Nontechnical Loss Detection using Neural Architecture Search and Outlier Detection

Ke Fei[1], Qi Li[1*], Can Cui[2], Xue chen[2], Xinxin Xu[2], Benshan Xue[2], Weifeng Cai[2]

[1] Chongqing University, Chongqing, 400044, China
[2] State Grid Cooperation of China, Shang Qiu, Henan Province, 476000, China

Abstract. Electricity supply is essential to economy growth and improvement of people's life. For a long time, illegal electricity theft not only affects the supply of power, but also causes significant economic loss. Traditional techniques for detecting electricity theft are inefficient and time-consuming. Data-based detecting algorithms become a new solution. This article analyses the features of electricity consumption, current, voltage and opening records under various electricity theft modes and proposes a new simulation method for electricity theft users. Based on the simulation dataset, a feature extraction method based on neural architecture search (NAS) is proposed. The advantage of this feature extraction model is demonstrated in the comparison experiments with other feature extraction model. Finally, the effectiveness and accuracy of the electricity theft detection method based on NAS model and outlier detection are verified through an industrial case study.

1 Introduction

1.1. Problem Definition

Electricity supply is important for social economy and people's life. Electricity theft is the main cause of non-technical losses. For a long time, illegal electricity theft has not been eradicated, electricity theft not only causes huge economic losses to the country, but also affects the reliability of power supply. It is reported that the global annual loss caused by electricity theft is $96 billion 1. Traditional methods of checking for electricity theft are inefficient and time-consuming. In recent years, the Advanced Metering Infrastructure (AMI) has been built in the smart grid, and the large amount of data it collects provides the foundation for data-driven theft detection algorithms 2.

1.2. Related work

Data-based algorithms can be divided into two categories: supervised learning algorithms and unsupervised learning algorithms. Algorithms that use label information (information about user types obtained through manual checking) are called supervised learning, and methods that do not use label information are called unsupervised learning.

Salman et al. 3 used boosted C5.0 decision tree algorithm to classify normal customers and electricity theft customers which used electricity consumption data of 1,033,051 customers provided by Pakistan Electricity Corporation. The authors used Pearson's chi-squared test to select features. Nizar et al. 4 used extreme learning machine to reveal abnormal behavior highly associated with electricity theft. This method extracts customer behavior patterns from past electricity consumption and uses the extracted behavioral features to reveal whether the behavior is abnormal electricity consumption. Tacón et al. 5 use Transductive SVM (TSVM) to identify customers with abnormal electricity consumption. Punmiya 6 proposed the use of extreme gradient boosting (XGBoost), categorical boosting (CatBoost) and light gradient boosting method (LightGBM) algorithms to detect highly anomalous power consumption behavior.

For unsupervised learning algorithms, Spirić et al 7 used fuzzy logic to identify potential electricity theft customers. It used the electricity consumption data of customers from 2003 to 2017. It established fuzzy suspicion based on the relationship between electricity consumption data and time series data. Then the fuzzy logic is used to calculate the suspicious value of each electricity user. If the suspicious value exceeds a certain threshold, the customer is considered a suspected electricity user. Similar work was presented by Viegas et al 8, which used fuzzy-based distance to check whether the distance of electricity users has significantly exceeded the normal electricity user prototype. They used electricity consumption data from four thousand Irish households. Krishan et al 9 proposed an anomaly detection method with a unique combination of principal component analysis (PCA) and DBSCAN clustering to detect electricity theft. Ramos 10 proposed to use OPF (optimum-path forest) for electricity theft detection.

From the above literature review, it can be found that most of the data-driven based literature mostly uses only

[*] Corresponding author's e-mail: qi.li@cqu.edu.cn

customer electricity consumption as the data for extracting electricity theft characteristics. While smart meters can collect various electrical data, such as voltage, current, opening records, etc., other data are less explored. When building a model, using the appropriate features as input has a significant impact on both accuracy and robustness. It is common practice to use unsupervised algorithms to perform feature extraction and then to use these features for electricity theft detection. However, most feature extraction algorithms are only applicable to a single data type and dimension (one-dimensional or two-dimensional electricity consumption data), and lack optimization.

1.3. Contribution

Our contribution to the issues summarized above is as follows:

(i) In this paper, we analyze various ways of electricity theft at the physical device level. Based on the analysis, electricity theft simulations are performed on the power consumption, current, voltage and opening records.

(ii)The feature extraction model is automatically constructed and optimized using neural architecture search (NAS).

(iii)Based on the extracted features, the performance of four outlier detection algorithms on electricity theft detection is compared and analyzed.

2 Methods

2.1 Analysis of electricity theft patterns

Based on the data provided by the State Grid, electricity theft can be classified into three categories: attacking meters to steal, changing connections to steal and unmetered electricity theft. Since the electricity consumption data of unmetered theft cannot be collected, data analysis cannot be performed. In this paper, we focus on two types of theft: attacking meters to steal and changing the connection to steal.

The attack on the meter is mainly through the modification of the circuit to reduce the current and voltage so that the electricity is less counted. Since the meter uses the phase line current to calculate power, modifying the current sampling circuit will cause an imbalance between the phase line current and the neutral line current. The electricity thief needs to open the meter cover in order to modify the circuit, so there will be an opening record. After the modification of the circuit is completed, the user will remain in a continuous state of electricity theft.

For changing the connection to steal electricity, it is mainly by making the phase line current measurement wrong or inaccurate, or lowering the measured voltage value to make the meter record value smaller. Except for the way of attacking secondary circuit which will have opening record, several other ways have no opening record. For borrowing neutral line theft, disconnecting neutral line theft and energy-saving device theft, as all add a controller that can switch between theft and non-theft, its theft has an intermittent nature.

The characteristics of electricity consumption, voltage, current and opening records in various electricity theft modes are shown in Table 1.

Table 1. Electricity theft patterns

Modes	Consumption	Current	Voltage	Opening record	Changing Pattern
Modify meter parameters	Decline	Normal	Normal	No	Continuous
Borrowing neutral line	Zero	Imbalance	Normal	No	Intermittent
Energy saving	Decline	Normal	Normal	No	Intermittent
Modification of current sampling circuit	Decline	Imbalance	Normal	Yes	Continuous
Modification of voltage sampling circuit	Decline	Normal	Decline	Yes	Continuous
Breaking neutral line	Decline	Normal	Decline	No	Intermittent
Destruction of the rear partition	Decline	Imbalance	Normal	No	Continuous
Attacking the secondary circuit	Zero	Normal	Normal	Yes	Continuous

2.2 Simulation of electricity theft patterns

The dataset used in this paper consists of 48 days of electricity consumption data from 5,000 normal customers, provided by the ISSDA11. 500 customers are randomly selected for the electricity theft simulation. The customer data is a tensor of shape (500, 48, 48, 4), where 500 is the number of customers, the first 48 is 48 days, the last 48 is 48 half hours per day, and 4 is 4 characteristics information, which are electricity consumption, phase neutral current abnormality degree, voltage abnormality degree and opening record

Table 2 shows the simulation algorithm for energy saving device theft, and the other theft modes are similar.

Table 2. Borrow neutral line

```
Borrow neutral line theft simulation:

Input: $X = (X_1, X_2, X_3, X_4)$ ,
($X_i \in \mathbb{R}^{48\times48}$, $X_2 = X_3 = X_4 = 0$)
$For\ k\ from$ 1 $to$ 48,  do
    $X_1[k, m : n] \leftarrow$ random(0， 1) $\times X_1[k, m : n]$
    $X_2[k,\ m : n] \leftarrow$ random(0， 1) $\times 1$
    ($m\ and\ n\ are\ random$)
$V_1 \leftarrow X_1$
$V_2 \leftarrow X_2$
$V_3 \leftarrow X_3$
$V_4 \leftarrow X_4$
Return $V$= $(V_1, V_2, V_3, V_4)$
```

Where X represents a normal user data, which is a 48×48×4×1 tensor, and V represents the user data after electricity theft simulation.

2.3 Feature extraction model by neural architecture search

The data feature extraction model is automatically generated and optimized using NAS. The NAS framework used in this paper is Auto-Keras 12. Auto-Keras utilizes Bayesian optimization to guide architectural changes. Bayesian optimization is used mainly to reduce the amount of computation required for architecture search and to improve the model performance. The algorithm ultimately yields the optimal model structure. This optimal model is then used to perform feature extraction.

The feature extraction model built in this paper using neural architecture search is similar to the Auto-Encoder 13. The main difference is that the entire feature extraction model structure is generated by a neural architecture search algorithm. The training process of the feature extraction model is shown in Figure 1.

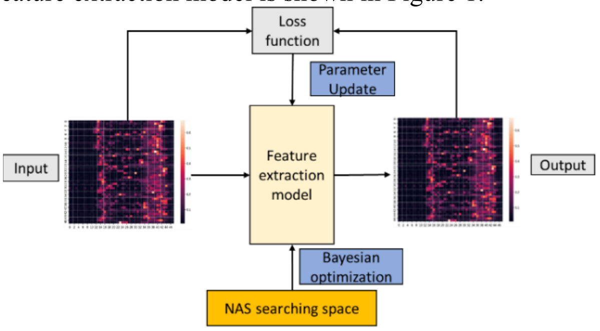

Figure 1. Feature extraction model training process

As shown in Figure 1, the feature extraction model uses the input as label information. The neural architecture search algorithm first picks an architecture from the search space based on Bayesian optimization and then keeps changing the architecture of the model until the loss value of the architecture is minimized. Since the structure of the model is variable and the input can be a tensor, it can be compatible with multiple dimensions of data.

2.4 NTL Detection workflow

The implementation of the NTL detection algorithm optimized using NAS is shown in Figure 2. Because the feature extraction model is a multi-layer structure, the output of one of the layers can be taken as features.

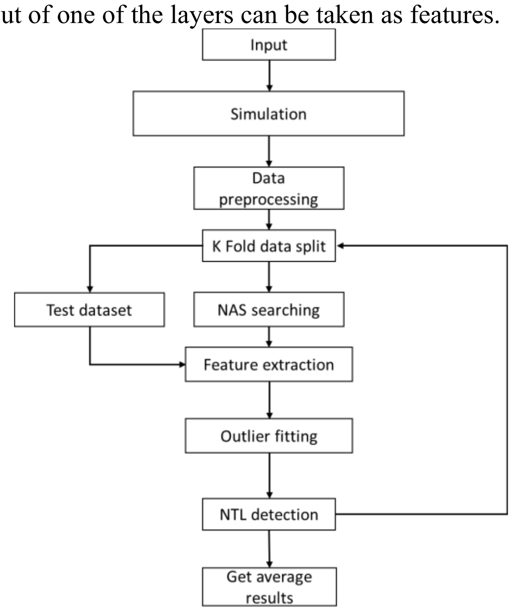

Figure 2. NTL detection flow

3 Results & Discussion

3.1 Metrics

This paper adopts Precision, Recall and F1 score as the evaluation indexes, which are defined as follows.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall \ = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = \frac{2 \times \ Precision \ \times \ Recall}{Precision \ + \ Recall} \tag{3}$$

Where TP is the correctly classified electricity theft customer, FN is the incorrectly classified electricity theft customer, and FP is the incorrectly classified normal customer.

3.2 Results and discussion

Based on the outlier algorithm, the Kernel PCA 14 feature extraction algorithm, the Auto-Encoder 13 feature extraction algorithm and the NAS-based feature extraction algorithm are compared. The experimental results are shown in Table 3, where the percentage of outlier division is 9% and the outlier detection algorithm is a clustering-based local outlier detection algorithm.

Table 3. Feature extraction algorithm results

Feature extraction	Precision	Recall	F1
NAS	**0.88**	**0.8712**	**0.8756**
Kernel PCA	0.7857	0.8461	0.8148
Auto-Encoder	0.6224	0.753	0.6815

The NAS-based outlier algorithm achieves 88.00% Precision, 87.12% Recall, and 0.8756 F1 score, while the Kernel PCA-based model achieves 78.57% precision, 84.61% recall, and 0.8148 F1 score, respectively. It can be seen that the NAS-based feature extraction algorithm outperforms the Kernel PCA-based feature extraction algorithm and the Auto-Encoder based feature extraction algorithm. The reason is that the NAS algorithm can optimize the parameters according to different data characteristics and has better nonlinear feature extraction capability.

Based on the NAS feature extraction model, the paper compares the clustering-based local outlier detection algorithm (CBLOF) 15, the histogram-based outlier detection algorithm(HBOS)16, the angle-based outlier detection algorithm (ABOD) 17, and the isolation forest algorithm (iForest) 18.

Table 4. Results for four algorithms with different dividing points

Algorithm	Dividing points	Precision	Recall	F1
CBLOF	8 %	**0.9101**	**0.8019**	**0.8526**
	9 %	**0.88**	**0.8712**	**0.8756**
	10 %	**0.8378**	**0.9207**	**0.8773**
ABOD	8 %	0.594	0.594	0.594
	9 %	0.5535	0.6138	0.5821
	10 %	0.5271	0.6732	0.5913
HBOS	8 %	0.8764	0.7722	0.821
	9 %	0.82	0.8118	0.8159
	10 %	0.7567	0.8316	0.7924
iForest	8 %	0.8651	0.7623	0.8105
	9 %	0.81	0.8019	0.8059
	10 %	0.7657	0.8415	0.8018

As can be seen from Table 1, CFLOF (9%) has a Precision of 88.00%, Recall of 87.12%, and F1 score of 0.8756, which is the optimal performance compared to the other three outlier detection algorithms. When the division point is 8%, the Precision of CBLOF reaches 91.01%, Recall reaches 80.19%, and F1 score is 0.8526, which can identify theft more accurately, but the detection rate of theft customers decreases. When the division point is 10%, Precision is 83.78%, Recall is 92.07%, and F1 score is 0.8773. The precision decreases and the recall increases. The choice of division points should be based on the actual situation. The division points can be increased when it is necessary to completely exclude customers who steal electricity, and reduced when it is necessary to save the cost of manual inspection.

4 Conclusion

The paper finds that phase and neutral current imbalance, voltage reduction and opening records are highly correlated with power theft. And based on the analysis, a new simulation method of power theft is proposed.

The paper compares the NAS-based feature extraction model with the Kernel PCA feature extraction model and the Auto-Encoder feature extraction model. The NAS-based outlier algorithm achieves 88.00% Precision and 87.12% Recall, which are significantly better than Kernal PCA and Auto-Encoder models. Based on the NAS model, the clustering-based local outlier detection algorithm, the histogram-based outlier detection algorithm, the angle-based outlier detection algorithm, and the isolation forest algorithm are compared. The results show that the clustering-based local outlier detection algorithm is the best.

How to use neural architecture search to build a superior feature extraction model and improve the efficiency of neural architecture search becomes the research direction afterwards.

Acknowledgement

The Authors are grateful to State Grid Shang Qiu Electric Power Company for their support for this work.

References

1. 96 Billion is Lost Every Year to Electricity Theft. Accessed: May 8, 2017. [Online]. Available at: https://www.prnewswire.com/newsreleases/96-billion-is-lost-every-year to-electricity-theft-300453411.html
2. Chen Z, Meng D, Zhang Y, et al. Electricity Theft Detection Using Deep Bidirectional Recurrent Neural Network[C]//2020 22nd International Conference on Advanced Communication Technology (ICACT). IEEE, 2020: 401-406.
3. Salman Saeed M, Mustafa M W, Sheikh U U, et al. An efficient boosted C5. 0 Decision-Tree-Based classification approach for detecting non-technical losses in power utilities[J]. Energies, 2020, 13(12): 3242.
4. Nizar A H, Dong Z Y, Wang Y. Power utility nontechnical loss analysis with extreme learning machine method[J]. IEEE Transactions on Power Systems, 2008, 23(3): 946-955.

5. Tacón J, Melgarejo D, Rodríguez F, et al. Semisupervised approach to non technical losses detection[C]//Iberoamerican Congress on Pattern Recognition. Springer, Cham, 2014: 698-705.
6. Punmiya, R., & Choe, S. (2019). Energy theft detection using gradient boosting theft detector with feature engineering-based preprocessing. IEEE Transactions on Smart Grid
7. Spirić J V, Stanković S S, Dočić M B. Identification of suspicious electricity customers[J]. International Journal of Electrical Power & Energy Systems, 2018, 95: 635-643.
8. Viegas J L, Esteves P R, Vieira S M. Clustering-based novelty detection for identification of non-technical losses[J]. International Journal of Electrical Power & Energy Systems, 2018, 101: 301-310.
9. Krishna V B, Weaver G A, Sanders W H. PCA-based method for detecting integrity attacks on advanced metering infrastructure[C]//International Conference on Quantitative Evaluation of Systems. Springer, Cham, 2015: 70-85.
10. Ramos C C O, de Sousa A N, Papa J P, et al. A new approach for nontechnical losses detection based on optimum-path forest[J]. IEEE Transactions on Power Systems, 2010, 26(1): 181-189.
11. ISSDA. Data from the Commission for Energy Regulation – <http://www.ucd. ie/issda>.
12. Jin, H., Song, Q., & Hu, X. (2019). Auto-Keras: An Efficient Neural Architecture Search System. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.
13. Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]//Proceedings of the 25th international conference on Machine learning. 2008: 1096-1103.
14. Hoffmann H. Kernel PCA for novelty detection[J]. Pattern recognition, 2007, 40(3): 863-874.
15. He, Z., Xu, X. and Deng, S., 2003. Discovering cluster-based local outliers. Pattern Recognition Letters, 24(9-10), pp.1641-1650.
16. Goldstein, M. and Dengel, A., 2012. Histogram-based outlier score (hbos): A fast unsupervised anomaly detection algorithm. In KI-2012: Poster and Demo Track, pp.59-63.
17. Kriegel, H.P. and Zimek, A., 2008, August. Angle-based outlier detection in high-dimensional data. In KDD '08, pp. 444-452. ACM.
18. Liu, F.T., Ting, K.M. and Zhou, Z.H., 2008, December. Isolation forest. In International Conference on Data Mining, pp. 413-422. IEEE.