Power Grid Fault Diagnosis Based on Improved Deep Belief Network

Yihui Zhang¹, Yuansheng Zhang², Libin Wen¹, Zhimei Cui¹, Yini He² and Guangshi Liu¹

1 Guangxi Power Grid Electric Power Research Institute, Nanning 530023, China
2 Dispatch and Control Center of Guangxi Power Grid, Nanning 530023, China
2782259972@qq.com

Abstract. It is of great significance to quickly and accurately identify power system faults. This paper introduces the idea of deep learning into power system fault diagnosis research and proposes a fault diagnosis model based on improved deep confidence network. Construct a set of 30-dimensional original features that can reflect the fault characteristics of the power system as the model input, and the fault diagnosis result is the model output. Use multi-layer Boltzmann machines to train the mapping relationship between grid faults and system features. Finally, the extreme learning machine is used to supervise the labeled samples to modify the network parameters. Different system failure scenarios were set up on the New England 10-machine 39-node system to test the diagnosis ability of the model. Simulation results show that the improved deep belief network has a strong feature extraction capability. The improved deep belief network has higher accuracy and faster speed in fault categories, fault areas, and fault locations than common artificial intelligence methods.

1. Introduction

With the increasing complexity of power system topology and operation, fast and accurate power grid fault diagnosis has been one of the research focuses[1]. At present, fault diagnosis methods can be roughly divided into three categories which are traveling wave technology, based on impedance and artificial intelligence applications. The methods based on the traveling wave theory is to analyze the fault by analyzing the position and time diagram of the current or voltage movement, which include wavelet transform [2], FFT [3] and other methods. The methods based on impedance measurement are divided into time domain and phasor domain. In the methods using the phasor domain [4], the purpose of fault diagnosis is to calculate through current and voltage data or phasors. The time domain method [5] is based on the establishment of differential equations for transmission lines. These methods are limited to the characteristics of the transmission line or sampling frequency of the data.

In recent years, research on power system fault diagnosis based on machine learning have made preliminary progress. Neural networks [6], support vector machines [7] and random forest methods [8] are common methods in fault diagnosis. These traditional machine learning methods are correct in some scenarios but it cannot extract the characteristic quantity of the power grid well, so it cannot be widely used. Compared with traditional machine learning methods, deep learning can mine deep and complex associations in the data. At present, there are few related researches on power system fault diagnosis using deep learning methods.

On the other hand, existing researches usually is usually limited to fault categories or fault location. Aiming at the problem, this paper proposes a power grid fault diagnosis model based on improved
deep belief network. Different system fault scenarios were set up on the New England 10-machine 39-node system to test the model’s diagnosis ability. Simulation results show that the improved deep belief network has higher accuracy and faster speed in fault categories, fault areas, and fault locations than common artificial intelligence methods.

2. Improved Deep Belief Network
The Deep Belief Network (DBN) [9] consists of an additional top-level supervised back-propagation neural network and a multi-level unsupervised Restricted Boltzmann Machine (RBM). RBM is a two-layer bidirectional recursive neural network, which consists of a layer of visible units $v$ representing input and a layer of hidden units $h$ representing hidden variables. Its basic structure is shown in Figure 1.

At the same time, RBM is an energy-based model. When the visible $v$ and hidden $h$ units are binary variables with states 0 and 1, the energy function of the unit is:

$$E(v, h) = - \sum_{i} \sum_{j} v_i w_{ij} h_j - \sum_{i} c_i v_i - \sum_{j} b_j h_j$$  \hspace{1cm} (1)

Where $w_{ij}$ is the connection weight of the first visible unit $i$ and the second hidden unit $j$, $v_i, h_j$ is the state of the visible unit $i$ and the hidden unit $j$, and $c_i$ and $b_j$ are the offset of the visible unit $i$ and the hidden unit $j$. To train RBM, get the parameters $w, c, b$ from the training data. The state probability of RBM is determined by its energy function, which is:

$$P(v, h) = \frac{1}{Z} \exp(-E(v, h))$$  \hspace{1cm} (2)

Where $Z$ is the normalization factor, also called the partition function.

The activation probabilities of visible and hidden units are:

$$P(h_j = 1) = \sigma(b_j + \sum_i w_{ij} v_i)$$  \hspace{1cm} (3)

$$P(v_i = 1) = \sigma(a_i + \sum_j w_{ij} h_j)$$  \hspace{1cm} (4)

Where $\sigma$ is the activation function, this paper chooses sigmoid activation function.

In traditional DBN, the RBM is pre-trained layer by layer, and the output of the hidden layer of the RBM is the input of the visible layer of the next RBM. In the tuning stage, the last layer of the network is trained in a supervised manner, the difference between the actual output and the expected output is estimated, and back propagation is made layer by layer, which takes a long time and easily falls into a local optimum. This paper replaces the traditional back-propagation neural network with an ELM classifier at the top level of the DBN, which is more efficient than traditional multilayer neural networks. The structure of improved Deep Belief Network is shown in Figure 2.

In the last layer of the ELM network, when the sample input is $X = [x_1, x_2, x_3, ..., x_n]$ and the sample output is $T = [t_1, t_2, t_3, ..., t_n]$. The goal is to minimize the output error, which is:

$$obj: e = \min \sum_{j} \beta \sigma(w_j x_j + b) - t_j$$  \hspace{1cm} (5)

Use the least squares method to calculate the output weight, that is:

$$\beta = (WX + B)^+T$$  \hspace{1cm} (6)

\[\text{(Figure 1. The structural of RBM.)}\]

\[\text{(Figure 2. The structure of improved Deep Belief Network.)}\]
Where \((WX + B)^+\) is the Moore-Penrose generalized inverse of \(WX + B\). Huang proposed the ELM with equality optimization constraint [10], which can not only minimize the training error, but also minimize the output weight. The objective function containing the equality optimization constraints can be expressed as:

\[
\text{obj} : \min \frac{1}{2} \| \beta \|^2 + \frac{1}{2} C \sum_{i=1}^{n} e_i^2
\]  

(7)

\[
\beta = (WX + B)^T \left( \frac{1}{C} + (WX + B)(WX + B)^T \right)^{-1} T
\]  

(8)

Where \(C\) is a penalty parameter, and the Lagrange method is used to solve the objective function into an unconditional optimization problem.

3. Fault diagnosis model of power grid based on Improved DBN

3.1. Data generation and feature extraction

One of the key steps in the power system fault diagnosis using deep learning models is data generation and feature extraction. In the power system time domain simulation software BPA program, inputing grid system parameters and setting fault conditions to obtain the original data set and classify. In the selection of the feature, the feature must meet the requirements that can correctly reflect the system fault status. According to relevant literature and time-domain simulation results, this paper selects 30 original features, as shown in the Table 1. Where \(T_0\), \(T_f\), and \(T_{cl}\) represent the normal operating time, the time when the fault occurs, and the time when the fault is removed.

| Number | Features |
|--------|----------|
| F1-F6  | System active load, reactive load, active output, reactive output, average bus voltage, generator maximum power angle difference at \(T_0\) |
| F7-F11 | The maximum, minimum, variance, root mean square error of rotor acceleration and maximum acceleration generator rotor angle at \(T_f\) |
| F12-F14| The maximum, minimum, average value of generator acceleration power at \(T_f\) |
| F15-F18| Maximum angular difference, angular velocity difference, rotor maximum kinetic energy value, maximum kinetic energy difference at \(T_f\) |
| F19-F26| Maximum angle difference, angular velocity difference, angular acceleration difference, kinetic energy difference, rotor angular increment, speed deviation, angular velocity increment at \(T_{cl}\) |
| F27-F30| average kinetic energy, Maximum kinetic energy and its rotor angle, maximum rotor angle and its kinetic energy at \(T_{cl}\) |
### 3.2. Model diagnosis process

In this paper, an improved DBN model is used to identify faults. The specific process is shown in the Figure 3. It mainly includes three parts: data set generation, diagnosis model training and model performance evaluation.

### 3.3. Model performance evaluation

In order to evaluate the evaluation results more accurately, four indicators including the miss alarm rate (MAR), false alarm rate (FAR) and accuracy (ACC) are introduced to evaluate the results of the system evaluation:

\[
MAR = \frac{FP}{FP + TN} \tag{9}
\]

\[
FAR = \frac{FN}{TP + FN} \tag{10}
\]

\[
ACC = \frac{TP}{TP + FP + FN + TN} \tag{11}
\]

Where FN represents the number of samples where the actual system is correct but the model evaluates as incorrect. The concepts of TP, FP, and TN are similar to FN.

It takes too long to train and test the recognition model, which is why it cannot be widely used. This article uses the evaluation time-consuming index $T_{sum}$ evaluation model to evaluate the speed, which is the total time for applying model training and testing samples.

![Figure 3. The structural of improved DBN.](image)

4. Case study results and discussion

#### 4.1. Fault area diagnosis sample set

This paper uses the New England 10-machine 39-node system as the example system for time-domain simulation. The system includes a total of 10 generators and 46 lines. In order to reflect the real and variable power system operating state, nine load levels of 0.8Pn, 0.85Pn, ..., 1.15Pn, 1.2Pn are simulated, and the generator output is adjusted accordingly to ensure system power balance and voltage stability. Set fault conditions for the transmission line. The fault locations are located at 0%, 20%, 50%, and 80% of the line. The fault category is three-phase short-circuit fault. The fault duration is 0.1s, 0.16s, 0.22s, and the simulation software is PSD-BPA. According to the example system and fault conditions, a total of 4,968 valid samples were obtained. In order to explore the model’s ability to identify fault areas, the system is divided into four areas, as shown in the Figure 4.

#### 4.2. Fault categories diagnosis sample set

The New England 10-machine 39-node system is still used as the example system for time-domain simulation. Under nine different load levels, 4 categories of faults (LG fault, LL fault, LLG fault, LLL
fault) are set on the transmission line, where letter L stands for line and letter G stands for ground. The fault locations are located at 20%, 50%, and 80% of the line, respectively, and the fault durations are 0.1s, 0.16s, and 0.22s. To ensure sample balance, randomly select 3000 samples for each of faults categories and the total sample set has 12,000 samples.

4.3. Fault location diagnosis sample set

The lines BUS1-BUS2 in the system are selected for simulation. The fault locations are located at 0%, 1%, and 2% ..., 98%, 99% of the line. The fault categories is three-phase short-circuit fault, and the duration of the fault is 0.05s, 0.1s, 0.15s, ..., 0.35s, 0.4s. A total of 7,200 samples were obtained and divided into four equal parts.

4.4. Results and discussion

| Model      | Fault location diagnosis | Fault categories diagnosis | Fault area diagnosis |
|------------|--------------------------|-----------------------------|----------------------|
|            | MAR(%) FAR(%) ACC(%)     | MAR(%) FAR(%) ACC(%)        | MAR(%) FAR(%) ACC(%) T_{sum}(s) |
| BPNN       | 1.79 5.35 94.57          | 3.05 9.23 90.79             | 5.58 17.65 82.44 48.95 |
| DT         | 1.68 5.06 94.92          | 2.08 6.25 93.75             | 4.63 13.92 86.11 30.84 |
| SVM        | 0.45 1.35 98.64          | 0.73 2.16 97.84             | 1.55 4.79 95.33 522.49 |
| KNN        | 0.76 2.31 97.71          | 0.65 1.96 98.04             | 3.65 11.05 89.00 1.82  |
| ELM        | 0.76 2.35 97.64          | 0.87 2.62 97.38             | 1.45 4.45 95.56 0.4023 |
| DBN        | 0.38 1.17 98.85          | 0.61 1.84 98.16             | 1.79 5.89 94.67 427.31 |
| improved DBN | 0.31 0.93 99.07        | 0.53 1.59 98.41             | 0.55 1.83 98.33 11.20 |

Back propagation neural network (BPNN), decision tree (DT), support vector machine (SVM), K-nearest neighbor (KNN), DBN, ELM, and improved DBN are applied on the sample set. The test PC was configured as an Intel Core i7-8700 CPU. The simulation results are shown in the Table 2. From the comparison of simulation results based on examples, the following conclusions can be drawn:

1) The accuracy of the improved DBN model is higher than that of DBN and ELM, and the total evaluation time is between DBN and ELM. This is because the traditional DBN model uses a layer of feedforward neural network as a classifier, which needs to repeatedly fine-tune the parameters to reduce errors and increase the training time. The ELM model has a single hidden layer structure, and its feature extraction capability does not perform as well as the DBN model.

2) Traditional machine learning models (BPNN, DT, SVM) are simple in structure and show a certain classification effect, but the evaluation results are worse than the DBN model. Especially in the diagnosis of the fault area, the diagnosis effect of the improved DBN model is much better than other models. The evaluation result of the improved DBN model is significantly better than the traditional machine learning model.

4.5. Evaluation of model feature extraction capability

In this paper, the visual tool t-SNE is used to visually verify the improved DBN model for fault diagnosis and feature extraction. T-SNE converts the distance between data points into a probability. It is one of the best visualization methods at present.

Taking the fault diagnosis data sample set as an example, the Figure 5 is the original sample space, and the improved DBN model feature learning sample space is shown in the Figure 6. It can be seen that in the original sample space, there is a large amount of mixture and overlap of each categories of sample, which is difficult to classify. However, the sample space after learning of the improved DBN model has basically classified the data set, and the classification requirements can be met by applying a simple classifier, showing the improved feature extraction and learning capabilities of the improved DBN.
5. Conclusion
This paper proposes a power system fault diagnosis model based on improved DBN, which combines the feature extraction capabilities of DBN with the computational advantages of ELM. Finally, the model performance was tested in the New England 10-machine 39-node system, and the t-SNE tool was used to demonstrate model feature learning capabilities. According to the simulation results, it can be seen that compared with traditional machine learning, the improved DBN model has a lower false positive rate and false negative rate in the diagnosis of fault areas, fault categories, and fault locations, ensuring a higher accuracy and the recognition speed shows excellent feature learning ability. The results verify the correctness and effectiveness of the improved DBN model. The focus of future research is to explore the application of deep learning considering complex and variable real power grids.

6. References
[1] Aleem S A, Shahid N and Naqvi I H 2015 Methodologies in power systems fault detection and diagnosis Energy Systems 6(1):85-108
[2] Shaik A G and Pulipaka R R V 2014 A new wavelet based fault detection, classification and location in transmission lines International Journal of Electrical Power & Energy Systems 64(2015):35-40
[3] Mamis M S, Arkan M and Keles C 2013 Transmission lines fault location using transient signal spectrum International Journal of Electrical Power & Energy Systems 53(2013):714-718
[4] Izykowski J, Molag R, Rosolowski E and Saha M M 2006 Accurate location of faults on power transmission lines with use of two-end unsynchronized measurements IEEE Transactions on Power Delivery 21(1):627-633
[5] Suonan J, Gao S, Song G, Jiao Z and Kang X A 2010 novel fault-location method for HVDC transmission lines IEEE Transactions on Power Delivery 25(2):1203-1209
[6] Gururajapathy S S, Mokhlis H and Illias H A 2017 Fault location and detection techniques in power distribution systems with distributed generation: A review Renewable and Sustainable Energy Reviews 74(2017):949-958
[7] Livani H and Evrenosoglu C Y 2014 A machine learning and wavelet-based fault location method for hybrid transmission lines IEEE Transactions on Smart Grid 5(1):51-59
[8] Samantaray S R 2013 A systematic fuzzy rule based approach for fault classification in transmission lines Applied Soft Computing 13(2013):928-938
[9] Zhang C, He Y, Yuan L and Xiang S 2018 Analog Circuit Incipient Fault Diagnosis Method Using DBN Based Features Extraction IEEE Access 6(2018)23053-23064
[10] Huang G, Wang D and Lan Y 2011 Extreme learning machines: a survey International Journal of Machine Learning and Cybernetics 2(2):107-122

Acknowledgments
This work was supported by Guangxi Power Grid Electric Power Research Institute’s technology project 0470002019030103DL00005.