Research and Application of Power System Data Anomaly Identification Based on Time Series and Deep Learning

Chao Li¹,², Xiaolei Li¹,², Xingyu Liu¹,², Lin Zhao¹,², Dajun Xiao³, Xuliang Hao¹,², Jun Xie³ and Xialing Xu³

¹ NARI Group Corporation (State Grid Electric Power Research Institute), Nanjing 211106, China
² Beijing Kedong Electric Power Control System Co., Ltd., Beijing 100085, China
³ Central China Branch of State Grid Corporation of China, Wuhan 430077, China

*Corresponding author email: lixiaolei3@sgepri.sgcc.com.cn

Abstract. Abnormal data in the power system will reduce the accuracy of system state estimation and affect the safe operation of the power dispatch system. This paper proposes a data anomaly identification model based on time series and neural network, which establishes time series for various measuring points of the control master station, creates a time series group of associated measuring points based on the network topology, and extracts the sample characteristics of the time series. The neural network model is used to realize the intelligent identification of normal data and abnormal data. The neural network recognition results are compared with normal distribution and DBSCAN density clustering methods to verify the abnormal recognition performance of the neural network. Using a provincial power grid dispatch center operating data set as a training and testing sample, it verifies the advancement of the proposed method in the comprehensive performance of anomaly detection recall rate and precision rate and its feasibility in actual system application.

Keywords: Abnormal data identification; Time series data; Neural Networks.

1. Introduction

With the increasing scale of power grid, the internal structure of power grid and system operation mode become more complex, intelligent and automatic, which become the direction and trend of power system development. Finally, the whole process of power system production can be monitored in full view [1], which puts forward a higher demand on the data quality of power system. However, in the actual production operation, the interference of various external environment or the abnormality of electrical equipment will lead to the measurement error, which will affect the data analysis, and then affect the decision-making [2]. Therefore, the anomaly detection of power grid dispatching data becomes more and more important.

At present, there are two kinds of outlier detection algorithms for power system data time series: outlier sequence point detection and outlier sequence detection [3]. Anomaly detection of time series points is the measurement of the value of the unit data points in time series. The main detection methods are: statistical outlier detection, clustering analysis based on outlier detection, support vector machine based on outlier detection, and so on [4]. The common point of this type of method is to pay attention to whether the value of a single time sequence point exceeds the corresponding threshold, but it often ignores the influence of data change trend on the judgment of time series abnormality, which is one-sided. For this reason, many researchers have done a lot of research on anomaly sequence detection. Combining the distance-based algorithm with the density-based Algorithm, the GMBR-DD
Anomaly Detection Algorithm is proposed by Sun Meiyu\cite{5}. Yu Yufeng and other people\cite{6} proposed a time series anomaly detection algorithm based on sliding window prediction. Reference\cite{7} combined the idea of segmentation and segmented the time series based on important points such as extreme points in the time series, and proposed a PLR-TSIP algorithm, which has a high degree of fit and high anomaly detection accuracy. But too many split points will affect its computing efficiency. The neural network model based on deep learning can consider both time series and nonlinearity of data\cite{8}, and the model has a wide range of application. The Model Algorithm has been applied to power market and load forecasting\cite{9-11}, showing good forecasting accuracy and better performance of robustness. However, the application of this method to data anomaly identification is still rare.

In this paper, an anomaly identification method based on time series and depth learning is proposed. Firstly, a time series is established according to the real-time data, the historical sampling data and the related calculation data of the automatic master station, and the time series group of the related measurement points is created based on the network topology. Secondly, the time series feature representation method is used to extract the time series feature of power data, so as to compress the high-dimensional time series into the low dimensional feature series. Finally, the intelligent identification of normal and abnormal data is realized based on the neural network model. The method has been verified and applied in a provincial power regulation and control network, and good results have been achieved.

2. Time Series

2.1. Establish Time Series Group

The network topology of power system makes the time series of correlated test points exist correlation relationship. It is helpful to improve the discrimination accuracy of anomaly detection task\cite{12}. In this paper, based on the network topology, a time series group is set up to identify the anomaly of the target and its related points. This method can mine and utilize the correlation information of multi-dimensional sequence, and it has strong interpretability and reduces the cases of wrong judgment and missing judgment.

In addition, the time series data points cannot be “peer-to-peer” matched because the station data is sent to the automation master station in the form of up-feed. Therefore, it is necessary to extract features and reduce dimensions of time series. Considering the stable characteristics of power system data, there is no need to process the time series in a sliding window. Only four time-domain features of length, mean, maximum and minimum and one frequency-domain feature of standard deviation of time series in the sliding window are extracted, so that the change trend of time series data in the sliding window is expressed as a 5-dimensional feature vector to realize data compression.

2.2. Data Preprocessing

In this paper, the sliding window based on time is adopted. Firstly, the real-time data, historical sampling data and related calculation data of each measuring point in the master station of control automation are analyzed, and the time series model is established.

The slide window in the same time segment stores the time series of the target measurement point and its associated measurement point, extracts the 5-dimensional features of the target measurement point time series and its associated measurement point time series, and takes them as the input of the neural network model. Then, the old data in the window is cleared, and the new time series is stored, and then repeat it. The flow is shown in Figure 1.
3. Experimental Analysis

3.1. Neural Network Structure
The time series data of master station are extracted and encapsulated into samples. Suppose there are m samples in total, then the input sample is a two-dimensional array $X_{m,n}$ of $m \times n$. The abnormal state of the data is taken as the output characteristic value of the sample.

The population sample is divided into training sample P and test sample T. Training samples are used to train the model, and test samples are used to verify and score the trained model.

The neural network model can be designed as follows: the model contains an input layer and can have any number of neurons. There are 6 perceptual layers (hidden layer), which contain 14, 12, 10, 8, 5 and 3 neurons respectively; one output layer and two neurons. Each sensing layer and output layer has a linear weight matrix $W$ and a bias matrix $B$, which are used to calculate the input data linearly.

Forward training: Select the ReLU function as the activation function. Let $W_k$, $B_k$ and $A_k$ represent the weight matrix, bias matrix and nonlinear activation function of the kth perceptual layer respectively.

So, the output value of layer k:

$$Z_k = A_k (W_k \cdot Z_{k-1} + B_k)$$  \hspace{1cm} (1)

Reverse feedback: for the loss value between the forward training calculation result and the actual result, the gradient descent method is used to carry out the reverse feedback regression layer by layer, so that the value of the loss function calculated by the training model is gradually reduced and converged to the minimum. Loss value of layer k:

$$D_k = D_{k+1} \cdot \frac{\partial C}{\partial W}$$  \hspace{1cm} (2)

Where $\frac{\partial C}{\partial W}$ is the partial differential of the output function of layer k + 1 to $W_{k+1}$.

The new weight matrix $W_k$ and offset matrix $B_k$ of each layer can be calculated inversely according to the above formula. Then retraining the next batch of samples.

When the training sample P is completely trained, the test sample T is brought into the training model to calculate the accuracy of the results.

3.2. Experiment Procedure
The following uses different methods to detect the abnormal active power of a line within 24 hours.

3.2.1. Normal distribution anomaly detection
Normal distribution anomaly detection is a traditional method for single point anomaly data detection in power system. There is the standard deviation of measurement error $\sigma$. Under normal measurement conditions, the error is more than $\pm 3\sigma$. The probability of the measured value is only 0.27%, which is a small probability event. If the error exceeds this range, it is considered as abnormal data.

The 24-hour positive and negative power of the line are taken as samples for ks normal verification.
KstestResult1(statistic=0.05077265229335348, pvalue=0.14181721398059965)
KstestResult2(statistic=0.05447467385382282, pvalue=0.13179118737450612)

Figure 2. Distribution histogram of sample data. The P values were greater than 0.05, indicating that the two samples were in line with the normal distribution, as shown in Figure 2. According to ± 3 \( \sigma \) for the boundary conditions, the samples are classified and predicted.

3.2.2. DBSCAN density clustering anomaly detection
DBSCAN is a kind of clustering algorithm based on density. The normal data in power system usually does not change dramatically; The abnormal jump data is often isolated in the whole sample. Based on this characteristic, DBSCAN density clustering method can be used for data anomaly detection. Sample selection remains unchanged. The samples are clustered by DBSCAN. The results are shown in Figure 3.

Figure 3. Clustering the samples with DBSCAN. When the clustering radius is set to 0.8, the prediction effect is the best.

3.2.3. Neural network anomaly detection
Using the 3.1 training model, the first 10 days data of the target sample is taken as the training sample, and the gradient descent step of the loss function in the reverse feedback is set as 0.005. The initial weight matrix and offset matrix of each hidden layer are obtained by a standard normal distribution random function. The training samples are trained for 100,000 rounds, and the number of samples in each round is 30. The curve of loss value during training is as follows in Figure 4.

Figure 4. Loss curve of neural network training process.
The final loss value is less than 1.5, and the loss curve tends to be flat, indicating that the model has been trained mature. Save the model, simulate and predict the target sample.

3.3. Results Comparison
The performance of anomaly detection algorithm is evaluated by recall and precision. An effective anomaly detection method needs to maintain high precision and recall at the same time. F1 score is an index used to measure the accuracy of binary classification model in statistics. It considers the accuracy and recall of the classification model. The F1 score can be regarded as a harmonic average of model accuracy and recall, with the maximum value of 1 and the minimum value of 0. The closer F1 is to 1, the better the performance of anomaly detection algorithm in terms of accuracy and recall. The results are as follows in Table 1.

| Statistical results | normal distribution | DBSCAN | Neural network |
|---------------------|---------------------|--------|----------------|
| Actual normal       | Normal prediction   | Abnormal prediction | Normal prediction | Abnormal prediction | Normal prediction | Abnormal prediction |
| 1116                | 963                 | 153    | 1033           | 83                | 1098           | 18               |
| Actual abnormal     | 324                 | 195    | 129            | 217               | 107            | 59               | 265             |
| Accuracy rate       | 75.8%               | 79.2%  | 94.7%          |                   |                |                  |
| Precision rate      | 83.2%               | 82.6%  | 94.9%          |                   |                |                  |
| Recall rate         | 86.3%               | 92.6%  | 98.4%          |                   |                |                  |
| F1 value            | 84.7%               | 87.3%  | 96.6%          |                   |                |                  |

Compared with the above four data anomaly detection methods, when the sample data is large enough, the neural network algorithm has wider applicability, higher accuracy and relatively independent of human experience.

4. Application Examples

4.1. Selection of Sample Points
Jump is a common data anomaly in the actual power dispatching and production system. A/D conversion error, station end acquisition device anomaly, frequent switching between main and standby machines or dual host can cause data jump fault in the master station. The data anomaly identification method based on time series and deep learning proposed in this paper is connected to the production system of a provincial power grid dispatching center. The data generated by the actual operation system is used to verify the effect of jump fault detection.

In this section, the training data set and test data set used in the algorithm model are selected from the actual production data set. Taking 10s as the time length of sliding window, 100,000 labeled data are selected as the training sample set, 37,180 normal samples are selected for algorithm model test, and 5,000 fault samples including data jump are selected, and the jump rates are ≥ 500%, ≥ 400%, ≥ 300%, ≥ 200%, ≥ 100% (which can cover all common data jump degrees of master station).

4.2. Result Analysis
The training sample set and the test sample set with different jump rates are substituted into the model algorithm, and the performance indexes for different jump rates are shown in Table 2.
Table 2. Experimental results of jump fault detection.

| Jump Rate | Accuracy Rate | Precision Rate | Recall Rate | F1   |
|-----------|--------------|----------------|------------|------|
| ≥500%     | 0.984        | 0.977          | 0.993      | 0.985|
| ≥400%     | 0.973        | 0.971          | 0.991      | 0.981|
| ≥300%     | 0.961        | 0.967          | 0.991      | 0.970|
| ≥200%     | 0.952        | 0.956          | 0.988      | 0.972|
| ≥100%     | 0.947        | 0.949          | 0.984      | 0.965|

The results show that the proposed method has different performance for faults with different jump rates. The higher the jump rate is, the higher the F1 value is, and vice versa. Among them, the accuracy rate of 0.977 and recall rate of 0.993 are obtained for jump faults with jump rate ≥ 500%, and F1 value reaches 0.985; the accuracy rate of 0.949 and recall rate of 0.984 are obtained for the jump fault whose jump rate is more than or equal to 100%. The F1 value reaches 0.966. In addition, the overall accuracy was 0.947 to 0.984. The results show that the data anomaly identification method based on time series and deep learning has very good comprehensive performance and has good application effect in power dispatching automation system.

5. Conclusion

Based on the analysis of the data characteristics of power dispatching production system, this paper proposes a method of data anomaly recognition based on time series and deep learning, which extracts feature from time series data and combines with correlation points, which greatly improves the solvability of the sample features. The experimental comparison shows that the method is advanced in the comprehensive performance of abnormal data detection. It is verified in the actual operation system. In the follow-up research work, in order to further improve the comprehensive performance and applicability of anomaly detection, the following two aspects can be improved: one is to consider the data upload time in the sample characteristics to avoid the interference of different curve shapes and different time distributions on the model algorithm; Second, different anomaly recognition models can be established for different characteristic physical quantities to improve the accuracy of anomaly recognition.

Acknowledgments

Thank you for the science and technology projects of State Grid Corporation of China aid financially ‘Research on key Technologies of improving data quality and monitoring reliability of dispatching automation master station system from multiple angles’. (5108-202040021A)

References

[1] Zhang Dongxia, Miao Xin, Liu Liping, et al. Research on development strategy for smart grid big data[J]. Proceedings of the CSEE, 2015, 35(1): 2-12(in Chinese).
[2] ZHU Qianwen, YE Lin, ZHAO Yongning, LANG Yansheng, SONG Xuri. Methods for elimination and reconstruction of abnormal power data in wind farms[J]Power System Protection and Control., 2015, 43(3):38-45 (in Chinese).
[3] Zhang Qi. Reserch on anomal y detection for streaming time series.[D], Shandong University, 2019.
[4] WangJing, A Survey on Power System Bad Data Detection and Identification. Electricity and energy, 2015, 36(6): 813-817.
[5] SUN Meiyu. Research on discords detect on time series based on distance and density. Computer Engineeringand Applications, 2012, 48(20): 11-17.
[6] YU Yufeng, ZHU Yuelong, WAN Dingsheng. Time series outlier detection based on sliding window prediction[J]. Journal of Computer Applications, 2014, 34: 2217-2220.
[7] SUN Z Y, DONG L L, MA Y J. A time series segmentation algorithm based on important
[7] HUANG Heng, WANG Zhiyong. Short-term load forecasting based on a hybrid model of ANN and SVM [J]. Acta Physica Sinica, 2006, 55(2): 375-382.

[8] KIM H Y, WON C H. Forecasting the volatility of stock price index: a hybrid model integrating LSTM with multiple GARCH-type models [J]. Expert Systems with Applications, 2018, 103: 25-37.

[9] DUDEK G. Pattern-based local linear regression models for short-term load forecasting [J]. Energy, 2016, 130: 139-147.

[10] GOU Xuan, XIAO Xianyong. Short-term electricity price forecasting model based on empirical mode decomposition and LSTM neural network [J]. Journal of Xi'an University of Technology, 2020.

[11] CHEN Zhuo, SUN Longxiang. Short-Term electrical load forecasting based on deep learning LSTM networks [J]. Electronic Technology, 2018, 047(1): 39-41.

[12] DING Xiao-Ou, YU Sheng-Jian, WANG Mu-Xian, WANG Hong-Zhi, GAO Hong, YANG Dong-Hua. Anomaly Detection on Industrial Time Series Based on Correlation Analysis [J]. Journal of Software, 2020, 31(3): 726-747.