Theories, Applications and Trends of Non-technical Losses in power utilities using Machine Learning

Tiantian Zhang[1,2] ,Rongfang Gao[1,2] , Shaohua Sun [1,2]
1. College of Computer Science and Technology, Xi'an Shiyou University
2. Xi'an, China
ttzhang@tulip.academy, gaorf@xsyu.edu.cn, shhsun@tulip.academy

***Abstract*—With the popularization of smart meter and electric information collecting system, the application of machine learning is also more widely in non-technical loss (NTL) detection. In this paper, the concept and solution of NTL is introduced. Among them, the contributions are the application of machine learning and concluding privacy-preserving methods in NTL detection. This paper introduces the application of data acquisition, data pre-processing and data mining modeling in NTL detection from the perspective of data mining, and sums up the applicable scenarios and characteristics of each method. Next, it lists the potential risks caused by privacy leakage of power users, as well as the privacy protection methods currently proposed in NTL detection, and analyzes these method. Finally, the current status of NTL detection is summarized, and the future research trends of NTL detection in feature processing, real-time monitoring, large-scale data sets and other issues are discussed.**

Keywords—non-technical loss; machine learning; privacy-preserving; electricity theft

I. Introduction

In the smart grid environment, with the popularization of smart meter and user information collection system, it conduces to collect user's electricity consumption data more efficiently and accurately, which is the foundation of NTL detection using machine learning. There are two types of power loss: Technical Loss (TL) and Non-technical Loss (NTL). TL occur naturally and are caused because of power dissipation in transmission lines, transformers, and other power system components. NTL are principally caused by fraud activities deliberately performed by the consumers, and lead to a series of additional losses, including damage to grid infrastructure and reduction of grid reliability [1]. TL is unavoidable, but we can reduce NTL through appropriate means. This section that introduce basic concept, causes, effects of NTL, aim to having a preliminary cognition.

A. The Concept of NTL

Generation, Transmission, and Distribution of electricity involve many losses. Generally, power-related losses are technically easy to measure, but transmission and distribution (T & D) losses are difficult to be quantified accurately. This illustrates the involvement of some non-technical parameters in T&D of electrical energy. Non-technical losses are defined as losses that occur outside of technical losses in T&D of electrical energy [2]. Under normal circumstances, the electrical energy generated by the power company should equal the energy consumed. However, the actual situation is different, because the energy loss is a comprehensive result of T&D. Davidson defined these energy losses in the following equations [3-6].

Energy Losses

$$E_{Loss} = E_{Delivered} - E_{Sold} \tag{1}$$

Revenue Loss due to technical losses

$$C_{Loss} = U_{ElectricityCost} \times E_{Loss} + M_{MaintenanceCost} \tag{2}$$

Non-technical Loss

$$C_{NTL} = C_{Loss} - C_{TechnicalLoss} \tag{3}$$

In formula (1), E_{Loss}, $E_{Delivered}$, E_{Sold} represent energy losses, delivered energy, sold energy respectively. In formula (2), C_{Loss}, $U_{ElectricityCost}$, $M_{MaintenanceCost}$ represent revenue loss due to technical losses, electricity cost, maintenance cost respectively. In formula (3), C_{NTL}, $C_{TechnicalLoss}$ represent non-technical loss, technical loss.

B. Causes of NTL

NTLs mainly consist of power theft and customer management processes in which there exist a number of means of consciously defrauding the utility concerned. NTLs include the following activities [7].

(1) Tamper with meters to make it record a lower consumption reading.

(2) Using electrical energy bypass meters or otherwise making illegal connections.

(3) To collude with internal employees to make billing irregularities, such as opening a lower-cost bill and changing the reading of the meters.

The first two behaviors are more common in the case of most electricity anomalies. Electricity theft is an important cause of NTLs. The detection of electricity theft can help power companies to reduce NTLs. the power measurement in the power calculation formula can be represented as follows:

$$\mathrm{P} = \mathrm{UI}\cos\varphi \tag{4}$$

In formula (4), P, U, I, φ represent power, voltage, current, power factor respectively.

According to formula (4), meter measurement is decided by power, voltage and current. Power thieves achieve their aim by changing any one of three conditions. Additionally, changing the structure of meters is also a method. Therefore, there are several common ways to steal electricity [8]:

978-1-5386-1803-5/18/$31.00 ©2018 IEEE

(1) Under-current method: through a variety of ways to make the current electricity through meters smaller or even 0, so that meters to record a lower reading.

(2) Under-voltage method: through a variety of ways to make voltage coil of voltage meter loses voltage or reduce the voltage, so that meters to record a lower reading.

(3) Difference expansion method: through expanding the error of the meter, so as to achieve the purpose of electricity theft. When using this method to electricity theft, the internal structure of meters needs to be changed or to destroy it by other means.

(4) Phase-shift method: through changing the normal phase relationship between the voltage and current in the meter, which causes the meter to slow down or even reverse, so that less power can be recorded.

(5) Strong AC magnetic field method: Because some electronic meter uses the stepper motor to count, and the stepper motor counter in the strong alternating magnetic field will automatically count, change the direction of the magnetic field, the counter will quickly down count, so will result in less electricity, so as to implement electricity theft.

(6) No meter method: connect wires to use electricity energy without the electric company's authorization, or bypass the meter to use it.

With the popularization of smart meters, the means of electricity theft are also improving, such as high-frequency and high-voltage electricity theft, high-power wireless signals. High-frequency and high-voltage power theft is through interference with the meter's internal workflow, destroying the working curve of the meter, making meters reducing accuracy, thereby can't work normally [9-10]. High-power wireless signal stealing electricity is interference with the meter CPU, making the meter less or not count. This method can also restore metering at any time. Because these high-tech means of electricity theft are relatively invisible and difficult to be detected, and their operation time is short, which has brought great difficulties to detect electricity theft.

C. Effects of NTL

The impact of NTL is mainly reflected in the following aspects:

(1) Lead to voltage overload, affecting performance or even damage to equipment.

(2) Large amounts of non-technical damage may result in tripping, resulting in disruption to the customer's power supply.

(3) Unpredictable additional load caused by NTL may result in power outage and blackout during peak periods.

(4) In the energy market, the NTL has caused huge economic losses to the power companies, forced by the surviving utility companies to recover the costs, so they have to impose excessive tariffs on all consumers, but this is actually for the honest consumers It is extremely unfair.

(5) Illegal electricity theft also easily lead to security risks such as electric shock, and even cause death.

(6) Improper operation of distribution feeders may put the entire community at risk because these wires may spark and cause fire in extreme weather conditions.

Next, this paper will presents a rigorous study about detection techniques using machine learning, and methods of privacy-preserving in power utilities through five sections. Among them, section 2 is data acquisition, which introduces the source and composition of power data in NTL detection. Section 3 is data pre-processing, which describes importance and methods of feature selection in pre-processing. Section 4 is data mining modeling, which analyses relative methods of NTL detection from supervised learning, unsupervised learning. Finally, Section 5 conclude the paper, and proposed the scope of future work.

II. Data Acquisition

Generally, through the power metering automation system, power information collection system and other power metering system can be collected to the phase current, voltage, power factor, other load data and terminal alarm information. Abnormal alarm information and electricity load data can reflect the user's electricity situation, while the inspection staff will also find out electricity thieves through the online inspection system and on-site inspection, then record in the system.

The original data related to electricity theft are mainly power load data, terminal alarm data, default information and user file information, etc. The data including characteristic of electricity theft can usually be classified in two aspects:

(1) According to user data: the main user name, user ID, power address, power type, power consumption, measurement methods.

(2) According to data from the metering terminal: There are mainly time points, measurement points, electricity load, power consumption, line loss rate, power factor, alarm time and so on.

Among them, start and end time that power user implements illegal activities are the key time to characterize electricity thieves. In these time nodes, electricity load and terminal data will have some distinctive changes, so the sample data acquisition needs contain a range of data before and after critical time nodes.

III. Data Pre-Processing

In the machine learning field, a data pre-processing procedure is often needed to eliminate errors, noises, in-consistent data or missing data that are contained in the dataset. After dealing with missing values, data transformations, etc., we need to select meaningful features from the current dataset and input them into the model for training. Feature selection of the data set is mainly based on the following considerations:

(1) Redundant features will affect the potential rules that hinder the model from finding data. If there are many redundant features, it will also cause disaster of dimensionality and take up a lot of time and space, greatly reducing the efficiency of the algorithm.

(2) Removing irrelevant features will reduce the difficulty of learning tasks. Retaining the key features can find out potential laws intuitively.

Feature selection is also called feature subset selection, or attribute selection, which refers to selecting a subset of features from all the features to improve the performance of learning algorithm [11]. A typical feature selection process consists of four basic steps, namely subset generation, subset evaluation, stopping criterion and result validation [12], as described below:

(1) Subset generation: the process of gene-ration is a process of searching for a subset of features and is responsible for providing a subset of features for the evaluation function.

(2) Subset evaluation: evaluation function is a criterion to evaluate the quality of a feature subset.

(3) Stopping criterion: The stopping criterion is related to the evaluation function, which is generally a threshold. When the evaluation function reaches this threshold, the search will be stopped.

(4) Result validation: verify the validity of the selected subset of features on the validation dataset.

According to combination mode of subset evaluation and follow-up learning algorithm, feature selection can be divided into the following three categories:

(1) Embedded method: the main idea of embedded method is to learn the best attributes to improve the accuracy of the model when the model is established. In the embedded method, the feature selection algorithm itself is embedded into the learning algorithm as an integral part of the learning algorithm. The most typical is the decision tree algorithm, such as Quinlan's ID3, C4.5 and Breiman's CART algorithm. This kind of classifiers determine branch nodes with choosing the most effective feature to partition the data set. Earlier Filho et al. used the decision tree to identify the power company's customers for fraud detection, which greatly reduced the inspection cost [13].

(2) Filter method: the main idea of the filtering method is to score the features of each dimension, which is give weights to the features of each dimension, and then sort them according to weights. Most of the early feature selection algorithms belong to the filtering method. The evaluation criteria of the filtering method are obtained from the intrinsic properties of the data set itself, and unrelated to the specific learning algorithm, so it has good versatility. Researchers think that the more relevant features or feature subsets will get higher accuracy in the classifier. Monedero et al. used Pearson correlation coefficient to detect the typical NTL characterized by load sag [14]. The advantage of this method is that it eliminates the training steps of classifiers and has low complexity. Therefore, it is suitable for large-scale data sets and can quickly remove a large number of irrelevant features. It is very suitable as preliminary filter for features.

(3) Wrapper method: the main idea of the wrapper method is to consider the selection of subsets as a search optimization problem, which generates different combinations, then evaluates the combinations, compares with other combinations finally. NAGI proposed GA-SVM detection algorithm based on support vector machine and genetic algorithm, which detected power theft by the global optimal parameters given by genetic algorithm [15-16]. Ramos et al. proposed modeling optimization problems based on natural phenomena, and combined Optimum-Path Forest (OPF) algorithm with Particle Swarm Optimization (PSO) algorithm [17-18], harmony search (HS) algorithm [19-20] and Gravitational Search Algorithm (GSA) [21-22], we found that PSO-OPF performs the best in accuracy and HS-OPF is the fastest in detection speed [23-24]. Recently, Pereira et al. introduced the social-spider optimization (SSO) algorithm to solve the optimization task considering male and female spiders' cooperative behavior, and combined SSO and SVM to solve power theft [25].

Generally speaking, embedded methods are optimized in the local space, and its effect is relatively limited. Compared with the filtering method, the wrapper method usually finds feature subset that is the better classification performance. However, it also has the disadvantage that the feature selected by the wrapper method is not universal. When changing the learning algorithm, it is necessary to reselect features by this learning algorithm. Additionally, every time we evaluate the subset that needs to train and test the classifier, so the computational complexity of the algorithm is high. Especially for the large-scale data set, the execution time of the algorithm is longer.

IV. Data Mining Modeling

At present, NTL detection methods in the field of machine learning can be divided into two categories:

(1) Methods based on supervised learning: Essentially, it's approaches based on classification that are given the type of electrical activity of power users (normal or abnormal), and usually apply to static data sets and to dynamic data with very minimal changes. If the characteristic of the data changes greatly, the original classification model won't reflect the normal or abnormal behavior of the dataset. Therefore, it's often necessary to reselect the training set to establish the classification model, which adds many costs.

(2) Methods based on unsupervised learning: this method usually find the outlier by analyzing the relationship between users, if all the user types are unknown. It's usually based on the assumption that the number of normal samples is much larger than the number of abnormal samples. The methods of NTL detection mainly include density-based methods, distance-based methods and clustering-based methods.

A. Methods based on classification

The classification-based approach derives a classification model by training labeled data samples and then classifies one test sample into one type based on this model during the detection phase. It usually consists of two phases. The first one is the classification of labeled data samples. The second one is the classification of unlabeled data samples according to classification model that the first-phase derived in order to distinguish between normal and abnormal. NTL detection commonly used decision trees, support vector machines, neural networks and other algorithms based on classification algorithms.

The main advantage of decision trees is that once the model is established, understandable rules can be generated. In the learning phase, the decision tree model is established using the training data according to the principle of minimizing the loss function. In the forecasting phase, new data are classified using decision tree model. Early on, Gontijo et al. developed a classification system based on a decision tree to reduce the inspection costs of power companies. It preselected customers who are going to be inspected, resulting in an on-site inspection rate of 30% [26]. Monedero et al. detected fraud using the CART decision tree, combined with Bayesian network to detect customer spending patterns and merged suspicious clients detected by both two methods, for reducing the number of clients to be inspected [14].

Support vector machines have many unique advantages in a small number of samples, in non-linear and high-dimensional pattern recognition. The basic idea of SVM is:

there is a training sample, $\{(x_i, y_i)\}(i = 1,2,3,\Lambda\ ,m)$, which can be accurately classified by the hyperplane, wx + b =0. In $x^i \in R^n, y_i \in \{-1,1\}$, m is the number of samples, R^n is N-dimensional real number space.

The purpose of the algorithm is to get a hyperplane classifying dataset that has the largest distance with different sample points. Jie Liu et al. detected anomaly using C-SVM algorithm. They solved the optimal parameters through genetic algorithm, and acquired the results of real-time detection of user abnormalities. Meanwhile, with the update of historical data, the detection rate in the running process gradually reached the optimal [27]. Depuru proposed and implemented a data encoding technique, which input the encoded data into the support vector model, effectively reducing the complexity of evaluating instantaneous energy consumption data [28].

Extreme Learning Machine (ELM) is an algorithm proposed by Guangbin Huang to solve single hidden layer neural network. The most striking feature of ELM is faster than the traditional learning algorithm in learning speed, meanwhile, it can ensure the accuracy of learning. Online Sequential Extreme Learning Machine (OS-ELM) is an online incremental fast learning algorithm based on ELM for dynamic data applications. Nizar et al. compared and analyzed the SVM, ELM and OS-ELM algorithms in the NTL detection. It is found that detection speed of ELM was faster than SVM in batch mode and OSM-ELM in on-line learning mode, and accuracy of ELM and OS-ELM are higher than SVM [29-31].

In 1943, McCulloch and Pitts proposed the M-P neuron model [32]. In this model, neurons receive input signals from n other neurons. These input signals are passed through weighted connections. The total input value received by the neuron is compared to the threshold of the neuron, then processed by activation function to produce neuron output. Output is represented by y, $y = f(\sum_{i=1}^{n} \omega_i x_i - \theta)$. By linking many of these neurons to a certain hierarchy, we get a neural network. Based on the characteristics of electricity theft, Zheng Cao et al. constructed an anti-electricity theft index evaluation system and established an anti-electricity theft model using a three-layer BP neural network with multiple inputs and single outputs [33]. Depuru et al. proposed a coding technique with faster evaluating whether is illegal consumers, and a hybrid model of SVM and neural networks with faster classifying illegal consumers [34]. Mandava et al. proposed a new idea of using probabilistic neural network (PNN) to detect electricity theft. The model was based on electricity behavior to choose suspicious customers [35].

B. Methods based on density

The density-based approach derives the notion of density by calculating the distance between data and the number of data in a given range, then correlates the degree of anomaly of the data object with the density of the data around it. Breuning put forward the concept of local anomaly in 2000, and proposed an anomaly data mining algorithm based on Local Outlier Factor. This algorithm measure the anomaly of the data object by calculating the neighborhood density of in each data object and the ratio of the average density of data objects in this neighborhood. The greater the possibility of the data object being an anomalous data object if the local anomaly factor of the data object is larger. Chijie Zhuang et al. proposed an anomaly detection model based on unsupervised learning. Firstly it extracted a number of features that characterize the user's power consumption patterns. Secondly mapped each user to a two-dimensional plane by principal component analysis. Thirdly achieved data visualization to calculate the local outlier factor easily. Finally, grid processing is used to filter out the data points in the low density region, which improves significantly the efficiency of the algorithm [36].

C. Methods based on distance

The distance-based method is proposed by Knnor and Ng. They think that the abnormal data object is a data object whose distance from at least k data objects to itself is greater than a certain threshold d in the data set. The essence is that the abnormal data object is considered as an object whose neighbor is relatively sparse within the threshold d.

Chao Cheng et al. considered users who A, C phases current is balanced as the centroid of the object distribution, and determined the threshold according to the Euclidean distance distribution of each object from the centroid. Suspected targets is objects which is larger than the threshold. It provided effectively a new idea for anti-tattling analysis [37].

D. Methods based on cluster

The clustering-based method clusters the data set first, then determines whether the clustering cluster is the abnormal data. The method is highly targeted. Result mainly depends on the number of clustering clusters in the data set and the existence of abnormal data. Although the algorithm is suitable for large-scale data sets, the effect of abnormal detection on high-dimensional data is not very satisfactory. Gerbec et al. used fuzzy C-Means (FCM) clustering algorithm to cluster consumer information to get different consumer groups [38].

First and last, abnormal data mining based on supervised learning has higher detection rate than methods of unsupervised learning. However, there are also some shortcomings. For example, Training set based on supervised learning usually depends on the analysis and construction of experts in this field, which is expensive and the method is suitable when the normal and abnormal data volumes are averaged. It can't effectively solve classification problem of the data set with a small amount of abnormal data.

V. Privacy-Preserving of Power Consumer

It is noteworthy that the power company must know the user's privacy information, load profiles or instrument readings at some point for NTL detection. However, the disclosure of such information will be a serious violation of the user's privacy. In particular, the user's private information may be sold to interested third parties. For example, an insurance company may acquire user's power consumption pattern after obtaining information such as load profiles of users, so as to determine a risk level of a fire, thereby increase an insurance premium. Criminals may infer their daily behavior users through the user's power consumption pattern, then commit the crime [39-41].

For the first time, Salinas et al. considered the issue of user privacy risk in NTL detection, and developed a distributed privacy-preserving energy theft detection algorithm LUD based on LU decomposition and an LUDP algorithm based on local pivot LU decomposition. Among them, LUD is suitable for NTL detection of small networks. However, it may not be stable in large networks. LUDP is suitable for large networks. However, LUDP requires a higher execution time compared with LUD [42]. Both of these algorithms have their limitations. It's that the scene can only happen when the user steals at a constant rate. Therefore, the author proposes a QRD algorithm based on QR decomposition in a variable constant rate. Large-scale network can also work well, the algorithm also has similarly a higher computational complexity and communication complexity [43]. In 2016, the author also proposed the SEK and PPBE algorithm for privacy-preserving energy theft detection in micro grid [44]. SEK is a centralized NTL detection algorithm based on Kalman filter, which can effectively identify users who steal power but can't protect privacy. Then the author decomposes the Kalman filter into two parallel and loosely coupled filters to find the power detector based on the SEK algorithm. The convergence speed of the proposed algorithm is faster than SEK and the algorithm is more stable.

At present, the study of privacy-preserving in the field of NTL detection is still a minority. With the enhancement of user privacy awareness, the traditional NTL detection technology that relies on user historical data can't fully meet the needs of the market. The existing NTL detection methods referring privacy-preserving still has much room for improvement in reducing computational complexity, communication complexity and improving robustness.

VI. Conclusions

In this paper, we discuss and analyze the concept, method, effects and application of machine learning on NTL detection. The authors found that classification-based machine learning algorithms are widely used in NTL detection, while density-based, distance-based and cluster-based approaches are less commonly used. Because most of the research is supported by utility companies that can provide the characteristics of users enough, and the accuracy of modeling based on classification method is higher. Next, the existing privacy protection methods detection are all of high sampling rate in the field of NTL, and the data set itself is processed, which may easily lead to the drop of the detection accuracy. According to the privacy design principles based on data minimization, it's usually the most effective. The higher the sampling rate, the greater the risk of leakage customer privacy. Therefore, the privacy-preserving methods that does not depend on the high data collection frequency are the future research direction. We found that the NTL detection at home and abroad still exist the following problems:

(1) There is no systematic research on the feature processing of load sequences at different time scales.

(2) Most of the researches focus on the accuracy of model prediction, lack of further research on the computational efficiency of large-scale data sets.

(3) Most of the NTL detection model is based on historical data to establish, not based on changes in user data to make the appropriate adjustments, real-time is not high.

However, with the continuous development of machine learning technology, the research on NTL detection will be further deepened. The problem of NTL will be solved to the greatest extent possible, so that power companies can reduce the cost of electricity and improve transmission and distribution efficiency.

References

[1] Nizar A H, Dong Z Y, Jalaluddin M, et al. Load Profiling Method in Detecting non-Technical Loss Activities in a Power Utility[C]// Power and Energy Conference, 2006. PECon '06. IEEE International. IEEE, 2007:82-87.

[2] Depuru S S S R, Wang L, Devabhaktuni V. Electricity theft: Overview, issues, prevention and a smart meter based approach to control theft[J]. Energy Policy, 2011, 39(2):1007-1015.

[3] Davidson I E. Evaluation and effective management of nontechnical losses in electrical power networks[C]// Africon Conference in Africa, 2002. IEEE Africon. IEEE, 2002:473-477 vol.1.

[4] Figueiredo V, Rodrigues F, Vale Z, et al. An electric energy consumer characterization framework based on data mining techniques[J]. IEEE Transactions on Power Systems, 2005, 20(2):596-602.

[5] Lo K L, Zakaria Z. Electricity consumer classification using artificial intelligence[C]// Universities Power Engineering Conference, 2004. Upec 2004. International. IEEE, 2004:443-447 Vol. 1.

[6] Chicco G, Napoli R, Piglione F, et al. A Review of Concepts and Techniques for Emergent Customer Categorisation[J]. 2002.

[7] Smith T B. Electricity theft: a comparative analysis[J]. Energy Policy, 2004, 32(18):2067-2076.

[8] Zengming Liu.Study on Methods and Countermeasures of Anti-theft in Power Supply Enterprises [D]. North China Electric Power University (Baoding) North China Electric Power University, 2013.

[9] Xin Wang, Zhi Zhou Bi, Yanfeng Zhao, et al.Effects of high frequency radiation on the measurement of digital power meter [J] . Electronics Device, 2012, 35 (5): 571-574.

[10] Zhengkan Xia, Shihui Hu, Bingxiang Jiang.Design of a smart remote anti-stealing electrical inspection device [J] .Electric Application, 2015 (S1): 118-121.

[11] Dash, Manoranjan, and Huan Liu. "Feature selection for classification." Intelligent data analysis 1.3 (1997): 131-156.

[12] Liu, Huan, and Lei Yu. "Toward integrating feature selection algorithms for classification and clustering." IEEE Transactions on knowledge and data engineering 17.4 (2005): 491-502.

[13] Filho J R, Gontijo E M, Delaiba A C, et al. Fraud identification in electricity company customers using decision tree[C]// IEEE International Conference on Systems, Man and Cybernetics. IEEE, 2004:3730-3734 vol.4.

[14] Monedero I, Biscarri F, León C, et al. Detection of frauds and other non-technical losses in a power utility using Pearson coefficient, Bayesian networks and decision trees[J]. International Journal of Electrical Power & Energy Systems, 2012, 34(1): 90-98.

[15] Nagi J, Yap K S, Tiong S K, et al. Nontechnical loss detection for metered customers in power utility using support vector machines[J]. IEEE transactions on Power Delivery, 2010, 25(2): 1162-1171.

[16] Nagi J, Yap K S, Tiong S K, et al. Detection of abnormalities and electricity theft using genetic support vector machines[C]//TENCON 2008-2008 IEEE Region 10 Conference. IEEE, 2008: 1-6.

[17] Kennedy J, Eberhart R C, Shi Y. - Swarm Intelligence[J]. Swarm Intelligence, 2006, 2(1):475–495.

[18] Holland J H. Adaptation in natural and artificial systems[M]. MIT Press, 1992.

[19] Zong W G. Music-Inspired Harmony Search Algorithm[M]. Springer Berlin Heidelberg, 2009.

[20] Dorigo M, Birattari M. Ant colony optimization[M]//Encyclopedia of machine learning. Springer, Boston, MA, 2011: 36-39.

[21] Kennedy J, Eberhart R C. A discrete binary version of the particle swarm algorithm[C]// IEEE International Conference on Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation. IEEE, 2002:4104-4108 vol.5.

[22] R.J. Ellison, D.A. Fisher, and R.C. Linger, “Survivability: Protectiong Your Critical Systems,” IEEE Internet Computing, vol. 3, issue 6, pp. 55-63, 1999.

[23] Ramos C C O, Souza A N, Chiachia G, et al. A novel algorithm for feature selection using Harmony Search and its application for non-technical losses detection ☆[J]. Computers & Electrical Engineering, 2011, 37(6):886-894.

[24] Ramos C C O, Souza A N D, Falcao A X, et al. New Insights on Nontechnical Losses Characterization Through Evolutionary-Based Feature Selection[J]. IEEE Transactions on Power Delivery, 2012, 27(1):140-146.

[25] Pereira D R, Pazoti M A, Pereira L, et al. Social-Spider Optimization-based Support Vector Machines applied for energy theft detection[J]. Computers & Electrical Engineering, 2016, 49(C):25-38.

[26] Gontijo E M, Delaiba A C, Mazina E, et al. Fraud identification in electricity company customers using decision tree[C]//Systems, Man and Cybernetics, 2004 IEEE International Conference on. IEEE, 2004, 4: 3730-3734.

[27] Jie LIU, Yue-bin HOU, Nian LIU, et al.An intelligent technology for detecting abnormal power loss based on non-technical loss [J] .East China Electric Power, 2014, 42 (4): 650-656.

[28] Depuru S S S R, Wang L, Devabhaktuni V. Enhanced encoding technique for identifying abnormal energy usage pattern[C]// North American Power Symposium. IEEE, 2012:1-6.

[29] 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.

[30] Nizar A H, Dong Z Y, Zhao J H, et al. A Data Mining Based NTL Analysis Method[C]// Power Engineering Society General Meeting. IEEE, 2007:1-8.

[31] Nizar A H, Dong Z Y. Identification and detection of electricity customer behaviour irregularities[C]// Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES. IEEE, 2009:1-10.

[32] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity[J]. The bulletin of mathematical biophysics, 1943, 5(4): 115-133.

[33] Zheng Cao, Jingfei Yang, Xiaona Liu.Research and Application of BP Neural Network in Anti-stealing System [J]. Hydropower and Energy Science, 2011, 29 (9): 199-202.

[34] Depuru S S S R, Wang L, Devabhaktuni V, et al. A hybrid neural network model and encoding technique for enhanced classification of energy consumption data[C]//Power and Energy Society General Meeting, 2011 IEEE. IEEE, 2011: 1-8.

[35] Mandava S, Vanishree J, Ramesh V. Automation of Power Theft Detection Using PNN Classifier[J]. 2014.

[36] Chijie Zhuang, Bin, HU Jun Zhang, et al.Analysis of power user abnormal power mode based on unsupervised learning [J] .Proceedings of the CSEE, 2016,36 (2): 379-387.

[37] Chao Cheng, Hanjing Zhang, Zhimin Jing, et al.Study on anti-stealing based on outlier algorithm and electricity information collection system [J] .Power System Protection and Control, 2015 (17): 69-74.

[38] Gerbec D, Gasperic S, Smon I, et al. Allocation of the load profiles to consumers using probabilistic neural networks[J]. IEEE Transactions on Power Systems, 2005, 20(2): 548-555.

[39] Ruzzelli A G, Nicolas C, Schoofs A, et al. Real-Time Recognition and Profiling of Appliances through a Single Electricity Sensor[C]// Sensor Mesh and Ad Hoc Communications and Networks. IEEE, 2010:1-9.

[40] D. Dan, Y.Q. Zhang, “Research on Definition of Network Survivability,” Journal of Computer Research and Development, vol. 43(Suppl.), pp. 525-529, 2006.

[41] K.Q. Zhao, Set pair analysis and its preliminary applicatio. Hangzhou, Zhejiang Science and Technology Press, 2000.

[42] Salinas S, Li M, Li P. Privacy-preserving energy theft detection in smart grids[C]// Sensor, Mesh and Ad Hoc Communications and Networks. IEEE, 2012:257-267.

[43] Salinas S, Li M, Li P. Privacy-preserving energy theft detection in smart grids[C]// Sensor, Mesh and Ad Hoc Communications and Networks. IEEE, 2012:257-267.

[44] Salinas S A, Li P. Privacy-Preserving Energy Theft Detection in Microgrids: A State Estimation Approach[J]. IEEE Transactions on Power Systems, 2016, 31(2):883-894.