Design Engineering

ISSN: 1520-0191
Issue: 1 | Pages: 103 - 110

Electricity Stealing Behavior Detection Method based on BP Neural Network

Jinliang Liu[1,*], Guoping Li[1], Wenle Song[1], Di Liu[2], Tiemin Jiang[2]

[1]China National Grid, Cangzhou Power Supply Company, Cangzhou, 061000, China

[2]Department of computer science, North China Electric Power University, Baoding, 071000, China

*Corresponding Author: Jinliang Liu

Abstract:

In view of the outstanding problems of current equipment stealing, device specialization, concealed behavior, and large-scale implementation, this paper proposes a detection model of electricity stealing users to realize the screening of electricity stealing users. The model first preprocesses the user's electricity data to determine the evaluation indicators of electricity theft; then uses the PCA method to achieve the feature extraction, thereby improving the efficiency and reducing the load on the detection system; then using BP neural network, the network completes the diagnosis of the user's electricity stealing behavior, and realizes the judgment of normal users and electricity stealing users. Finally, the power consumption data of enterprise large users provided by a power grid company is used for experimental analysis to verify the validity and applicability of the model.

Keywords: *Anti-electric theft, BP Neural Network, Principal component analysis.*

I. INTRODUCTION

The electricity theft not only threatens the safety of power supply and disrupts the normal order of power supply, but also causes huge economic losses to the country and power supply enterprises. The annual losses caused by electricity theft in the country are around 20 billion yuan, and the detected electricity theft cases are less than 30% of the total electricity theft cases. The traditional electricity inspection and anti-theft leak detection work mainly rely on surprise inspection methods to combat electricity theft, which has inherent defects and deficiencies. As the problem of electricity theft becomes more prominent, the detection of electricity theft urgently needs to be improved. Various current anti-stealing technologies still leave room for some illegal users to steal electricity and destroy metering devices, mainly manifested in the problems of reliability, timeliness and accuracy. With the gradual popularization of electricity consumption information collection systems in China, the method of stealing electricity has developed into intelligent equipment, specialized means, concealed behavior, and large-scale high-tech electricity stealing. Existing anti-stealing technology can no longer thoroughly investigate and punish all methods of stealing electricity, and there is an urgent need to carry out research and application of intelligent anti-stealing technology [1-2].

ISSN: 1520-0191
Issue: 1 | Pages: 103 - 110

Existing power metering automation systems can collect power load data, and terminal alarm information such as abnormal power consumption. If these data information can be used to extract the key characteristics of users who steal electricity, they can construct electricity leakage. The user's identification model can automatically check to determine whether the user has electricity theft or leakage. At present, abnormal power consumption behavior detection can be completed from three perspectives based on system state, data-driven and game theory. Among them, data-driven methods can be divided into three types of methods based on classification, regression-based and cluster-based, including Algorithms such as Bayesian network, decision tree, outlier detection, BP neural network, KNN, and SVM have achieved good results [3-10]. Currently, most electricity theft occurs on 10kV and 380V power supply lines. According to the principle that electricity consumption abnormality occurs when electricity stealing occurs, abnormality is found through data analysis, and suspected objects are selected for inspection and treatment, which makes the anti-stealing work more targeted and can effectively combat electricity stealing.

In this paper, based on the three-phase voltage and three-phase current data of the electricity information collection system, the proposed electricity theft detection model is mainly composed of two parts, feature extraction based on principal component analysis (PCA) and feature matching based on neural network to improve electricity theft Check accuracy. This article compares the performance of SVM and compares the effects of different learning algorithms on the accuracy of electricity theft detection.

II. DATA PREPROCESSING

Preprocess the power load data of large enterprise users from 2018 to 2019 provided by a power grid company. The original data is to record the user's ABC three-phase current, ABC three-phase voltage, transformer capacity, rate, and maximum current every 15 minutes, The minimum current, the unbalance rate, and record the electricity consumption of the day once a day, and the data on the day when the user breaches the contract and steals electricity are marked in the data. The sample data includes the data of the users who steal electricity and the data of normal users. In order to make the sample data closer to the actual situation, most of the sample data is normal electricity consumption data, and a small part is the user electricity consumption data with electricity stealing phenomenon.

Data preprocessing involves two aspects: data cleaning and standardized processing. On the one hand, the data may be duplicated, missing, or even wrong, so the data needs to be preprocessed, mainly to delete duplicate information, fill in the missing information, and correct the wrong information; on the other hand, because the evaluation indicators are different in nature and usually have different dimensions and orders of magnitude. Therefore, need to standardize the original indicator data to ensure reliable results.

2.1 Data Filtering

Compared with working days, the electricity consumption during holidays will be significantly lower. Therefore, filter holiday electricity data to achieve the best data effect

ISSN: 1520-0191
Issue: 1 | Pages: 103 - 110

possible.

2.2 Missing Value Processing

It is found that the original measurement data is missing. If these values are discarded directly, the data may be less effective. So need to deal with missing values. In this paper, the Lagrange interpolation method is used to interpolate the missing values: first determine the dependent variable and independent variable from the original data set, extract the 5 data before and after the missing value, and form a group based on the 10 data taken out, and then use Lag The Lagrange polynomial interpolation formula interpolates all missing data in sequence until there are no missing values.

2.3 Data Transformation

2.3.1 Normalization of single-day data

The three-phase current and three-phase voltage data of the system is collected for 15 minutes, and the user power consumption is collected for daily. To ensure that the data dimensions are the same, the daily data needs to be reduced to a record of electricity consumption.

2.3.2 Join the trend indicator of power consumption decline

A few days before and after can be considered as a statistical window period, using the slope of the straight line fitting of the electricity as a measure, if the slope continues to decline with time, then the user is likely to steal electricity [12-13]. Five days before and after the day of statistics is set as the statistical window period, and the decline in power trend during these 11 days is calculated. First calculate the daily power trend in these 11 days, of which the power consumption trend on the i-th day is to consider the slope of the power consumption during the 5 days before and after, namely:

$$x_i = \frac{\sum_{l=i-5}^{i+5} (f_l - \bar{f})(l - \bar{l})}{\sum_{l=i-5}^{i+5} (l - \bar{l})^2} \quad (1)$$

Among them, xi is the power trend of the i-th day, fl is the power consumption of the l-th day, $\bar{f}$ and $\bar{l}$ are the average of 5 days before and after.

III. CONSTRUCTION OF ELECTRICITY THEFT IDENTIFICATION MODEL

3.1 Feature Extraction

If the data of electricity stealing indicators are directly analyzed and too many electricity stealing assessment indicators are selected, the data will be complicated, the network performance will be reduced, the system load will be increased, and the indicator data at each moment will affect and correlate with each other. Therefore, it is necessary to process the data of electricity theft indicators. PCA is the most commonly used linear dimensionality reduction method. It expresses the original multiple variables in a linear combination by using several principal components. PCA is to map n-dimensional features onto k (n>k) dimensional new

ISSN: 1520-0191
Issue: 1 | Pages: 103 - 110

orthogonal features. It can be proved that PCA is a linear dimensionality reduction method that minimizes the loss of original data information through some linear projection.

In this paper, the PCA method is used to extract the sample data from the high-dimensional data to extract feature data (note here that the dimensionality reduction is not performed on the electric leakage labeling data items). When the cumulative contribution rate of the data items after dimensionality reduction is 95%, then it can be considered that the data after dimensionality reduction has similar data effects to the data before dimensionality reduction.

3.2 Theft Identification Network

A three-layer BP neural network is used as the detection model. As shown in Fig 1. In this paper, the number of input layer nodes, output layer nodes and hidden layer nodes are 12, 1, and10.

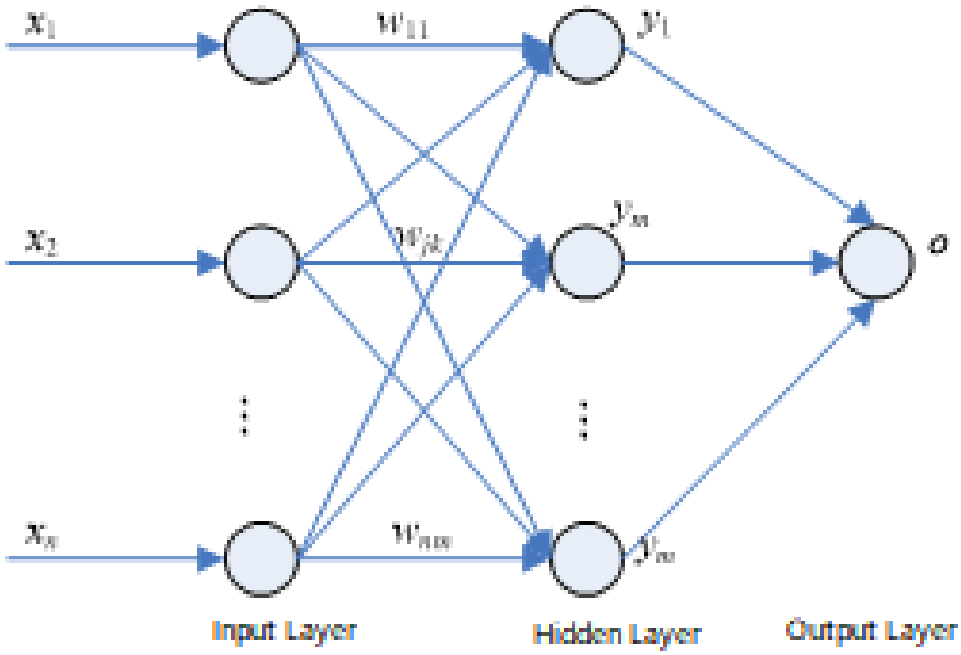

Fig 1 Theft identification network

In Figure 1,the input and hidden layer output vectors of the network are respectively x=(x1,x2,⋯ ,xn)T and y=(y1,y2,⋯ ,ym)T,The weight matrix between the input layer and the hidden layer and between the hidden layer and the output layer are respectively wjk(j=1,2,⋯ ,n,k=1,2,⋯,m) and w=(w1,w2,⋯,wk,⋯,wm)T,o is the output of the network.The input vector of the BP network is the dimensionality-reduced feature vector obtained in the feature extraction stage. The hidden layer uses Relu(x)=max(x,0) as the activation function.

In summary, the overall electricity theft detection model is shown in Fig 2.

ISSN: 1520-0191
Issue: 1 | Pages: 103 - 110

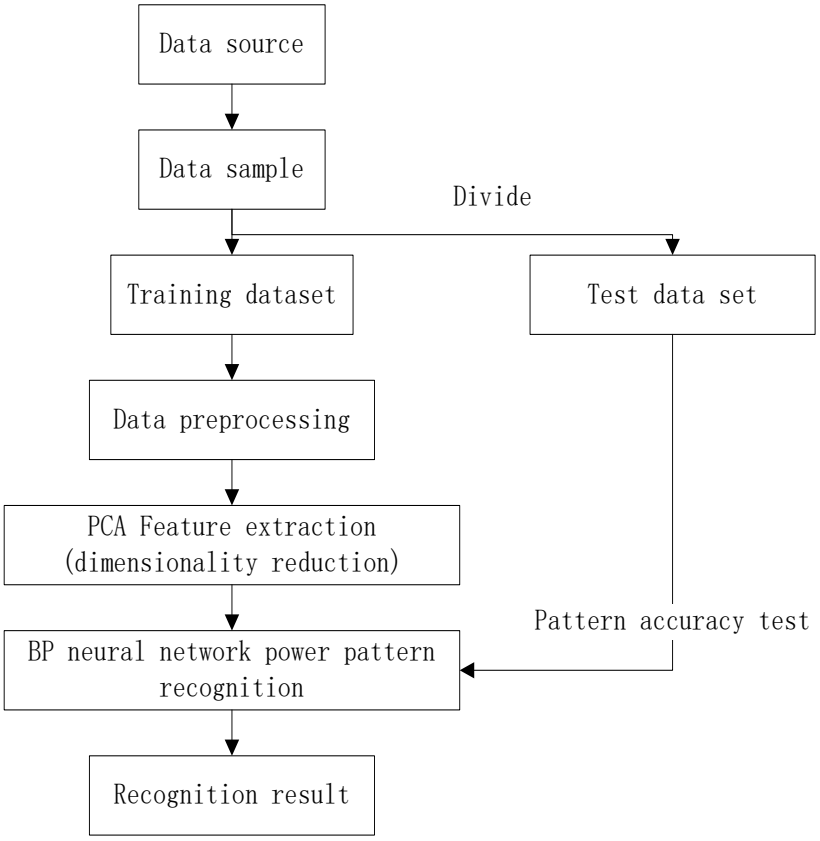

Fig 2: User stealing detection model

IV. EXPERIMENTAL ANALYSIS

This paper selects the power consumption data of power users from 2018 to 2019 provided by a power grid company, including data on the three-phase current, three-phase voltage and power consumption of the users and normal users, as well as the signs of whether the users steal the power. 730 records. The data collection interval is 15 minutes, 96 points a day, to clean the data and remove abnormal data. The goal of this article is to build a user identification model for electricity leakage, which can be used to implement user diagnosis.

In the experiment, the PCA was used to reduce the dimension of the 730 data processed for two years. When the cumulative contribution rate of the data items is 95%, it can be considered that the data before and after dimensionality reduction have similar data effects. As a result, the dimensionality reduced data is finally 13-dimensional, which contains the label items of electricity leakage. Each time before training the BP neural network model, the dimensionality reduction data is randomly shuffled, 80% of which are selected as training samples, the remaining 20% are used as test samples, and the number of iterations is 5000. After learning and training, the confusion matrix of the user electricity theft detection model adopted in this paper is shown in Fig 3.

ISSN: 1520-0191
Issue: 1 | Pages: 103 - 110

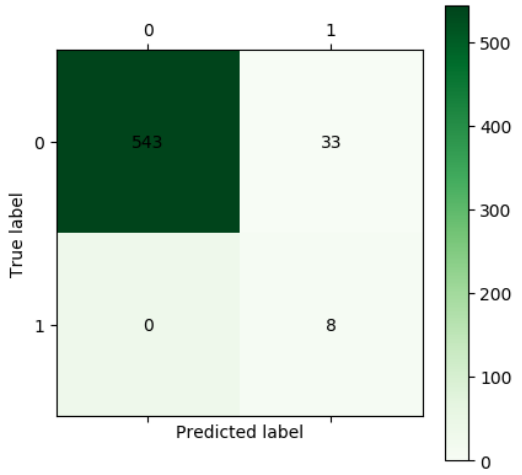

Fig 3: Confusion matrix based on BP neural network

It can be seen from Fig 3 that the classification accuracy rate based on BP neural network is 94.3%, the normal electricity consumption is misjudged as electricity leakage and accounts for 5.7%, and the electricity leakage behavior is misjudged as normal electricity consumption has not occurred. Fig 4 is the ROC curve based on the BP neural network.

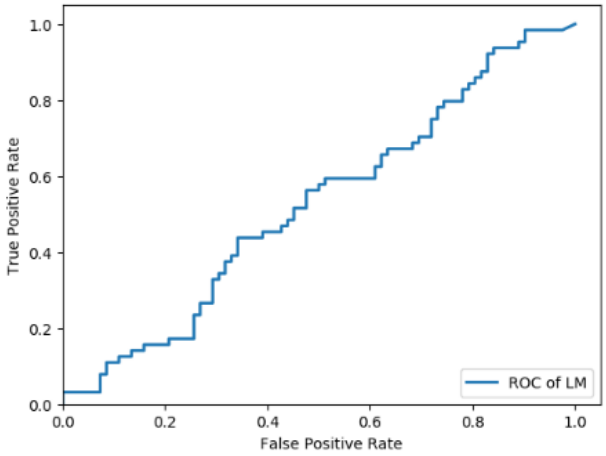

Fig 4: ROC curve based on BP neural network

In order to verify the performance of the model established in this paper, consider using Support Vector Machine (SVM) to identify the eigenvectors after dimensionality reduction. Through parameter optimization, the classification regression selection parameter of the support vector machine is 0, the kernel function type is RBF, the penalty coefficient is 0.03125, and the kernel function parameter is 0.5. The confusion matrix based on SVM is shown in Fig 5.

ISSN: 1520-0191
Issue: 1 | Pages: 103 - 110

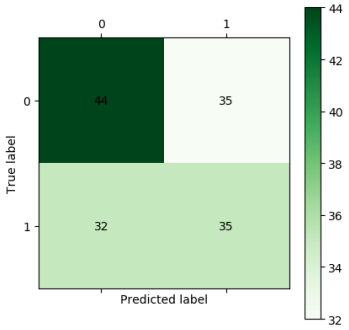

Fig 5: SVM-based confusion matrix

It can be seen from Fig 5 that the classification accuracy is 54.1%. The normal electricity consumption is mistakenly judged as stealing electricity and accounts for 44.3% of the normal electricity consumption. The electricity stealing behavior is mistakenly judged as normal electricity and is 47.8%.Fig 6 is the ROC curve based on the SVM model.

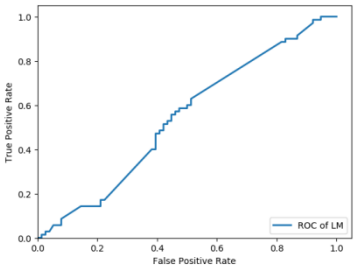

Fig 6: ROC curve based on SVM

It can be seen from the analysis that the classification accuracy rate based on the BP neural network detection model constructed is high, the misjudgment rate of normal power consumption behavior is relatively low, and the missed judgment rate of power stealing behavior has not occurred, which can basically meet the power application. The demand has practical significance for the management of electricity leakage. However, the performance of the SVM-based electricity theft detection model is not ideal.

V. CONCLUSION

Based on the characteristics of electricity stealing, this paper combines electricity stealing evaluation indicators with machine learning to construct a electricity stealing detection model and applies it to the identification of electricity stealing users. Experiments verify the feasibility and effectiveness of the proposed algorithm. The structure of the electricity stealing detection model proposed in this paper is relatively thin, which will affect the systematicity and accuracy of electricity stealing detection. Subsequent work will further improve the evaluation indicators and feature matching algorithms of electricity stealing to improve the accuracy and efficiency of the electricity stealing detection model.

ISSN: 1520-0191
Issue: 1 | Pages: 103 - 110

REFERENCES

[1] Hu Jiangyi, Zhu Enguo, Du Xingang, et al. (2014) Application status and development trend of electricity information collection system. Power System Automation (2): 131-135

[2] Wang Quanxing, Li Sitao (2016) Analysis and preventive measures of anti-stealing technology based on collection system. Electrical Measurement and Instrumentation 53(17): 78-83

[3] Sahoo S, Nikovski D, Muso T, et al. (2015) Electricity theft detection using smart meter data. Innovative Smart Grid Technologies Conference IEEE 1-5

[4] Chen Qixin, Zheng Kedi, Kang Chongqing (2018) Detection method of abnormal power consumption: review and prospects. Power System Automation 42(17): 189-199

[5] Kang Ningning, Li Chuan, Zeng Hu, Li Yingna (2017) Electricity theft detection using FCM clustering and improved SVR model. Journal of Electronic Measurement and Instrument 31(12): 2023-2029

[6] Cheng Chao, Zhang Hanjing, Jing Zhimin, etc. (2015) Research on anti-stealing electricity based on outlier algorithm and electricity information collection system. Power System Protection and Control 43(17): 69-74.

[7] Chen Tong, Fu Feng, Wang Jun, Chen Shuang (2016) CAPSO-BPNN-based early warning method for the operation status of metering devices. Electrical Measurement and Instrumentation 53(17): 65-70

[8] Wang Xinxia, Wang Ke, Jiao Dongxiang, etc. (2017) Research on anti-stealing electricity based on normal distribution outlier algorithm. Electrical Application 36(7): 60-65

[9] Zhuang Chijie, Zhang Bin, Hu Jun, etc. (2016) Detection of abnormal power consumption patterns of power users based on unsupervised learning. Chinese Journal of Electrical Engineering 36(2): 379-387.

[10] Yap KS, Tiong SK, Nagi J, et al. (2012) Comparison of supervised learning techniques for non-technical loss detection in power utility. International Review on Computers and Software 7(2): 626-636

[11] Xu Zhi, Li Hongjiao, Chen Jingjing (2017) User-stealing behavior prediction based on machine learning. Journal of Shanghai Electric Power University 33(4): 389-393

[12] Shi Yuliang, Rong Yiping, Zhu Weiyi (2018) Electricity stealing behavior recognition method based on analysis of electricity consumption characteristics. Computer Research and Development, 55(8): 1599-1608

[13] Zhang Liangjun (2015) Python data analysis and mining combat. Beijing: Mechanical Industry Press

[14] Chen Wenying, Chen Yan, Qiu Lin, etc. (2016) Analysis of anti-stealing electricity using big data technology. Journal of Electronic Measurement and Instrument 30(10): 1558-1566