Positive Active Power Outlier Detection based on One-Class SVM

Bin Ma
State Grid Hebei Electric Power Co.Ltd.
Shijiazhuang City, Hebei Province
mabin8736@163.com

Long Yuan
State Grid Hebei Electric Power Co.Ltd.
Shijiazhuang City, Hebei Province
531206360@qq.com

Shaozhe Xu,Kuanyun Zheng,Fuxing Huang
NARI Group Co., Ltd (*State Grid Electric Power Research Institute*)
Nanjing,Jiangsu, China
xushaozhe@sgepri.sgcc.com.cn

Runlong Li,Peisen Yuan
College of Information Science and Technology
Nanjing Agricultural University
Nanjing 210095, Jiangsu China
peiseny@njau.edu.cn

Abstract: The smart grid creates a large-scale intelligent energy delivery network. It requires a much more complex and high-dimensional data processing system. Thus, smart grid has become one of the most potential fields for machine learning application. This paper adopts an anomaly detection algorithm based on one-class SVM to realize anomaly detection task based on large-scale electrical energy data. One-class SVM model uses training data to train a hypersphere and calculate the distance between data points and center. The algorithm is able to process a large number of data quickly and has high reliability. In this paper, large-scale positive active power data is used as training and testing datasets. As the experiment shows, the algorithm has high detection efficiency and high accuracy.

Keywords—Outlier, Positive Active Power, One-Class SVM

I. INTRODUCTION

The smart grid can be considered as a modern electric power grid infrastructure with a larger scale and higher dimension[1]. Machine learning technologies can obtaining hidden knowledge from huge amounts of data and analyze big data effectively. The big data features of the smart grid match the research objects of machine learning. It is of great significance to ensure the accuracy and reliability of energy data and machine learning. Under such circumstances, machine learning algorithms has been widely used in power data anomaly detection.

The cause of abnormal energy data attributes to a measurement equipment failure or an ill-intentioned manipulation of electricity equipment[2]. The anomaly detection in the field of electricity can be divided into two kinds. The first way is to make in situ inspections and check equipment regularly. Nevertheless, the employment of professional inspectors causes a high cost. The second solution is to apply data mining techniques to detect anomaly. Such techniques proved to be success in multiple fields, including media and medical[3].

The application of data mining algorithms in positive active total power outlier detection is important to guarantee the accuracy of electricity energy data.

First of all, it is difficult to define a normal region because the boundary between normal and abnormal data is not accurate. Secondly, deliberate adversaries may pretend to imitate normal behaviors in order to conceal its purpose. Thirdly, there exists noise data similar to actual anomalies[4]. It increases the difficulty to recognize and remove the anomaly.

In the recent years, power outlier detection has became the heated topic of a number of surveys and articles. Wang et al.[5] introduced synchronization measurement information provided by the phasor measurement unit into outlier detection and present an algorithm used to detect topology error based on the synchronization measurement information. Wesin Alves [6] presented a methodology including a pre-processing algorithm and hybrid approach outlier detectors to deal with SCADA's (Supervisory Control and Data Acquisition) for big power data.

However, techniques shown above are mainly based on estimation, their calculation process is complicated and efficiency is not high, which is not suitable for detecting a large number of high-dimensional electrical energy data. In addition, anomaly detection algorithms only define concept of normal samples, so that the description of abnormal samples will not be optimized, which may cause detection errors or only a few anomalies detected.

Therefore, the application of machine learning technology in anomaly detection has gradually developed. According to two main categories of machine learning algorithms, it is divided into supervised[7] and unsupervised[8] anomaly detection algorithm.

In machine learning, one-class support-vector machines (One-class SVM) [9] is a supervised learning model that conducts data classification and data analysis. The object of outlier detection support vector model is data that only has one class which is defined as normal. The model infers the properties of normal cases and can judge data which is unlike the normal data from these properties.

In this paper, an anomaly detection algorithm based on One-class SVM is used to realize anomaly detection of large positive active total power data. This article uses the real electrical energy data from 2017/07/27 to 2019/06/11 of NARI Group to detect abnormal data on the positive active power.

978-1-7281-5748-1/20/$31.00 ©2020 IEEE

II. OUTLIER DETECTION

2.1 Anomalies in time series data

Anomalies, or outlier data are data points that exist but not clustered in the data set. Given an unlabeled data set $D=\{x_1,x_2,...,x_n\}$, train the model of the data as $P(x)$, where x is the characteristic variable. If it is the threshold $P(x_t)<\varepsilon$, where ε is the threshold, then it is considered abnormal[4].

Time series is a sequence of numerical data points collected in successive periods. Time series data mining can find the connection between the data point and time period. Time series data mining can also find some time series with a fixed sequence from a large number of data. If data in time series is abruptly changed at certain moment, the data point is likely to be an outlier. Time series anomaly detection technologies are widely used in fields such as stocks and weather. Chen et al.[10] adopted non-negative matrix factorization (NMF) to extract stock features, used wavelet transform and weighted fusion to find abnormal data points. Onal et al.[11] implemented *k*-means clustering based on public weather dataset and successfully extract anomalies.

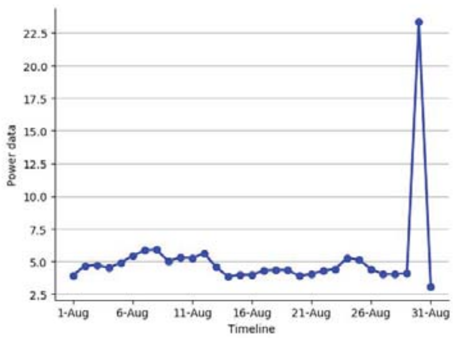

Figure 1 positive active power data

Take positive active power data for instance. Figure 1 is data of positive active total power consumption in August 2017. The year in the following figure is 2017. On August 30, there was a sudden change from the previous data point, and the power data on that day increased significantly, then this point is likely to be anomaly in the power time series data.

The typical outlier detection methods of sequence data can be divided into three categories:

(1) Outlier detection based on window

This method divides the sequence into specific fixed-length windows. But this method doesn't work well when the window cannot cover the abnormal sequence.

(2) Outlier detection based on proximity[13]

The distance between the sequence data can judge whether the data is abnormal. The prediction effect depends on the pros and cons of the proximity measurement method.

(3) Outlier detection based on prediction[14]

Data points with large differences between predicted data and real data can be regarded as abnormal, but the detection accuracy will be weaken when the prediction time is too long.

The algorithm in this paper is an outlier detection method based on proximity. It aims to learn a hyperplane that the normal and abnormal data points far away as possible from it.

2.2 Anomalies in power data

Many research projects aim to detect power data anomalies and find whether the power data performs normally in sequence. Catterson et al.[15] used Gaussian mixture model to detect abnormal power data of old transformers. The model is able to identify anomalies in transformer measurements and ignore normal data points with anomalous environmental factors. Jakkula et al.[16] built a novel statistical techniques to identify outliers and compared the performance with clustering methods. The new method achieved better stable results.

The data set in this paper is positive active total power data provided by a power enterprise. Active power is also referred to real power or true power. Active power is a form of powers in AC circuits and is the actual amount of power being dissipated or performs the useful work in the circuit[17]. The power oscillates between zero and positive maximum value, which produces positive active power. The experiment's purpose is to detect whether the data points in the data set are anomalies.

III. METHOD DESCRIPTION

3.1 One-class SVM

As a method for anomaly detection, one-class support vector machine trains a representational model with normal data. If new testing data is too different from the origin ones, it will be labeled as out-of-class.

The one-class support vector machine data set is like $D=\{(x_1,y_1),(x_2,\ y_2),\ldots,(x_n,y_n)\}$. x_i is the *i*-th input data point and y_i is the *i*-th output point which refers to the classification of the class. y_i equals to 1 or -1 ,which means normal or abnormal data.

The support vector domain description (SVDD)[18] is an algorithm to describe a kind of data points with same classification. It trains a hypersphere that gives a closed boundary around data points. SVDD aims to find a spherical surface with center a and radius R .

$$F(R,a,\xi_i)=R^2+C\sum_i \xi_i \tag{1.1}$$

Parameter C in Eq. (1.1) decides the trade-off between errors and volume. The introduction of ξ_i is to allow the distance from center a to x_i not strictly smaller than R, in order to allow outlier training data points. Thus this sphere should satisfy :

$$(x_i - a)^T (x_i - a) \le R^2 + \xi_i, \forall_i, \xi_i \ge 0 \qquad (1.2)$$

Then the Lagrangian multipliers is used, which is defined as Eq. :

$$L(R, a, \alpha_i, \xi_i) = R^2 + C\sum_i \xi_i - \sum_i \alpha_i \{R^2 + \xi_i - (x_i{}^2 - 2ax_i + a^2)\} - \sum_i \gamma_i \xi_i, \alpha_i \ge 0, \gamma_i \ge 0 \qquad (1.3)$$

L should be minimized with respect to R , a , ξ_i and maximized with respect to α_i , γ_i .After practice, to test a new data point z , distance to the sphere center should be calculated. If the distance is smaller than R , it can be judged as a part of this classification.

One-class SVM algorithms use a set of functions which are defined as kernel. The purpose of kernel is to transform data points into the required form for processing. Radial basis function (RBF) is the used type of kernel function as defined in Eq.(1.4). The function return the inner product between two points in a suitable feature space， which is low computational cost in very high-dimensional spaces.

In Eq.(1.4), gamma is the parameter which decides the data distribution after mapping.

$$k(x, z) = \exp(-gamma \cdot d(x, z)^2), gamma = \frac{1}{2 \cdot \sigma^2} \qquad (1.4)$$

One merit of one-class SVM is that the non-linear decision boundary can be made to recognize abnormal data points which cannot be divided through a linear function. As is shown in the Figure 2,any data points outside the one-class classifier dotted line are anomalies.

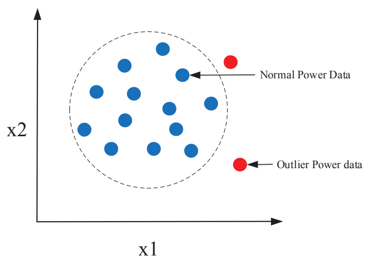

Figure 2 One-Class Classifier based outlier detection

3.2 Processing procedure

Our method has five steps: (1) Label the origin positive active power data with 1 or -1 based on whether they are anomalies. (2) Input the labelled PAP datasets to calculate a hypersphere. (3) Train our One-Class SVM model. (4) Test the performance on this model, use accuracy and recall rates to evaluate models. Then optimize the parameters like gamma to get better results. (5) Compare the model performance with other algorithms, robust covariance and local outlier factor. We will calculate indicators of each algorithm including anomalies number, accuracy and recall. The process figure is as below.

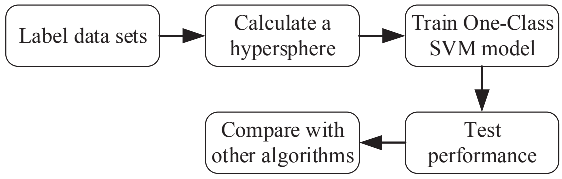

Figure 3 Processing steps of positive active power data outliers

IV. Experiments and Results

4.1 Data set

The data is 686 electric energy data records between 2017/07/27 and 2019/06/11, including electric energy recording time TIME and corresponding positive active total power data PAP. The abnormal data points are marked with -1 while normal points with 1. The data sets noted 18 anomalies manually.

4.2 Parameter settings

The parameter settings are based on data set. Radial basis function (RBF) is the most used type of kernel function because it has localized and limited response along the entire x-axis, which fits the features of anomaly detection datasets. We chose RBF as the kernel, and the training error is 0.026, to fit the real data situation. is As a parameter for a nonlinear support vector machine (SVM) with a Gaussian radial basis function kernel, Gamma is set to 0.011.

4.3 Results

Data set are trained into one-class SVM model to make it learn representation, calculate a hypersphere and output the predicting results.

Figure 4 Positive active power outlier detected results with the One-Class SVM

As is shown in the figure, the model detects 18 anomalies in the dataset, which is equal to the manually marked results and accuracy achieves 99.4169%. It shows that one-class SVM can correctly calculate a hypersphere and detect outlier data of positive active power data.

Then we compare the result performance with other two anomaly detection algorithms: robust covariance[19] and local outlier factor[20]. N-neighbors parameter of the number of neighbors used for *k*-neighbors queries in local outlier factor algorithm is set to 5. The results are shown below. Anomalies number refers to the predicted number of anomalies in the data set. The definition of accuracy:

$$accuracy = \frac{True\ numbers}{True\ numbers + False\ numbers} \quad (1.5)$$

The definition of recall:

$$recall = \frac{True\ positive\ numbers}{True\ positive\ numbers + False\ negative\ numbers} \quad (1.6)$$

Tab.1 Experiment comparison of electrical energy data anomaly detection with 3 algorithms

Algorithm	Anomalies Number	Accuracy/%	Recall/%
One-class SVM	18	99.4169%	88.8889%
Robust covariance	13	99.2711%	72.2222%
Local outlier factor	34	93.2945%	16.6667%

As is shown in the table, one-class SVM model has a high accuracy and recall rate. Local outlier factor and robust covariance algorithm detect more incorrect abnormal points. Accuracy of one-class SVM is higher than other two algorithms. Recall of one-class SVM is 1.23 times that of robust covariance, and the recall rate is 5.33 times that of local outlier factor.

V.CONCLUSION

Based on the One-class SVM, this paper detect the outlier of the positive active power data. The extensive experiments are conducted to mine the anomaly data of large-scale electric energy, and the results shown that our method is very suitable for positive active power outlier instance detecting.

REFERENCES:

[1] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "Smart grid technologies: Communication technologies and standards," *IEEE transactions on Industrial informatics,* vol. 7, pp. 529-539, 2011.

[2] J. Liu, J. Liu, H. Chen, R. Huang, Z. Li, P. Liu, Y. Guo, J. Shen, and Y. Wang, "Abnormal energy identification of variable refrigerant flow air-conditioning systems based on data mining techniques," *Applied Thermal Engineering,* vol. 150, pp. 398-411, 2019.

[3] D. Jiang, Z. Yuan, P. Zhang, L. Miao, and T. Zhu, "A traffic anomaly detection approach in communication networks for applications of multimedia medical devices," *Multimedia Tools and Applications,* vol. 75, pp. 14281-14305, 2016.

[4] V. Chandola, A. Banerjee and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR),* vol. 41, pp. 1-58, 2009.

[5] W. Zengping, Z. Jinfang and Q. Cheng, "Topology error identification based on synchronized measurements for power network," *Electric Power Automation Equipment,* vol. 32, pp. 1-8, 2012.

[6] W. Alves, D. Martins, U. H. Bezerra, and A. Klautau, "A Hybrid Approach for Big Data Outlier Detection from Electric Power SCADA System," *IEEE Latin America Transactions,* vol. 15, pp. 57-64, 2017.

[7] I. Steinwart, D. Hush and C. Scovel, "A Classification Framework for Anomaly Detection," *Journal of Machine Learning Research,* vol. 6, pp. 211-232, 2005.

[8] L. Cai, J. Chen, Y. Ke, T. Chen, and Z. Li, "A new data normalization method for unsupervised anomaly intrusion detection," *Journal of Zhejiang University SCIENCE C,* vol. 11, pp. 778-784, 2010-01-01 2010.

[9] D. T. Bui, T. A. Tuan, H. Klempe, B. Pradhan, and I. Revhaug, "Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree," *Landslides,* vol. 13, pp. 361-378, 2016.

[10] S. Chen, W. Zhang and M. Peng, "Anomaly detection based on Non-negative matrix factorization in stock market," *Rev. Téc. Ing. Univ. Zulia,* vol. 40, pp. 9-18, 2016.

[11] A. C. Onal, O. B. Sezer, M. Ozbayoglu, and E. Dogdu, "Weather data analysis and sensor fault detection using an extended IoT framework with semantics, big data, and machine learning," in *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 2037-2046.

[12] Y. Yu, Y. Zhu, S. Li, and D. Wan, "Time Series Outlier Detection Based on Sliding Window Prediction," *Mathematical Problems in Engineering,* vol. 2014, pp. 1-14, 2014.

[13] S. Saliminejad and N. G. Gharaibeh, "Proximity-Based Outlier Detection Method for Roadway Infrastructure Condition Data," *Journal of Computing in Civil Engineering,* vol. 30, p. 04015001, 2016.

[14] F. Angiulli, S. Basta and C. Pizzuti, "Distance-based detection and prediction of outliers," *IEEE Transactions on Knowledge and Data Engineering,* vol. 18, pp. 145-160, 2006.

[15] V. M. Catterson, S. D. McArthur and G. Moss, "Online conditional anomaly detection in multivariate data for transformer monitoring," *IEEE Transactions on Power Delivery,* vol. 25, pp. 2556-2564, 2010.

[16] V. Jakkula and D. Cook, "Outlier detection in smart environment structured power datasets," in *2010 Sixth International Conference on Intelligent Environments*, 2010, pp. 29-33.

[17] N. L. Tasfi, W. A. Higashino, K. Grolinger, and M. A. Capretz, "Deep neural networks with confidence sampling for electrical anomaly detection," in *2017 IEEE International Conference on Internet of Things,Green Computing and Communications, Cyber, Physical and Social Computing,Smart Data*, 2017, pp. 1038-1045.

[18] D. M. Tax and R. P. Duin, "Support vector data description," *Machine learning,* vol. 54, pp. 45-66, 2004.

[19] Y. Chen, A. Wiesel and A. O. Hero, "Robust shrinkage estimation of high-dimensional covariance matrices," *IEEE Transactions on Signal Processing,* vol. 59, pp. 4097-4107, 2011.

[20] H. Kriegel, P. Kröger, E. Schubert, and A. Zimek, "LoOP: local outlier probabilities," in *Proceedings of the 18th ACM conference on Information and knowledge management*, 2009, pp. 1649-1652.