Anomaly Prediction for Wind Turbines Using an Autoencoder with Vibration Data Supported by Power-Curve Filtering

Masaki TAKANASHI†, Member, Shu-ichi SATO†, Kentaro INDO†, Nozomu NISHIHARA†, Hiroki HAYASHI†, and Toru SUZUKI††, Nonmembers

SUMMARY The prediction of the malfunction timing of wind turbines is essential for maintaining the high profitability of the wind power generation industry. Studies have been conducted on machine learning methods that use condition monitoring system data, such as vibration data, and supervisory control and data acquisition (SCADA) data to detect and predict anomalies in wind turbines automatically. Autoencoder-based techniques that use unsupervised learning where the anomaly pattern is unknown have attracted significant interest in the area of anomaly detection and prediction. In particular, vibration data are considered useful because they include the changes that occur in the early stages of a malfunction. However, when autoencoder-based techniques are applied for prediction purposes, in the training process it is difficult to distinguish the difference between operating and non-operating condition data, which leads to the degradation of the prediction performance. In this letter, we propose a method in which both vibration data and SCADA data are utilized to improve the prediction performance, namely, a method that uses a power curve composed of active power and wind speed. We evaluated the method’s performance using vibration and SCADA data obtained from an actual wind farm.

key words: anomaly prediction, autoencoder, vibration data, SCADA data, power-curve filtering

1. Introduction

Wind energy power generation technologies have been widely studied and applied, as wind power is considered one of the most important environmentally friendly renewable energy sources. A crucial objective of wind-farm management is to minimize the downtime caused by main component failures. When a main component fails, the resumption of operation may be substantially delayed, as a result of the component’s long lead time and the transportation of its replacement from a distant location. This can cause significant production losses, and therefore, it is essential to predict the failure timing so that preparations to replace the component can be made in advance.

Many studies have been conducted on the detection of anomalies in wind turbines under these circumstances. In general, detection methods can be categorized into two major approaches: model-based and data-driven. In the data-driven approach, machine learning methods are applied to the data of condition monitoring systems (CMSs), which monitor components of wind turbines, or supervisory control and data acquisition (SCADA) data, which are used to monitor and control wind turbine operation. In modern wind turbines, standard SCADA systems are usually installed. SCADA data include meteorological conditions, such as wind speed and direction, component temperatures, and electrical measurements, such as the generated power. Autoencoder (AE)-based techniques for detecting anomalies in SCADA data have also been studied and have demonstrated improved detection performance [1]–[4]. Abundant anomaly data are not usually available for wind turbines, and thus, the main advantage of an AE is that it can be applied using unsupervised learning. Although SCADA data are favored because of the relatively low system installation cost involved, the data frequency is not very high, and they are occasionally unsuitable for predicting anomalies of critical components, such as gearboxes and shafts.

Apart from those utilizing SCADA data, studies have also been conducted using wind turbine condition monitoring data, such as vibration, acoustic, oil, and temperature data. The installation of monitoring equipment on wind turbines is more expensive than that of a SCADA system; however, this approach is expected to predict anomalies of the critical components effectively because of the high sampling rate. The use of Gaussian mixture models and their combination with deep neural networks to detect anomalies in the vibration data of wind turbines have been studied [5]–[8]. In [7], the authors developed a testbed of a generator motor and applied an AE for anomaly detection. In [8], the authors proposed a method to train a neural network to classify normal and abnormal data. They used the multivariate data obtained in a hidden layer of the trained network as input to the anomaly detection method. However, its efficacy when applied to real wind turbine generator data has not been proven, and the method requires a dataset containing normal and abnormal labels for learning.

We applied an AE to real wind turbine data. However, we found that the prediction performance deteriorated because we had attempted to learn the AE normal conditions using data that also included data of the non-operating condition in which power was not generated. To improve the prediction performance, the elimination of vibration data generated in non-operating conditions is thus considered a clearly necessary countermeasure. In this letter, we pro-
pose a method that utilizes SCADA data to exclude irrelevant data and improve the prediction performance.

The remainder of this paper is organized as follows. Section 2 describes an anomaly prediction method using an AE. Sections 3 and 4 present the specifications of the vibration and SCADA data used in this study, respectively. In Sect. 5, a method for improving the prediction performance by utilizing the power curve derived from the SCADA data is proposed. Section 6 presents our evaluation of the method’s performance when our processed data are used. Section 7 summarizes this letter.

2. Anomaly Prediction Using An Autoencoder

In this section, we briefly describe anomaly detection by means of an AE. A typical AE is shown in Fig. 1. An AE is a type of neural network, where the input is multivariate data and the same number of multivariate data are the output. The input data are propagated using the network coefficients, $W_i$, between two layers, and an activation function, $f(\cdot)$:

$$h_{i+1} = f(W_i h_i).$$

(1)

where $h_i$ denotes the values at the $i$-th hidden layer. A nonlinear function, such as a rectified linear unit (ReLU), sigmoid function, or hyperbolic tangent function, is usually employed as the activation function. This nonlinear functionality is necessary for employing multiple hidden layers in the network to handle the complicated nonlinear characteristics of the input data. The network coefficients to minimize the difference between the inputs and outputs should be obtained:

$$\text{minimize} \| x - \hat{x} \|^2$$

(2)

When the AE has $I$ hidden layers, the input variables, $x$, and the output variables, $\hat{x}$, satisfy

$$x = h_0$$

(3)

$$\hat{x} = W_I h_I$$

(4)

At this time, the number of variables in the middle layer of the network is set to a value that is smaller than the number of variables for the input and output. By creating this bottleneck layer in the network, it is possible to obtain network coefficients that extract low-dimensional characteristic data representing the input data.

Anomaly detection using an AE is frequently conducted through unsupervised learning, and network coefficients are learned using normal data in which no anomalies are assumed to occur. Next, during testing, the abnormal data are assumed to be located far from the normal data in the multivariate space. Abnormal data cannot be reconstructed correctly by an AE because it cannot capture the features of data that are not included in the learning process. This reflects a larger magnitude of the difference between the input and output variables. By tracking this reconstruction error, we can not only detect but also predict anomalies.

3. Vibration Data

We attempted to predict anomalies of a wind turbine by applying an AE to the vibration data obtained from sensors mounted close to the critical components. In practice, accelerometer sensors are deployed close to the main bearing, gear box, generator, etc. The specifications of the various sensors are usually different because different sampling frequencies, durations, etc. are suitable for analyzing their characteristics. In general, sensors’ sampling rates and sampling durations are set to several kilohertz and several seconds, respectively. To suppress the total number of the acquired data, the vibrations are recorded only a few times per day. By virtue of these settings, by applying an AE to the vibration data obtained from each of the components and comparing the results, we can predict anomalies and identify a malfunctioning component.

When we apply AEs to the vibration data in a straightforward manner, it is difficult to predict anomalies. This is because vibration does not explicitly determine whether a wind turbine is operating. Because the vibration data in the non-operating condition do not contain the vibration data generated from the critical components, such data are considered useless for differentiating normal from abnormal conditions. This causes the prediction performance of the AE-based anomaly detection method to deteriorate because it was trained also on non-operating condition data. Thus, it is expected that the AE-based prediction performance will be improved by eliminating the non-operating condition data. To achieve this, we propose using SCADA data to identify the operating condition data of wind turbines.

4. Supervisory Control and Data Acquisition Data

In modern wind turbines, standard SCADA systems are installed to capture operational status data, which include meteorological conditions, such as wind speed and direction, component temperatures, and electrical measurements, such as the generated power. The operating conditions can easily be estimated by using these operational data. Let us take the wind speed and active generated power as examples, as shown in Fig. 2. In this figure, the ideal power curve is represented by a solid curve. This curve is separated into three
stages: no power generation at low wind speed, linear power generation at middle speed, and constant power generation at high speed; the actual power curve differs somewhat from the ideal power curve. The dense points on the horizontal axis may be obtained when no power is being generated, although the wind is blowing. This situation occurs because power generation is suspended for reasons such as maintenance, precautionary protection, and electric power demand. When the active power is equal to zero, we can readily estimate that the wind turbine is not operating. By utilizing this curve, focusing on the linear and constant power generation areas, we can estimate the operation conditions and improve the prediction performance obtained by using the vibration data.

5. Power-Curve Filtering

In this section, we propose a filter preprocessing method to eliminate non-operational vibration data by utilizing the power curve. The prerequisite for applying this filtering technique is that wind speed and active power are included in the SCADA data. Here, we set specific filtering ranges for wind speed and active power. If the combination of wind speed and active power in a certain snapshot is within the specified range, the corresponding vibration data are used for anomaly prediction using the AE; otherwise, they are omitted. Figure 3 shows a schematic image of the filter preprocessing method, where active power, wind speed, and vibration data have been obtained. The range is set only to the linear power generation area shown within the solid box in the figure. The black circles and waveforms shown on the left hand side of the figure represent the observed SCADA data and the corresponding vibration data, respectively. The waveforms related to the black circles on the right hand side of the figure are retained and applied to an AE, whereas those related to the white circles are excluded. This is the power-curve filtering preprocessing method for eliminating the non-operating vibration data.

Figure 4 shows the entire process of the prediction technique. In this process, we employ multiple sensors mounted around a target component. First, the vibration data obtained from the sensors are filtered using the criterion mentioned above, and time- and frequency-domain vibration data are obtained by applying a short-windowed fast Fourier transform (FFT) algorithm. Then, the multiple vibration data are concatenated along the frequency domain. Subsequently, we apply a logarithmic transformation and an averaging over the time domain and obtain the frequency domain data, which are used as input data for an AE. The anomaly prediction is conducted using an AE.

6. Performance Evaluation

This section describes anomaly prediction conducted by applying our proposed method using the vibration and SCADA data captured from a real wind turbine. We attached sensors close to the main bearing, gearbox, and generator as the target components. In this evaluation, we applied a 1024-sample FFT to the vibration data of each sensor. Because we had obtained imaging data for the frequency region higher than 512, we used only the lower 512 frequency point data. The number of nodes in the input and output layers was 1536, because we utilized three sensors deployed around each target component. We employed an AE with five layers, where the number of nodes of the first, second, and third hidden layers were 512, 128, and 512, respectively. The ReLU function was applied as the activation function in the AE. We trained the AE using the first seven-month vibration data, from August 1st, year Y to March 1st, year Y+1, and attempted to predict anomalies for the subsequent five months. A warning appeared for June 16th, year Y+1.

We present the results obtained with and without power curve filtering in Figs. 5 – 7. We applied a filter, the value ranges of which were 5 to 12 for wind speed and more than 1 for active power. These ranges were determined to capture the data associated with the linear area in the power curve, where the wind turbine is considered to function con-
sistently. In this study, we simply evaluated the prediction performance using a reconstruction error that is equivalent to the square difference between the input and output of the AE:

\[ \text{Reconstruction error} = \| x - \hat{x} \|^2. \quad (5) \]

The error value can be obtained at most three times per day. We derived a median value calculated from the three consecutive error values and plotted them in the figures.

No large errors can be observed in the case of the main bearing and gearbox, shown in Figs. 5 and 6, when the filter was applied, whereas several instantaneous large errors appear when filtering was not applied. In terms of improving prediction performance, we observed that our method could be useful in terms of avoiding false predictions due to false peak error values. However, large errors could be observed approximately half a month before the warning appeared in Fig. 7. In this case, our proposed method outperformed the AE without filtering in terms of both improving the genuine peak and suppressing the false peak. This is because we can limit the wind turbine conditions to be learned by the AE. These results indicate that our filtering method is useful for anomaly prediction for critical components of wind turbines with an AE and vibration data.

7. Conclusion

In this letter, we proposed a power-curve filtering method combined with an AE-based anomaly prediction technique that uses vibration and SCADA data. Our performance evaluation using actual data demonstrated that the proposed AE method for predicting a wind turbine malfunction has a possibility to perform better than an AE method without filtering.

References

[1] K. Kim, G. Parthasarathy, O. Uluyol, W. Foslien, S. Sheng, and P. Fleming, “Use of SCADA data for failure detection in wind turbines,” Proc. Energy Sustainability Conference and Fuel Cell Conference, Aug. 2011.

[2] H.H. Yang, M.L. Huang, and S.W. Yang, “Integrating auto-associative neural networks with Hotelling’s $T^2$ control charts for wind turbine fault detection,” Energies, vol.8, no.10, pp.12100–12115, 2015.

[3] L. Wang, Z. Zhang, J. Xu, and R. Liu, “Wind turbine blade breakage monitoring with deep autoencoders,” IEEE Trans. Smart Grid, vol.9, no.4, pp.2824–2833, July 2018.

[4] G. Jiang, P. Xie, H. He, and J. Yan, “Wind turbine Fault detection using a denoising autoencoder with temporal information,” IEEE/ASME Trans. Mechatronics, vol.23, no.1, pp.89–100, Feb. 2018.

[5] J. Ogata and M. Murakawa, “Vibration-based anomaly detection using FLAC features for wind turbine condition monitoring,” Proc. 8th European Workshop on Structural Health Monitoring (EW-SHM), vol.3, pp.1952–1961, Aug. 2016.

[6] T. Hasegawa, J. Ogata, M. Murakawa, T. Kobayashi, and T. Ogawa, “Adaptive training of vibration-based anomaly detector for wind turbine condition monitoring,” Proc. IEEE Int. Conf. Prognostics and Health Management (ICPHM), June 2017.

[7] G. Jiang, H. He, P. Xie, and Y. Tan, “Stacked multilevel-denoising autoencoders: A new representation learning approach for wind turbine gearbox fault diagnosis,” IEEE Trans. Instrum. Meas., vol.66, no.9, pp.2391–2402, Sept. 2017.

[8] T. Hasegawa, J. Ogata, M. Murakawa, and T. Ogawa, “Tandem connectionist anomaly detection: Use of faulty vibration signals in feature representation learning,” Proc. IEEE Int. Conf. Prognostics and Health Management (ICPHM), June 2018.