An Intelligent Per-decision Method of State Grid Accidents in Big Data Environments

Dianyang Li¹, Chen Tan²,³, Hongzhe Wang¹, Da Han²,³, Ling Qin¹, GuangCheng Zhang²,³, Tianyue Lou²,³, Tian Xiu²,³

¹State Grid Liaoning Electric Power Supply Co., Ltd., Liaoning 110006, China
²Nari Group Corporation/State Grid Electric Power Research Institute, NanJing 211000, China
³BeiJing KeDong Electric Power Control System Co., Ltd., Beijing 100192, China

Abstract. Aiming at the problem of hard to per-alarm the state grid accidents in big data environment, this paper proposes an intelligent per-decision method of state grid accidents. This method divides the management of state grid data into two parts, i.e. analysis of state grid data and intelligent expert database. The part of analysis of state grid data consists of input layer, facility layer and accident layer. Establish the self-learning intelligent expert database by means of connecting weight values of input and output nodes and threshold values of output nodes. Compare with the per-data that comes from the part of analysis of state grid data and the data from intelligent expert database, then achieve the goal of per-decision of state grid accidents and dispatch the accidents. A case and simulation results show that this method is implementable. In addition, this method provides beneficial thinking to solve the state grid accidents per-alarm in big data environment.

1. Introduction

As global energy issues become more and more serious, SG (SG) has become a hot research topic in various countries [1]-[2]. The SG integrates modern and advanced Internet of Things technology, network technology, sensing and measurement technology, communication technology, computing technology, automation technology and intelligent control technology into the current physical grid to form a new SG. The SG establishes a panoramic perception network that completely covers the power generation, transmission, substation, power distribution, power consumption and dispatching of the power system. Therefore, the data generated in the grid operation and equipment detection grows exponentially and has gradually formed big data.

Big data is a hot topic in the information industry today. It is defined as the large amount of data that is too large to be obtained, managed, processed, and organized in a reasonable time by the current mainstream software tools for the positive purpose to help companies make business decision [3]-[5]. The business of big data in the SG is divided into three categories [6]: The first one is grid operation and equipment detection or monitoring data; the second is marketing data of power enterprises; the third is management data of power enterprises. This paper mainly studied the big data generated in grid operation and equipment detection or monitoring. The data generated in grid operation and equipment detection or monitoring is mainly in the form of telemetry and telesignalisation. The historical data of telemetry and telesignalisation from 2012 to the present has reached 14T. Fig. 1 shows the total amount of data since 2012, and according to the growth trend, the amount of data for 2016-2021 is obtained::
According to Fig. 1, by 2021, the total amount of data can reach 25T. There are many problems in analyzing and mining these huge data, such as large amount, various types, complex correlation, technical bottlenecks in analysis, insufficient utilization of data values, especially the difficult realization in real-time alarm through the data. Based on this, the article [7] proposed a cascading failure prediction method, which calculated the system power flow after the accident to determine the node voltage fluctuation and the overload of the circuit and to cut off the corresponding elements according to the determination; the article [8] put forward the use of two neural networks, one for the diagnosis of normal, overheating, partial discharge and arc faults, and the other for the diagnosis of faults involving cellulose degradation. There are still some such methods at home and abroad, but these methods do not consider the self-learning of SG data in the era of big data, and its predictive ability needs to be improved.

This paper mainly studied the SG data monitoring, equipment faults and event prejudgment under big data, proposed the monitoring and pre-judgment analysis algorithm for power grid events, established a self-learning intelligent expert database, and conducted pre-judgment and dispatching of SG events.

2. System structure model
The system was divided into two parts, the expert database and the grid data analyzer. The expert database was responsible for the data of the storage equipment failure and the occurrence of the grid event, and the grid data analyzer conducted the big data analysis. The specific structure is shown in Fig. 2:
The following section details the grid data analyzer:

Various detection devices distributed in the SG sent a huge amount of data of grid operation and equipment detection or monitoring to the data center. These data served as input signals X. After analyzing the data, various accident states of the power grid could be predicted, and these states were the output signal Y. When the power grid was running, different grid events were caused by different equipment faults, while different equipment faults were caused by changes of various monitoring and detection data. That was, the changes of various input monitoring data first affected the running state of the device, and the change of the running state of the device might generate different events. According to the above relationship and the trend of the grid data to big data, the data was processed as follows. The grid data analyzer was divided into three layers: one was the input layer, the second one was the device layer, and the third was the event layer. The numbers of nodes were \( n \), \( l \) and \( m \), respectively. \( n \) was the number of input signals, while \( l \) was the number of devices, and \( m \) was the number of grid events. The input layer performed data collection, collecting the device detection or monitoring data in the power grid; The device layer predicted the state of each device according to data from the input layer and the expert database, and outputted the device status information to the event layer; The event layer predicted the status of the event based on the data output by the device layer.

Where \( x_1, x_2, \ldots, x_n \) are the input nodes, indicating various monitoring signals in the power grid; \( h_1, h_2, \ldots, h_l \) are the processing nodes, \( h_1, h_2, \ldots, h_l \) are the operating states of various devices in the power grid; \( y_1, y_2, \ldots, y_m \) are the output nodes, indicating the states of different events in the power grid. The connection weight values of each node between the input layer and the device layer were set as \( q_{nx,h_l} \), while the connection weight values of each node between the device layer and the event layer were set \( q_{h_l,y_m} \), and the thresholds of the nodes of the device layer and the event layer were set \( a_{h_l} \) and \( b_{y_m} \) respectively. The sampling time was set as \( t \). The link direction of each node indicated the influence of the forward node on the backward node. The magnitude of the influence was determined by the connection weight values. The larger the weight value, the greater the influence of the forward node on the backward node. The threshold value represented the limits of the processing
node and the output node.

Self-learning intelligent expert stored the data of equipment failures and grid events, and compared them with the data obtained from the grid data analyzer, as described in Section 3.

3. Grid data analysis method

In the part of power grid equipment monitoring and operation analysis, an algorithm for power grid events monitoring and prejudgment analysis was proposed, which could realize the functions of analysis and prejudgment on equipment abnormality and grid operation event. The details are as follows.

3.1. Equipment abnormal analysis and prejudgment

According to the current time input value $X_i^t$ of the smart grid big data, the connection weight values $q_{x_i,h_j}^t$ between the input layer and the device layer, and the threshold $a_{h_j}^t$ of the device layer, the predicted output $H_{h_j}^t$ of the device layer can be obtained, that is, the predicted value of the device operation state:

$$H_{h_j}^t = f(\sum_{i=1}^{n} q_{x_i,h_j}^t X_i^t - a_{h_j}^t)$$  \hspace{1cm} (1)$$

In the above formula, $l$ is the device number of the device layer, that is $j = 1, 2, \ldots, l$, $f(x)$ is the excitation function of the device layer, whose function is to convert the data into a non-dimensional signal:

$$f(x) = \frac{1}{1 + e^{-x}}$$  \hspace{1cm} (2)$$

The connection weight values $q_{x_i,h_j}^t$ between the input layer and the device layer, as well as the threshold values $a_{h_j}^t$ of the device layer, are described in detail in Section 3.4.

3.2. Analysis and prejudgment of grid operation event

According to the predicted output $H_{h_j}^t$ of the device layer, the connection weight value $q_{h_j,y_k}^t$ between the device layer and the event layer, and the threshold $b_{y_k}^t$ of the event layer calculated in 2.1, and the predicted output $Y_{y_k}^t$ of the grid event state can be obtained:

$$Y_{y_k}^t = \sum_{i=1}^{l} H_{h_j}^t q_{h_j,y_k}^t - b_{y_k}^t$$  \hspace{1cm} (3)$$

Where $k = 1, 2, \ldots, m$. The connection weight values $q_{h_j,y_k}^t$ between the device layer and the event layer, as well as the threshold $b_{y_k}^t$ of the event layer, are described in detail in Section 3.4.

3.3. Calculation of the error between the predicted value and the actual value

According to the actual output $O_{y_k}^t$ and the predicted output $Y_{y_k}^t$ obtained from 2.2, we could calculate the instantaneous error $e_{y_k}^t$:

$$e_{y_k}^t = Y_{y_k}^t - O_{y_k}^t \quad \text{where } k = 1, 2, \ldots, m.$$

\hspace{1cm} (4)
3.4. Processing of weight values and thresholds

The weight values and thresholds were predicted values generated based on historical data, and the weight values represented the magnitude of the impact of each input on the output, while the thresholds represented the change limits of the outputs.

The connection weight values $q^{t+1}_{h_i,j}$ and $q^{t+1}_{h_j,j}$ at the next time were updated based on the instantaneous error $e^{t+1}_{h_j}$, the input value of the input layer, the output value of the device layer, and the historical connection weight value:

$$d^{t+1}_{h_i,j} = q^{t}_{h_i,j} + \eta H^{t}_{h_j}(1-H^{t}_{h_j})X^{t}_{h_i} \sum_{k=1}^{m} q^{t}_{h_j,j_k} e^{t}_{j_k}$$  \hspace{1cm} (5)

Where $i = 1,2,\ldots,n$; $j = 1,2,\ldots,m$.

$$q^{t+1}_{h_j,j_k} = q^{t}_{h_j,j_k} + \eta H^{t}_{h_j}e^{t}_{j_k}$$  \hspace{1cm} (6)

Where $j = 1,2,\ldots,l$; $k = 1,2,\ldots,m$; $\eta$ is the learning coefficient and can be set to a constant.

The thresholds $a^{t+1}_{h_j}$ and $b^{t+1}_{j_k}$ of the device layer and event layer were updated according to the network’s instantaneous prediction error $e^{t}_{j_k}$, the input values of the input layer, the output values of the device layer, and historical thresholds:

$$a^{t+1}_{h_j} = a^{t}_{h_j} + \eta H^{t}_{h_j}(1-H^{t}_{h_j}) \sum_{k=1}^{m} q^{t}_{h_j,j_k} e^{t}_{j_k}$$  \hspace{1cm} (7)

Where $j = 1,2,\ldots,l$.

$$b^{t+1}_{j_k} = b^{t}_{j_k} + e^{t}_{j_k}$$  \hspace{1cm} (8)

Where $k = 1,2,\ldots,m$.

4. Intelligent analysis expert database

When the predicted output of the device layer $H^{t}_{h_j} > \theta_{h_j}$, it indicates that the device has a fault, while $\theta_{h_j}$ is a constant, indicating the fault threshold. When the predicted output of the event layer is $Y^{t}_{j_k} > \theta_{j_k}$, it indicates that the grid has an accident $y_k$, while $\theta_{j_k}$ is a constant, indicating the accident limit value. When the device failed, the input values $X^{t}_{h_i}$ of the weight value $q^{t}_{h_i,j}$, the weight values $q^{t+1}_{h_i,j}$, the thresholds $a^{t+1}_{h_j}$, and the output $H^{t}_{h_j}$ of the processing layer, were stored in the expert database, and the dispatching and handling information and method when the device $h_j$ failed were recorded. The output values $H^{t}_{h_j}$, the weight values $q^{t+1}_{h_j,j_k}$, and the thresholds $b^{t+1}_{j_k}$ of the device layer whose $q^{t}_{h_j,j_k} > \beta$ when the power grid had an accident $y_k$, and the outputs $Y^{t}_{j_k}$ of the device layer were stored in the expert database, the dispatching and handling information and method were recorded when the power grid had an accident $y_k$. $\alpha$ and $\beta$ in the above description are state constants.

The above is the process of building a self-learning expert database. The self-learning function of the self-learning expert database is achieved by timely updating the weight values and thresholds. It is only necessary to input the monitored and detected signals, and historical weight values and thresholds, to predict the equipment fault status and the grid accident status.

The self-learning expert database could automatically record the main influence factors of equipment failure and grid accidents. When predicting equipment failure or grid accident again, the data was
compared with the data in the expert database. If they were consistent, pre-judged alarm of equipment failure or grid accident were sent out. If there was no consistent data in the expert database, the data was stored and alarmed. Through this process, a self-learning expert database was gradually established.

5. Application cases and simulation analysis
The nodes $x_1$ and $x_2$ are the temperature and humidity data around the transformer 1, respectively, and $x_3$ and $x_4$ are the temperature and humidity data around the transformer 2. $h_1$ is the operating state data of the transformer 1, and $h_2$ is the operating state data of the transformer 2, while $y_1$ is the data of the freezing disaster accident state. According to the data in the established self-learning expert database, the connection weight value of $x_1$ and $x_2$ is 0, i.e. $q_{x_1,h_1}' = 0$ and $q_{x_2,h_2}' = 0$, similarly, $q_{x_3,h_1}' = 0$ and $q_{x_4,h_1}' = 0$. Since the monitoring signal data in the power grid constitutes big data, the amount of information is very large. For the convenience of discussion, it can be assumed that the connection weight values of other input signals to the transformer 1 and the transformer 2 are also zero.

When the temperature data $X_{x_1}' = 30°C$ around the transformer 1, the change of the humidity around the transformer 1 has little effect on the transformer. According to the formula (5), $q_{x_1,h_1}'$ is small. The threshold $d_{h_1}'$ can be obtained according to the formula (7), and from the formula (1) it can be found that the change of $H_{h_1}'$ is not obvious, and thus the change of $Y_{y_1}'$ is not obvious. When the temperature data $X_{x_1}' = -2°C$ around the transformer 1, the change of the humidity around the transformer 1 has a great influence on the transformer. According to the formula (5), it can be found $q_{x_1,h_1}'$ is very large. The threshold $d_{h_1}'$ can be obtained according to the formula (7), and from the formula (1) it can be found that the change of $H_{h_1}'$ is not obvious. According to other data from the grid, the changes of $Y_{y_1}'$ are obvious and a freezing disaster accident may occur.

At the time of $t_1$, a freezing disaster occurs, and the input data $X_{x_1}'$, $X_{x_1}'$, $X_{x_1}'$, $X_{x_1}'$, weight values, thresholds $q_{x_1,h_1}'$, $q_{x_1,h_1}'$, $q_{x_1,h_1}'$, $q_{x_1,h_1}'$, $a_{x_1}'$, $a_{x_1}'$, $b_{x_1}'$ and the output data $H_{h_1}'$, $H_{h_1}'$, $Y_{y_1}'$ by the device layer and the event layer are stored in the expert database. When the data similar with that at time $t_1$ is generated at time $t_n$, it is possible to warn of a freezing disaster at the moment, and the dispatcher can take relevant actions to process the alarm.

According to the remote measurement data of East China, the above application cases are simulated and calculated, shown as follows:
As shown in Fig. 3, according to the data of East China, the state changes of transformer 1 and that of freezing disaster with humidity are simulated at $30^\circ C$ and $-2^\circ C$, respectively. Since the change curves of transformer 1 and transformer 2 are very close, only the one of transformer 1 is simulated. In Fig. 3, the state of the transformer 1 and that of the freezing disaster incident at $30^\circ C$ are close to zero, indicating that no freezing disaster will occur. At $-2^\circ C$, the state of the transformer 1 and that of the freezing disaster have dramatic changes at the humidity of 75%, and the state value grows as the humidity increases. When the state of the transformer 1 reaches $\theta_{\text{th}}$, a freezing disaster occurs.
Fig. 4 shows the changes of weight values with humidity for at $30^\circ C$ and $-2^\circ C$. Fig. 5 displays the changes of the thresholds of transformer 1, transformer 2 and that of freezing disaster with humidity. In Fig. 4, the weight values increase with the increasing humidity. At $-2^\circ C$, the weight value from the input layer to the device layer and that from the device layer to the event layer are larger because of the great influence of humidity on equipment and freezing disaster events at low temperatures. The thresholds decrease with increasing humidity in Fig. 5.

6. Conclusions
This paper proposed a pre-judgment analysis algorithm for grid event monitoring, and carried out analyses of application cases and simulation calculation. The analysis showed that in the era of big data, the self-learning expert database of smart grid played a key role in monitoring and warning of power equipment and power grid accidents, with positive practical significance. In addition, by updating the weight values and thresholds of each node in time, the algorithm could comprehensively and accurately consider the impact of monitoring and controlling variables on equipment operation state and grid event state, and realize timely alarm function, which has practical application value for future research on smart grid.

References
[1] YU Yixin, LIU Yanli, Challenging issues of smart grid[J]. Automation of Electric Power Systems,2015,39(25):1-5.
[2] Vineetha, C.P. Smart grid challenges, issues and solutions[C]//Intelligent Green Building and Smart Grid (IGBSG). April 23-25,2014,Taipei:1 – 4.
[3] ZhANG Yin, CHEN Min, LIAO Xiaofei. Big data application: a survey[J]. Journal of computer research and development, 2013, 50(suppl): 216-233.
[4] Zibin ZHENG Service-Generated Big Data and Big Data-as-a-Service: An Overview[C]// Big Data (BigData Congress). June 27-July 2,2013.,Santa Clara, CA.:403 – 410.
[5] Xhafa, F. Processing and Analytics of Big Data Streams with Yahoo!S4[C]// Advanced Information Networking and Applications (AINA). March 24-27,2015,Gwangiu:263 – 270.
[6] ZHANG Dongxia, MIAO Xin, LIU Liping, ZHANG Yang, LIU Keyan. Research on Development Strategy for Smart Grid Big Data[J]. Proceedings of the CSEE, 2015, 35(1):2-12.
[7] CHEN Yuan, LU Jinling, YU Chengyang, LIU Haijun. Cascading Contingency Forecasting of Complex Power System[J]. Electric Power Science and Engineering,2008,24(3): 9-12.
[8] Zhang, X .Ding, Y Liu, and P . J . Griffin, An Artificial Neural Network Approach to Transformer Fault Diagnosis[J], IEEE Transactions on Power Delivery,1996,11(4):1836-1841.