International Core Journal of Engineering Volume 5 Issue 11, 2019
ISSN: 2414-1895 DOI: 10.6919/ICJE.201910_5(11).0029

Wind Turbine Fault Diagnosis Based on Compression Acquisition and Depth Confidence Network

Chenggang Zhen [a], Weijia Zhao [b]

School of control and computer engineering, North China Electric Power University, Baoding, 071000, China;

[a]251811399@qq.com, [b]zhaowei_jia@163.com

Abstract

Considering the large amount of condition monitoring data in wind turbine data acquisition system, a deep learning method based on compressed sensing for deep confidence network is proposed. Firstly, the compressed sensing theory is used to sample and process the original data, and then the processed data is trained through the deep confidence network to construct the deep confidence network model. Finally, the trained model is simulated and analyzed. The simulation results verify the validity of the deep confidence network learning method based on compressed sensing.

Keywords

Wind turbine, deep belief network (DBN), compressed sensing sampling , fault diagnosis.

1. Introduction

In recent years, the state has vigorously advocated clean energy, and wind energy plays an increasingly important part in clean energy. With the continuous expansion of the scale of wind farms, wind turbines are also facing a series of problems. Because the wind turbine has been in a bad environment for a long time, bearings, gears, rotors and other vulnerable components [2]. Therefore, this paper proposes a deep confidence network fault detection method based on compressed sensing, which reduces the amount of data by compressing the original data, and uses deep confidence network to realize fault classification and recognition. Experiments show that this method can effectively identify the fault types and improve the accuracy of fault classification.

2. Theoretical Basis

2.1 Compressed Sensing Theory

Compressive sensing theory was proposed by E.J. Candes, T. Tao, J. Romberg and Donoho in 2004. Once the theory of compressed sensing is put forward, it has a great impact on the field of signal processing. Compressed sensing theory is a new sampling principle. It uses random sampling to obtain discrete samples of signals on the premise that the sampling frequency is far less than the minimum sampling frequency stipulated by Nyquist sampling law, and then reconstructs the signals perfectly by non-linear reconstruction algorithm.

Compressive sensing theory refers to the theory that the signal is compressible or sparse in a transform domain. The transformed high-dimensional signal can be projected into a low-dimensional space by an observation matrix which is not related to the transform basis. Then the original signal can be reconstructed with high probability from these small projections by solving an optimization problem. It can be proved that such projection contains sufficient information of reconstructed signal.

Compressed sensing theory mainly includes three parts:

ISSN: 2414-1895 DOI: 10.6919/ICJE.201910_5(11).0029

First,sparse representation of signals。Generally, most signals have sparse characteristics. A dense signal can also be mapped to a sparse signal by a linear transformation matrix. Sparse representation of signals is to use as few atoms (or bases, columns of dictionary matrix Ψ) as possible to represent signals in a given dictionary matrix Ψ, so as to remove redundant information of signals. Sparse signal is more convenient for further processing, such as compression, coding and so on. The research hotspots of signal sparse representation mainly focus on sparse decomposition algorithm, over-complete dictionary and the application of sparse representation. Given a dictionary matrix Ψ, each column of $\{\psi_i\}_{i=1}^{N}$ is an atom. The signal $x \in R^N$ is represented by the atom of the dictionary matrix Ψ:

$$x = \sum_{n=1}^{N} s_n \psi_n = \Psi s \tag{1}$$

Among Ψ is an orthogonal matrix of N×N. $s \in R^N$ is a sparse signal, in which the number of non-zero elements K (sparsity) in S is much smaller than the signal length N, i.e. k≪N. In practical application, the value close to 0 can also be set to 0. The transform matrices used in sparse decomposition of signals usually include DFT, DCT and wavelet bases.

Second,the measurement matrix should be designed to ensure the minimum information loss of the original signal X at the same time of low dimension. Considering the noise-free situation, if x is a dense signal, the observation vector is expressed as $\mathrm{y} = \Phi x = \Phi\Psi s$. We call $\mathrm{A} = \Phi\Psi$ a sensing matrix or a sensing matrix. If x is a sparse signal, the observation vector is expressed as $y = \Phi x$, and then $A = \Phi$ is a sensing matrix, also known as an observation matrix.

Last, A signal recovery algorithm is designed to recover the original signal of length N without distortion from M observations. Signal reconstruction theory is the core of compressed sensing theory, considering the model:

$$y = Hx + w \tag{2}$$

Where $x \in R^N$ denotes sparse signals, $H \in \mathbb{R}^{M\times N}$ denotes observation matrices, and $y \in R^M$ denotes observation vectors. $w \in R^M$ is noise interference. Unlike other reconstruction theories, in compressed sensing, the length of the observation vector is much smaller than the length of the target signal, namely $M \ll N$. Since the target signal x is sparse, the sparsity is set to K (the number of non-zero elements in the target signal), which correspondingly is equivalent to solving the $M \times k$ equations. Traditional reconstruction algorithms of compressed sensing theory include greedy iteration algorithm, matching pursuit algorithm, orthogonal matching pursuit, subspace pursuit and so on.

2.2 Deep Confidence Network

Deep Belief Nets (DBN) is a kind of neural network. It can be used not only in unsupervised learning, similar to a self-coding machine, but also in supervised learning as a classifier[4]. As a neural network, neurons are an indispensable part of it. DBN consists of several layers of neurons, and its components are restricted Boltzmann machine (RBM).

RBM is a kind of nerve perceptron, which consists of a dominant layer and a recessive layer. The neurons in the dominant layer and the recessive layer are fully connected in two directions. As shown in Figure 1:

ISSN: 2414-1895 DOI: 10.6919/ICJE.201910_5(11).0029

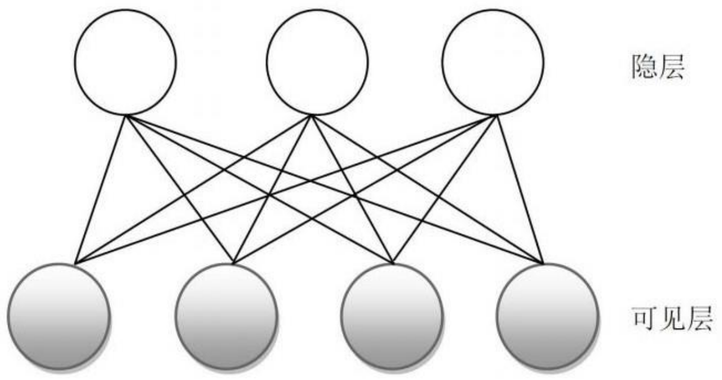

Figure 1. DBN structure

RBM can be expressed as a bipartite graph model. All the nodes in the visible layer and the hidden layer have two states: 1 in the active state and 0 in the inactive state. The meanings of 0 and 1 states here represent which nodes the model will select to use, which nodes are in active state are used, and which nodes are not in active state are not used. The activation probability of the nodes is calculated by the distribution function of the nodes in the visible layer and the hidden layer.

In RBM, *v* represents all visible units and *h* represents all hidden units. To determine the model, only three parameters of the model $\theta = \{W, A, B\}$, can be obtained. Weight matrix *W*, visible layer element offset *A*, hidden layer element offset *B*.

Assuming that a RBM has *n* visible units and *m* hidden units, v_i is used to represent the *i* visible unit, h_j is used to represent the *j* hidden unit, and its parameter form is as follows:

$W = \{w_{i,j} \in R^{n\times m}\}$,$W_{i,j}$ denotes the weight between the visible unit *i* and the hidden unit *j*.

$A = \{a_i \in R^m\}$, Where a_i represents the bias threshold of the *i* visible unit.

$B = \{b_j \in R^n\}$, Where b_j represents the offset threshold of the *j* visible unit.

For a given set of *(v, h)* values, assuming that both visible and hidden layer elements obey Bernoulli distribution, the energy formula of RBM is as follows:

$$E(v,h \mid \theta) = -\sum_{i=1}^{n} a_i v_i - \sum_{j=1}^{m} b_j h_j - \sum_{i=1}^{n}\sum_{j=1}^{m} v_i W_{i,j} h_j \quad (3)$$

Among them, $\theta = \{W_{i.j}, a_i, b_j\}$ are the parameters of RBM model. The energy function indicates that there is an energy value between the values of each visible node and the values of each hidden layer node.

After exponentiation and regularization of the energy function, the joint probability distribution formulas of the set of nodes in the visible layer and the set of nodes in the hidden layer under a certain state *(v, h)* can be obtained.

$$P(v,h \mid \theta) = \frac{e^{-E(v,h|\theta)}}{Z(\theta)} \quad (4)$$

$$Z(\theta) = \sum_{v,h} e^{-E(v,h|\theta)} \quad (5)$$

$Z(\theta)$ is a normalization factor or partition function, which represents the summation of the energy exponents of all possible states of the set of nodes in the visible layer and the hidden layer.

Likelihood function is often used to derive the parameters. The joint probability distribution $P(v,h \mid \theta)$ is known. By summing all the states of the hidden layer node set, the edge distribution $P(v \mid \theta)$ of the visible layer node set can be obtained.

ISSN: 2414-1895 DOI: 10.6919/ICJE.201910_5(11).0029

$$P(v \mid \theta) = \frac{1}{Z(\theta)} \sum_h e^{-E(v,h|\theta)} \tag{6}$$

Edge distribution represents the probability that the set of nodes in the visible layer is in a certain state distribution.

The fault diagnosis flow of DBN is shown in Figure 2:

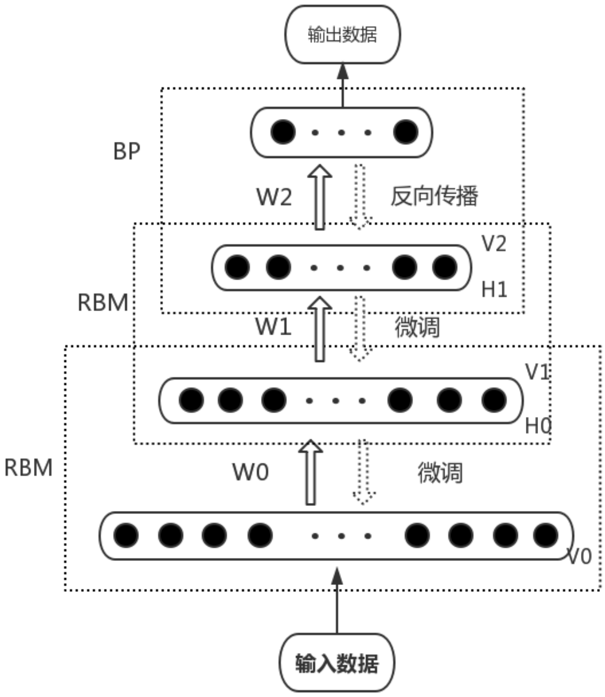

Figure 2. DBN Fault Diagnosis Process

Firstly, each layer of RBM network is trained independently and unsupervised to ensure that the feature information is preserved as much as possible when the feature vectors are mapped to different feature spaces; (H0 can be regarded as the visible layer of H1).

In the last layer of DBN, BP network is set up to receive the output eigenvector of RBM as its input eigenvector, and the entity relationship classifier is trained supervisively.

Each layer of RBM network can only ensure that the weights in its own layer are optimal for the mapping of the layer's eigenvectors, not for the whole DBN's eigenvectors. Therefore, the back propagation network also propagates error information from top to bottom to each layer of RBM, and fine-tunes a DBN network.

The process of training model of RBM network can be regarded as the initialization of weights of a deep BP network, which makes DBN overcome the shortcomings of BP network which is easy to fall into local optimum because of randomly initializing weights and long training time. This can be explained intuitively. The BP algorithm of DBNs only needs a part of the weights parameter space. Compared with the forward neural network, the training speed is faster and the convergence speed is faster.

3. Experiments and Results

Because of the shortcomings of large data and traditional methods in fan health detection system, this paper presents a method of data compression sensing signal acquisition to realize the compression acquisition of fault data in SCADA system. Meanwhile, a deep neural network based on deep confidence network is constructed to realize fault self-realization by learning the original data. Motion extraction and intelligent diagnosis of health status.

ISSN: 2414-1895 DOI: 10.6919/ICJE.201910_5(11).0029

It is found that the vibration signal of the fan is sparse, and the random Gaussian matrix can be used to realize the projection in the transform domain, which greatly reduces the redundancy of the original data. The compressed data becomes more valuable and can be input into the deep confidence network for in-depth learning.

After in-depth learning of compressed data, all parameters of DBN are initialized by layered pre-training of a large number of unlabeled data. Finally, after using BP algorithm to optimize, it can converge to the local extreme point more quickly, and realize the intelligent diagnosis of mechanical faults.

4. Conclusion

This paper presents a new method for fault diagnosis of wind turbines. Based on the full data-driven diagnosis algorithm, this method effectively utilizes the data in SCADA system, and does not need to model the fan in advance. Experiments show that the method is more effective than other algorithms in the detection of fan faults. The research work in this paper contributes to further verifying the possibility of complete data-driven diagnostic methods.

References

[1] Li Mengshi, Yu Da, Chen Ziming, Xia Hou Kaishun, Li Juyan, Ji Tianyao. Fault Diagnosis Method for Wind Turbine Generators Based on Deep Confidence Network [J].Journal of Electrical Machinery and Control, 2019, 23 (02): 114-122.

[2] Zheng Xiaoxia, Chen Guangning, Ren Haohan, Li Dongdong. Fault warning of fan vulnerable components based on improved VMD and deep confidence network [J]. Vibration and shock, 2019, 38 (08): 153-160+179.

[3] DONOHO D L. Compressed sensing ［J］. IEEETransactions on Information Theory, 2006, 52 (4) : 1289- 1306.

[4] Chen Baojia, Liu Haotao, Xu Chao, Chen Fafa, Xiao Wenrong, Zhao Chunhua. Application of Deep Confidence Network in Gear Fault Diagnosis [J]. China Mechanical Engineering, 2019, 30 (02): 205-211.