LETTER

Anomaly Prediction for Wind Turbines Using an Autoencoder Based on Power-Curve Filtering

Masaki TAKANASHI(1), Member, Shu-ichi SATO(1), Kentaro INDO(1), Nozomu NISHIHARA(1), Hiroto ICHIKAWA(1), and Hirohisa WATANABE(1), Nonmembers

SUMMARY Predicting the malfunction timing of wind turbines is essential for maintaining the high profitability of the wind power generation business. Machine learning methods have been studied using 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 have attracted significant interest in the detection or prediction of anomalies through unsupervised learning, in which the anomaly pattern is unknown. Although autoencoder-based techniques have been proven to detect anomalies effectively using relatively stable SCADA data, they perform poorly in the case of deteriorated SCADA data. In this letter, we propose a power-curve filtering method, which is a preprocessing technique used before the application of an autoencoder-based technique, to mitigate the dirtiness of SCADA data and improve the prediction performance of wind turbine degradation. We have evaluated its performance using SCADA data obtained from a real wind-farm.

key words: anomaly detection, autoencoder, power-curve filtering

1. Introduction

Wind energy power generation technologies have been studied and applied widely as one of the most important renewable energy sources that are very environment-friendly. One of crucial objectives of wind-farm operation is to minimize the downtime caused by main component failures. Once a main component fails, it may take much time to resume operations as a result of the component’s long lead time and its transportation to a distant location, which can cause significant losses in production. Therefore, one must prepare the component for replacement in advance of the failure by predicting the timing of the component’s failure.

Under these circumstances, many studies on the detection of anomalies in wind turbines have been conducted. They are largely categorized into two major approaches: the model-based approach and the data-driven approach. In the data-driven approach, machine learning methods are employed using condition monitoring system data, such as vibration data, and supervisory control and data acquisition (SCADA) data. Gaussian mixture models and their combination with deep neural networks were studied to detect anomalies in the vibration data of wind turbines [1]–[3].

Autoencoder (AE)-based techniques for detecting anomalies in SCADA data have also been studied and have demonstrated better detection performances [4]–[6]. The main advantage of an AE is that it can be applied through unsupervised learning because abundant anomaly data of wind turbines are not usually available.

Although AEs have been proven to be beneficial in the detection of anomalies in wind turbines, they are limited to the case of relatively stable SCADA data plotted close to the ideal data which can be determined by the specification of wind turbines. To the best of our knowledge, studies involving deteriorated SCADA data, which are plotted far from the ideal data, are yet to be reported. A major reason for the deterioration is that some sensors employed into turbines continue to measure values even though the operation of wind turbines is intently suspended for some operational reasons. This can cause the relationships among the variables to degrade. As a result of data degradation, AE-based detection suffers from degradation of its performance in anomaly detection. Therefore, we need to suppress the data degradation to maintain the detection performance. In this study, we propose a filtering preprocess based on power-curve filtering to alleviate this problem. Power-curve filtering comprises two variables: wind speed and active power, which have a typical relationship in each wind turbine. Power-curve filtering eliminates variables that are significantly far from the typical relationship.

In this study, we briefly introduce an AE for anomaly detection and the SCADA data used in this study in the following sections. Next, we propose a power-curve filtering method and demonstrate its potential to predict the deterioration of a turbine through evaluation using obtained SCADA data.

2. Anomaly Detection Using an AE

In this section, we briefly describe anomaly detection using an AE. A typical AE is shown in Fig. 1. An AE is a type of neural network that inputs multivariate data and has the same number of multivariate outputs. Input data are propagated with the network coefficients, \( \mathbf{W}_i \), between two layers and an activation function, \( f(\cdot) \), as follows:

\[
\text{AE}(\mathbf{x}) = f(\mathbf{W}_i f(\mathbf{W}_i f(\mathbf{x})))
\]
where \( h_i \) denotes the values at the \( i \)-th hidden layer. A non-linear function such as rectified linear unit (ReLU), sigmoid function and hyperbolic tangent function is usually employed as the activation function. This non-linear functionality is necessary to employ multiple hidden layers in the network to deal with the complicated non-linear characteristic of the input data. The network coefficients should be obtained to minimize the difference between the inputs and the outputs, as follows:

\[
\text{minimize} \| x - \hat{x} \|^2 \tag{2}
\]

When the AE has \( I \) hidden layers, the input variables, \( x \), and the output variables, \( \hat{x} \), satisfy the following:

\[
x = h_0 \tag{3}
\]

\[
\hat{x} = W_I h_I \tag{4}
\]

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

3. SCADA

Modern wind turbines install standard SCADA systems to capture operational status data. These data include meteorological conditions, such as wind speed and direction, component temperatures, and electrical measurements, such as generated power. These data are quite beneficial in the analysis of the current health condition of wind turbines, and they can be useful in anomaly detection using a data-driven approach. However, some of these data could be relatively deteriorated as a result of deriving representative values, such as average and maximum, for over 10 min to reduce the overall volume of the data to be sent.

Let us take wind speed and active generated power for example. The operation of wind turbines is intentionally suspended because of some operational reasons such as maintenance, preventive measure for predicted harsh weather, total power generation management inside an entire wind-farm and so on. Meanwhile, the wind sensor measures the wind speed even though the wind turbine is suspended. This causes the relationships among the multivariate values to deteriorate, which results in the deterioration of anomaly detection performance. Figure 2 presents a relationship between wind speed and active generated power, which is called a power-curve, obtained from one of our wind-farms through 10-min averaging. In the figure, the ideal power-curve is represented by a solid curve. This curve is separated into three stages: no power generation in low wind speed, linear power generation in middle speed, and constant power generation in high speed. However, the actual power-curve deviates from the ideal power-curve. The dense points on the x-axis can be obtained in the case where there are no power generations even though the wind blows over 10 minutes. The dense points distributed among wind speed [15.0, 20.0] with active power [200,800] can be obtained in the case where power generations are suspended in the middle of 10 min while the wind continues to blow. Because this SCADA system derives the averaged values within 10 min, the relationships between the obtained values degrade as a result of the mixture of operating and stopping times. Such a complicated pattern has detrimental impacts on the AE for capturing the features of the obtained SCADA data compactly, thus deteriorating the performance of anomaly detection and prediction.

4. Power-Curve Filtering

In this study, we propose a filter preprocessing method based on the power-curve to reduce the effect of noisy data. As a
prerequisite, wind speed and active power need to be included in SCADA data to apply this filtering technique. Here, we set particular filtering ranges for wind speed and active power. If the combination data of wind speed and active power in a certain snapshot are within the ranges, these values and the corresponding operational data are used for anomaly detection using the AE. Otherwise, they are left out. Figure 3 shows an example of a schematic image of the filter processing method, where we have active power, wind speed, and nacelle temperature. The range contains only the linear power generation area in the power-curve. The black circles shown on the left side of the figure represent the observed data. The ones shown on the right side of the figure are applied to the AE, and the white ones are excluded. This is the power-curve filtering preprocessing method for alleviating the effect of the deteriorated SCADA data.

5. Performance Evaluation

This section describes anomaly prediction using the AE when the power-curve filtering method is applied in the obtained SCADA data. In this evaluation, we employ an AE whose number of layers is five. The number of nodes of the input and output layers, respectively, is 29. The variables used for detection are shown in Table 1. The number of nodes of the first, second, and third hidden layers are 15, 9, and 15, respectively. The ReLU function is applied as the activation function in the AE. We train the AE using the first one-year SCADA data and attempt to predict anomalies for the following ten-year data. A malfunction occurs in June year Y+5 (Y+5/6).

We present the results obtained without and with power-curve filtering in Figs. 4 and 5, respectively. We apply a filter whose ranges are 5 to 12 in wind speed and 0.1 to 1400 in active power. These areas were determined to catch the data associated with the linear area in the power-curve, where the wind turbine is considered to be functioning consistently. In this study, we simply evaluate the prediction performance using a score that is equivalent to the square difference between the input and the output of the AE, as follows:

\[ score = ||x - \hat{x}||^2. \]  

The score value can be obtained every 10 min. We derive a median value from the score values obtained on each day, and a three-day moving average is then applied and plotted in the figures.

A large score appears after around Y+5/6, as shown in Fig. 4. This is because some variables drastically changed after the malfunction. However, we cannot observe an increase in the score just before the malfunction date. In contrast, even though the large score seen in Fig. 4 disappears, a gradual increase in the score can be observed before the malfunction date shown in Fig. 5. This indicates that small changes can be identified by removing noisy effects in the

| Table 1 | Variables used for prediction. |
|---------|-------------------------------|
| Active power | Yaw angle | Wind speed |
| GenTermVolt | ACVolt | PF |
| GenCurrent | PitchDem | WindDiff |
| PitchDeg | UseKw | Lubricate oil pressure |
| RPM | WindSpeed1 | WindSpeed10 |
| GenTemp4 | GenTemp6 | GenTemp10p |
| GenTempRev | HighSpTemp | AmbientTemp |
| NacellTemp | MaxWindSpeed | SGateDdeg |
| WindDir10 | Servo | RotorRpm |
| GenOutput1 | GenOutput10 | |

Fig. 3 Schematic image of the filter processing method.

Fig. 4 Reconstruction error without the filter.

Fig. 5 Reconstruction error with the filter.
To clarify which variable contributes to this increase, we derived the score for each variable. Figure 6 shows the score for lubricate oil pressure, which mostly contributed to the increase in the overall score. As seen in the figure, the error slightly increases before $Y+5/6$. In Fig. 7, we also show the raw values of lubricate oil pressure input into and output from the AE. In this case, the maximum of the input value is normalized as 1. The input value gradually decreases before the malfunction date, whereas the output value constantly remains at approximately 0.8. The output value can be regarded as the value estimated using the other variables. Because the input value degrades from the estimated value, a component malfunction related to this variable is considered to occur. These results indicate that our filtering method is helpful in the prediction of malfunctions.

6. Conclusion

In this study, we proposed a power-curve filtering method used with an AE-based anomaly detection technique to reduce the effect of degraded SCADA data. Through performance evaluation using actual collected data, we demonstrated the prediction of malfunctions in wind turbines by eliminating the degraded data using the proposed method.

References

[1] 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.
[2] 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.
[3] 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.
[4] H.H. Yang, M.L. Huang, and S.W. Yang, “Integrating auto-associative neural networks with Hotelling $T^2$ control charts for wind turbine fault detection,” Energies, vol.8, no.10, pp.12100–12115, 2015.
[5] 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.
[6] 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 Feb. 2018.