判断出的异常样本量最少,同时其 b 值最大,代表 GAN 方法在故障月份判断出的故障样本量最多。由此可知虽然基于 GAN 方法检测早期故障的敏感度降低了,但是该方法的鲁棒性却增强了。此外,相较 DNN 和 MD 的方法,基于 GAN 的方法在拥有鲁棒性的优势下,在预警时间上有更好的表现。

4 结 论

现有风电机组状态监测方法主要是建立在失效物理

表 1 不同监测方法结果对比表

Table 1 Comparison of the monitoring results for different approaches

方法	健康月份平均 超阈值样本量 a	故障月份超阈 值样本量 b	预警时间
MD	6	3 749	2017-02-01 00:55:00
DNN	17	2 687	2017-02-01 00:10:00
PCA	13	2 870	2017-01-28 05:40:00
OC-SVM	3	3 782	2017-01-28 03:20:00
GAN	1	4 378	2017-02-01 00:10:00

模型和异常数据分析的基础上,然而实际中失效物理模型构建困难,异常数据样本量相对较少,但是 SCADA 系统监测并积累了海量机组正常运行的状态数据,因此基于生成对抗网络模型提出一种完全基于正常运行状态的 SCADA 数据建模的风电机组健康状态监测方法。随后,分别利用一台故障风电机组和一台健康风电机组的 SCADA 数据对基于 GAN 的状态监测方法进行了验证分析。结果表明,基于 GAN 的方法能够比 SCADA 系统提早 5 d 检测出故障风电机组的异常情况,并且比 MD、PCA、OC-SVM、DNN 等方法在风电机组异常检测上具有更好的性能,所提方法对健康风电机组的监测结果符合实际情况。

参考文献

- [1] Global Wind Energy Council. Global wind report 2018: Annual market update[R]. GWEC, 2019.
- [2] 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(5): 1042-1053.
JIN X H, SUN Y, SHAN J H, et al. Fault diagnosis and prognosis for wind turbine: An overview [J]. Chinese Journal of Scientific Instrument, 2017(5):1042-1053.
- [3] JIN X, QIAO W, PENG Y, et al. Quantitative evaluation of wind turbine faults under variable operational conditions [J]. IEEE Transactions on Industry Applications, 2016, 52(3): 2061-2069.
- [4] ANAYA M, DIBADUIZA D A, FORERO E, et al. An acousto-ultrasonics pattern recognition approach for damage detection in wind turbine structures [C]. 20th Symposium on Signal Processing, Images and Computer Vision, 2015:1-5.
- [5] 李丹, 邱颖宁, 冯延晖. 风力发电系统报警信号故障诊断[J]. 太阳能学报, 2017, 38(11): 3138-3143.
LI D, QIU Y N, FENG Y H. Alarm fault diagnosis of

- wind turbine system [J]. *Acta Energiæ Solaris Sinica*, 2017, 38(11): 3183-3143.
- [6] YANG W X, COURT R, JIANG J S. Wind turbine condition monitoring by the approach of SCADA data analysis[J]. *Renewable Energy*, 2013, 53: 365-376.
- [7] GILL S, STEPHEN B, GALLOWAY S. Wind turbine condition assessment through power curve copula modeling[J]. *IEEE Transactions on Sustainable Energy*, 2012, 3(1): 94-101.
- [8] MEIK S, ILMAR F S, SOFIANE A. Wind turbine condition monitoring based on SCADA data using normal behavior models [J]. *Applied Soft Computing*, 2013(13): 259-270.
- [9] CHEN N Y, YU R R, CHEN Y, et al. Hierarchical method for wind turbine prognosis using SCADA data[J]. *IET Renewable Power Generation*, 2017, 4(11): 403-410.
- [10] 刘帅, 刘长良, 甄成刚, 等. 基于群体多维相似性的风机齿轮箱预警策略[J]. *仪器仪表学报*, 2018, 39(1): 180-189.
- LIU SH, LIU CH L, ZHEN CH G, et al. Fault warning strategy of wind turbines gearbox based on group multi-dimensional similarity[J]. *Chinese Journal of Scientific Instrument*, 2018, 39(1): 180-189.
- [11] KUSIAK A, VERMA A. Monitoring wind farms with performance curves [J]. *IEEE Transactions on Sustainable Energy*, 2013, 4(1): 192-199.
- [12] BADIHI H, ZHANG Y M, HONG H. Wind turbine fault diagnosis and fault-tolerant torque load control against actuator faults[J]. *IEEE Transactions on Control Systems Technology*, 2015, 23(4): 1351-1372.
- [13] 郭鹏, 刘琳. 多变量风电机组功率曲线建模与监测研究[J]. *电网技术*, 2018, 42(10): 3347-3354.
- GUO P, LIU L. Modeling and monitoring of multivariable wind turbine power curve[J]. *Power System Technology*, 2018, 42(10): 3347-3354.
- [14] 贾子文, 顾煜炯. 基于数据挖掘的风电机组齿轮箱运行状态分析[J]. *中国机械工程*, 2018, 29(6): 650-658.
- JIA Z W, GU Y J. Wind turbine gearbox operation state analysis based on data mining[J]. *China Mechanical Engineering*, 2018, 29(6): 650-658.
- [15] GOODFELLOW I J, POUGET J, MIRZA M, et al. Generative adversarial nets [C]. *Neural Information Processing Systems*, 2014, 2: 2672-2680.
- [16] 毛典辉, 李子沁, 蔡强, 等. 基于 DCGAN 反馈的深度差分隐私保护方法[J]. *北京工业大学学报*, 2018, 44(6): 870-877.
- MAO D H, LI Z Q, CAI Q, et al. Tickling deep differential privacy protection method based on DCGAN[J]. *Journal of Beijing University of Technology*, 2018, 44(6): 870-877.
- [17] 赵光权, 刘小勇, 姜泽东, 等. 基于深度学习的轴承健康因子无监督构建方法[J]. *仪器仪表学报*, 2018, 39(6): 82-88.
- ZHAO G Q, LIU X Y, JIANG Z D, et al. Unsupervised health indicator of bearing based on deep learning[J]. *Chinese Journal of Scientific Instrument*, 2018, 39(6): 82-88.
- [18] 温江涛, 闫常弘, 孙洁娣, 等. 基于压缩采集与深度学习的轴承故障诊断方法[J]. *仪器仪表学报*, 2018, 39(1): 171-178.
- WEN J T, YAN CH H, SUN J D, et al. Bearing fault diagnosis method based on compressed acquisition and deep learning [J]. *Chinese Journal of Scientific Instrument*, 2018, 39(1): 171-178.
- [19] 郭楚珊, 郭鹏, 杨锡运. 基于三适应度粒子群算法的风速威布尔分布参数估计[J]. *太阳能学报*, 2019, 40(1): 206-212.
- GUO CH SH, GUO P, YANG X Y. Wind weibull distribution parameter estimation based on three fitness PSO algorithm[J]. *Acta Energiæ Solaris Sinica*, 2019, 40(1): 206-212.
- [20] 廖菲, 邓华, 曾琳, 等. 南海北部海面风速概率分布特征[J]. *海洋学报*, 2018, 40(5): 37-47.
- LIAO F, DENG H, ZENG L, et al. The probability distribution of sea surface wind speeds over the northern South China Sea[J]. *Acta Oceanologica Sinica*, 2018, 40(5): 37-47.
- [21] 戴迪昊, 钱泽平, 丁明月. 基于 Weibull 分布的最大风速变点估计[J]. *大学数学*, 2018, 34(5): 12-18.
- DAI D H, QIAN Z P, DING M Y, et al. Change point estimation of annual maximum wind speed based on Weibull distribution [J]. *College Mathematics*, 2018, 34(5): 12-18.
- [22] EDILSON M A, GABRIEL A, AUGUSTO, P, et al. Q-Weibull applied to Brazilian hydropower equipment[J]. *IEEE Transactions on Reliability*, 2019, 68(1):

122-132.

- [23] KINGMA D, BA J. Adam: A method for stochastic optimization [C]. International Conference on Learning Representations, 2015: 1-15.
- [24] WANG L, ZHANG Z, LONG H, et al. Wind turbine gearbox failure identification with deep neural networks [J]. IEEE Transactions on Industrial Informatics, 2017, 13(3): 1360-1368.
- [25] JIN X, MA E W M, CHENG L, et al. Health monitoring of cooling fans based on Mahalanobis distance with mRMR feature selection [J]. IEEE Transactions on Instrumentation and Measurement, 2012, 61(8): 2222-2229.
- [26] STETCO A, DINMOHAMMADI F, ZHAO X, et al. Machine learning methods for wind turbine condition monitoring: A review [J]. Renewable Energy, 2019, 133: 620-635.

作者简介



E-mail: xhjin@zjut.edu.cn

Jin Xiaohang (Corresponding author) received his B. Sc. degree from Zhejiang University of Technology in 2003, M. Sc. degree from Shanghai Jiao Tong University in 2006 and Ph. D. degree from City University of Hong Kong in 2014. Now, he is an associate professor in the College of Mechanical Engineering, Zhejiang University of Technology. His research interests include signal analysis, reliability evaluation, health monitoring, fault diagnosis and prognosis, and intelligent maintenance for electromechanical equipment.

金晓航 (通信作者), 2003 年于浙江工业大学获得学士学位, 2006 年于上海交通大学获得硕士学位, 2014 年于香港城市大学获得博士学位, 现为浙江工业大学机械工程学院副教授, 主要研究方向为机电装备的信号分析、可靠性评估、健康监测、故障诊断与预后、智能维护。