Chapter 15
Research on Identification Model of Special Transformer Stealing Electricity Based on Vector Similarity Matching Algorithm

Xin Xuan, Yazhuang Cao, Shuming Wang, and Tingyan Jiang

Abstract In today's society, the act of stealing electricity not only loses a large amount of electricity fee income, but also brings property losses to the state grid, and at the same time, it is extremely easy to lay hidden security risks. The characteristics of the means of stealing electric power have shown the development from the private construction to the specialization, concealment and networking. Internet big data, this paper "precise marketing" thinking, according to the special transformer users power samples form special transformer power fingerprint, and special transformer users of electricity characteristic signal form of special transformer power use fingerprint vector similarity comparison, summed up the special transformer power matching algorithm based on vector similarity identification model. In practice, the algorithm is efficient, simple and accurate, and can effectively assist the inspectors to identify electric larceny.

15.1 Introduction

With the development of State Grid Information Construction and the popularization of smart meters, a large number of users' real-time data such as voltage, current and power have been accumulated. How to accurately identify and locate the suspected power stealing users through the abnormal power consumption signals in the user's electrical signal data is a research hotspot at home and abroad [1, 2]. At present, a large number of professionals has carried out research on electricity theft identification, but there are still some problems such as low practicability, low accuracy and low-calculation efficiency. Based on this, the research on the identification model of electric larceny of special transformer based on vector similarity matching algorithm proposed in this paper can be used to establish the power stealing identification model of special transformer based on the existing multi-source data of marketing specialty by applying the method of big data mining. By using the identification model of

X. Xuan · Y. Cao (✉) · S. Wang · T. Jiang
Beijing CHINA-POWER Information Technology Co., Ltd. State Grid Information & Telecommunication Group, Beijing 100085, China
e-mail: cyaz011@163.com

L. C. Jain et al. (eds.), *Smart Communications, Intelligent Algorithms and Interactive Methods*, Smart Innovation, Systems and Technologies 257,
https://doi.org/10.1007/978-981-16-5164-9_15

transformer stealing electricity, the abnormal situation of voltage, current and power of users can be monitored, and various means of stealing electricity such as private capacity increasing and bypassing measurement can be identified, and the suspected power stealing users can be found.

15.2 Analysis of Power Stealing of Special Transformer

15.2.1 Summary of Power Stealing Methods for Special Transformer

There are many ways to steal electricity from special transformer users. The essence of stealing electricity is to avoid measurement and reduce measurement by changing the metering device. The main idea of stealing electricity is to start from the formula of electric quantity measurement: as shown in formula 15.1, the power stealing user can change the parameters such as voltage, current and power factor to achieve the goal of no or less measurement of electric quantity and complete the theft In addition, some electricity stealing users steal electricity through direct off meter wiring to avoid measurement and special transformer capacity increase. To sum up, the power stealing methods of special transformer users are summarized as follows: (1) voltage loss method, (2) under current method, (3) differential expansion method, (4) phase shift method and (5) no meter method [3].

$$Q = U \times I \times \cos\varphi \times T \tag{15.1}$$

Formula 15.1: electricity metering formula.

Note: Q—power consumption, U—voltage, I—current, cos φ—power factor, T—time.

15.2.2 Summary of Power Consumption Characteristic Signals of Special Transformer Users

At present, a provincial power company of State Grid can collect the electricity consumption signals of special transformer users, such as electric quantity, voltage, current, power, power factor, etc., and the acquisition frequency is once per hour. Using big data thinking, through the electric quantity interface table, voltage interface table, current interface table and power interface table, these characteristic signals of electricity consumption are summarized to form a "power consumption profile," and the judgment conditions of each feature vector are used to scan the "power consumption fingerprint database," which can further effectively judge whether the user has abnormal electricity consumption behavior.

The characteristics of normal electrical signals of users are as follows:

(1) Voltage: the voltage value and the power supply voltage value are basically equal, and the value remains basically unchanged.
(2) Current: the current value changes with the change of external load, but the three-phase current basically keeps balance.
(3) Power factor: the power factor value is basically stable, and the phase angle is stable.
(4) Line loss: the line loss value will fluctuate within a certain range, generally less than 7%.

The power, voltage, current, power, power factor and other electrical signals of special transformer users will have certain changes in their daily use. This change has nothing to do with whether the user has electricity stealing behavior. However, if the change of power consumption signal exceeds a certain range of values, we can judge it as abnormal power consumption and form a "power fingerprint database." Through further analysis, we can judge the power consumption abnormality know whether it has stealing electricity.

15.3 Current Situation of Electricity Stealing Samples of Special Transformer Users

In many references [4–7], the traditional data mining methods are described in detail, and the process basically follows the crsip-dm process. However, according to the current situation of Zhejiang electric power company of State Grid, there are few electricity stealing samples of special transformer users. If the traditional data mining algorithm training model is used to identify the power stealing of special transformer users, there will be some inevitable problems, such as the construction of the training set, the model's recall and precision rate.

Therefore, the idea of "user portrait" based on the "precision marketing" thinking of Internet big data [8, 9] and the idea of "user portrait" based on the characteristics of certain users' online shopping behavior can also be applied in the field of anti-power stealing. Through the analysis of multi-source data of marketing specialty, this paper extracts the characteristic vector and vector threshold of electricity stealing behavior of national historical electricity stealing users, describes the "user portrait" and "stealing fingerprint" of suspected electricity stealing, realizes the accurate positioning of suspected users, and forms the fingerprint database of stealing electricity.

Due to the small number of electricity stealing samples of special transformer, it is particularly important to summarize the electricity consumption characteristics of special transformer users for small sample analysis. By summarizing the existing electricity stealing samples, the power stealing fingerprints of special transformer users are formed, as shown in Fig. 15.1.

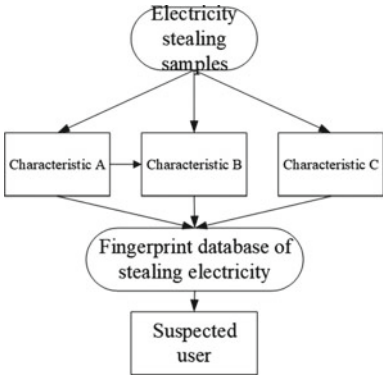

Fig. 15.1 Identification process of stealing electricity based on small sample

15.4 Power Stealing Identification of Special Transformer Based on Vector Similarity Matching Algorithm

15.4.1 Vector Similarity Comparison Method

The definition of similarity measure is to compare the degree of similarity between individuals. If the value of similarity measure is larger, the similarity between individuals will be greater; otherwise, it will be opposite. In a two-dimensional space right angle, the meaning of cosine similarity is: the similarity between two vectors $a(x_1, y_1)$ and $b(x_2, y_2)$ is expressed by the cosine value of the angle between vectors. When the cosine value of the included angle is closer to the value of 1, the angle between the two vectors is closer to 0 degree, which means that the two vectors are more similar, as shown in Fig. 15.2.

The included cosine formula of any pair of vectors in two-dimensional space is shown in Formula 15.2.

$$\cos(\theta) = \frac{a \cdot b}{|a| * |b|} = \frac{(x_1, y_1) \cdot (x_2, y_2)}{\sqrt{x_1^2 + y_1^2} \times \sqrt{x_2^2 + y_2^2}} = \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2} \times \sqrt{x_2^2 + y_2^2}} \quad (15.2)$$

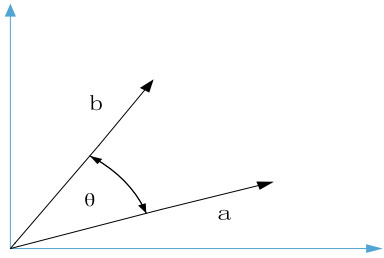

Fig. 15.2 Comparison of similarities between vectors a and b

Formula 15.2: Formula for calculating the included cosine of two-dimensional space vector.

Note: a and b represent two different vectors, (x_1, y_1) and (x_2, y_2) are the coordinates of vector a and b, respectively.

The n-dimensional space vector can be inferred from the two-dimensional space vector. Assuming that a and b are two vectors in n-dimensional space, the cosine formula for the angle between alpha and beta can be obtained as shown in Formula 15.3.

$$\cos(\theta) = \frac{a \cdot b}{|a| * |b|} = \frac{\sum_{i=1}^{n}(x_i \times y_i)}{\sqrt{\sum_{i=1}^{n}(x_i^2)} \times \sqrt{\sum_{i=1}^{n}(y_i^2)}} \tag{15.3}$$

Formula 15.3: Formula for calculating included cosine of n-dimensional space vector.

Note: a and b are two vectors in n-dimensional space, $(x_1, \ldots, x_n)$ and $(y_1, \ldots, y_n)$ are the coordinates of vector a and b, respectively.

15.4.2 Identification of Power Theft by Special Transformer Users

Summarize the samples of power theft in history, form the special transformer user power theft fingerprint library, set it as the reference standard, and transform the user power signal to be detected into the power vector with the same dimension as the special transformer power theft fingerprint library to detect whether the user has power theft at a certain time. The specific identification process is shown in Fig. 15.3. Read several electric characteristic vectors in electric quantity interface table, voltage interface table, current interface table and power interface table to form user's electric picture. Scan user's electric picture into special transformer's electric fingerprint database by threshold value of electric anomaly signal index of special transformer user, form special transformer's electric fingerprint database and compare it with special transformer's electric stealing fingerprint database for vector similarity. Match, complete the algorithm and output the list of suspected power theft from special transformers.

15.4.3 Algorithmic Process Analysis

By matching the vector similarity between the "power fingerprint library" and the "power theft fingerprint library," the power user can be judged whether there is power theft. The special transformer power theft identification process based on the vector similarity matching algorithm is as follows:

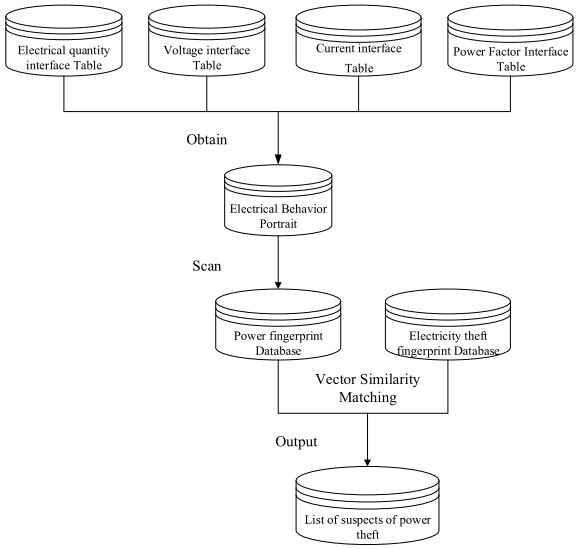

Fig. 15.3 Identification of power theft by dedicated transformer users based on vector similarity

(1) Summarize the existing sample of power theft by users and form special transformer power theft fingerprint library;
(2) Collect the user's power consumption data and form the user's power consumption picture;
(3) Construct user's electric picture into special transformer power fingerprint database by threshold value of user's electric abnormal signal index of special transformer;
(4) Carry out vector similarity calculation between special transformer power fingerprint database and special transformer power theft fingerprint database;
(5) Output the result to the suspect scoring table to form the suspect list of power theft of special transformer.

15.4.4 Application Results of Algorithmic Model

Through field verification in Zhejiang Province, it is shown that the special transformer power theft identification model based on vector similarity matching algorithm can detect abnormal power consumption behaviors of users in time and accurately identify power theft users. The algorithm is efficient, concise and accurate, and can effectively assist inspectors in identifying power theft.

15.5 Conclusion

The identification model of special transformer power theft based on vector similarity matching algorithm can effectively solve the problems of less samples of power theft from special transformer users. At the same time, the algorithm is simple and easy to understand, which can effectively improve the calculation efficiency and facilitate operator to understand and optimize parameters. The following research directions can be carried out: (1) adding feature components and describing power images in detail; (2) optimizing threshold parameters of users' power anomaly signal indicators; (3) expanding special transformer power theft fingerprint library to enhance the recognition of different power theft methods by algorithm. Improve the convenience and accuracy of model identification for power theft.

References

1. Yadi, T.: Research on anti-theft technology application of electric power information acquisition system. Mater. Decoration **12**(42), 211–212 (2016)
2. Bo, L., Ning, G., Hongmiao, C.: Anti-anti-theft scheme for power supply enterprises. Rural Electrician **26**(4), 48 (2018)
3. Zhongjun, M., Yunruo, X.: Discussion on electricity theft and anti-electricity theft in new form. Commun. World **25**(5), 244–246 (2018)
4. Junxing, X., Chuan, L., Yingna, L.: Electricity theft analysis based on BP neural network optimized by genetic algorithm. Software **38**(11), 18–23 (2017)
5. Hu, Z.: Cluster analysis of electric energy metering data and research on detection of power theft. China (2017)
6. Chuan, L.: Intelligent Clustering Analysis and its Application. Science press, China (2016)
7. Salini, B., Deepti, M.: Applying CHAID algorithm to investigate critical attributes of secured interoperable health data exchange. Int. J. Electron. Healthc. **8**(1), 25–50 (2015)
8. Hongyan, Z.: Discussion on application of power supply marketing service based on big data. Electromech. Inf. **18**(15), 169–171 (2018)
9. Zeying, H., Jiachen, X.: Precision marketing of big data—a tool for efficient marketing. Modern Mark. (Operating Version) **24**(6), 101 (2016)