



A data-based approach for online smart meter calibration

Fangxing Liu^{1,2}, Chengbin Liang^{2,3}, Qing He^{1,2}

¹ Harbin Institute of Technology, Harbin, P. R. China

² National Institute of Metrology, Beijing, P. R. China

³ College of Electrical and Information Engineering, Hunan University, Changsha, P. R. China

ABSTRACT

The smart meter has been considered the key element in the smart grid. The current smart meter verification method mainly focuses on on-site and laboratory detection and fails to test all deployed meters. Online calibration is a novel approach that calculates meter error by analyzing meter reading data. By presenting a comprehensive survey and basic model of online smart meter calibration, this study proposes a recursive algorithm in order to estimate meter error. Concrete steps are subsequently illustrated. The following testing case shows the higher accuracy of the method. Finally, some issues related to the application of online calibration are discussed.

Section: RESEARCH PAPER

Keywords: Smart meter; online calibration; meter error

Citation: Fangxing Liu, Chengbin Liang, Qing He, A Data-Based Approach for Smart Meter Online Calibration, Acta IMEKO, vol. 9, no. 2, article 6, June 2020, identifier: IMEKO-ACTA-09 (2020)-02-06

Editor: Alexandru Salceanu, Technical University of Iasi, Romania

Received January 17, 2020; **In final form** March 24, 2020; **Published** June 2020

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work is supported by the National Key R&D Program of China under Grant No.2016YFF0201201.

Corresponding author: Qing He, e-mail: heqing@nim.ac.cn

1. INTRODUCTION

The smart meter is an important component in a smart grid. The performance and measurement accuracy of smart meters is directly related to power grid safety and energy trade settlement [1]. Consequently, it is particularly important to determine the operation error state of smart meters [2]-[3]. A smart meter is not the same as a traditional electricity meter. In addition to the basic function of metering electricity, a smart meter has two-way communication and multiple measurement modules, such as the data transmission and anti-stealing module. It represents the direction of the development of the intelligent terminal of future energy-saving smart grids. The widespread popularity of smart energy meters enables the collection of a large amount of electricity data, such as energy consumption, voltage, and current data. In China, 460 million smart electricity meters have been installed and in US and UK, the deployed meter number is 70 million and 2.9 million [4].

With the increasing number of deployed smart meters, the verification of massive meters tends to be an important task for energy companies. As a significant indicator of a meter's working status, a meter with the error that it exceeds the official specific threshold will be classified as an abnormal meter. At present, on-site and laboratory checking are two main ways in which power companies verify meter accuracy [5]. If a meter needs to be

detected, the electrician should remove the meter from the site and install a new one, and they should then send the removed meter to professional agencies to test its metrological character. It is a time-consuming and costly work to detect millions of smart meters. Many smart meters are designed to be used for more than ten years. In this way, an energy company can determine in which year the energy meters should be replaced. For example, China Smart Grid removes meters in the eighth year after the meter's installation in order to reduce the testing costs. Furthermore, some energy providers conduct random sampling to pick up parts of meters to test their performance. The staff carry instruments to conduct on-site field tests at the meter installation site. Then the status of every meter can be inferred from the on-site test results. However, the disadvantage of these two existing methods is that they cannot verify the meter error of all meters. As a result, it is difficult to decide whether the smart meter should continue to be used or if it should be removed.

Advanced Metering Infrastructure (AMI) is the basic component of a smart grid, which enables large-scale measurement data collection and transmissions in a power grid. With the rapid development of AMI, an increasing amount of data can be acquired from a smart meter [6]-[8]. Smart meter data analytics has become a popular research topic, saturated with the increasing volume of data collected by meters. Many novel

methods and algorithms have been developed to analyse meter data in order to discover the internal information. Energy data analysis includes consumer profiling, load forecasting, electricity pricing, identification of irregularities, metering, and real-time operations [9]-[12]. Consumption data analytics is helpful for constructing a consumer behaviour model and forecasting consumer consumption levels. Due to the emergence of AMI and the broad application of enormous intelligent meters, it is possible to carry out more accurate load forecasting by using a large amount of household load data [13]-[15]. The energy pricing strategy can be updated according to consumption data analytics because the energy data shows the peak and troughs of daily electricity usage. Abnormal data can be detected by mining the inner relationship among energy usage data. Furthermore, data analytics is a powerful and effective method when implementing the monitoring the equipment in the power grid [16].

Online calibration for smart meters is a novel method that computes meter errors by processing meter reading data. Due to the increased digitisation of smart meters, multiple types of energy data can be stored and transmitted to energy data centres. Those multiple types of data contain information about energy consumption, voltage, current, time, location, and installation condition. To analyse the data and estimate the meter error, online calibration is a new idea that can solve the problem of ineffective meter detection. By analysing the energy usage data, the errors of all meters in a low-voltage system can be calculated at the same time. Then, a meter error that exceeds the threshold will be classified as an abnormal meter.

There are few studies on online smart meter calibration. Korhonen [17] introduced a method by which energy consumption data were used to evaluate meter error and energy loss. However, this method does not work well in a large energy loss rate system. In [18], Kong proposed a double-parameter based least squares algorithm to estimate meter errors. In [19], an algorithm combined with cluster theory and regularization was proposed to calculate meter error. This method performs better when the system energy loss rate is low. In [20], Peng adopted the Monto Carlo method to estimate combined meter error.

In the relevant work on this topic, energy data analytics has been widely used in load forecasting and abnormal equipment detection. Online calibration is a novel idea that estimates meter error by analysing meter measurement data. The traditional meter calibration process includes the verification of the initial installation of electricity meter; regular on-site and laboratory meter error testing; periodic replacement; and so on. These steps are extremely complex and costly when detecting a huge number of smart meters, while online calibration can mitigate this shortcoming well.

In this article, a method is proposed for the online estimation of meter error. This method provides a theoretical basis for the online calibration of large-scale smart meters. The meter's deployment structure and low-voltage system model are analysed. Our meter error estimation process has three steps. First, we filter the abnormal data and cluster the data with similar features. Then, an approach based on recursive least squares is adopted to estimate meter error. Finally, we compare the estimated meter error with a threshold to determine an abnormal meter.

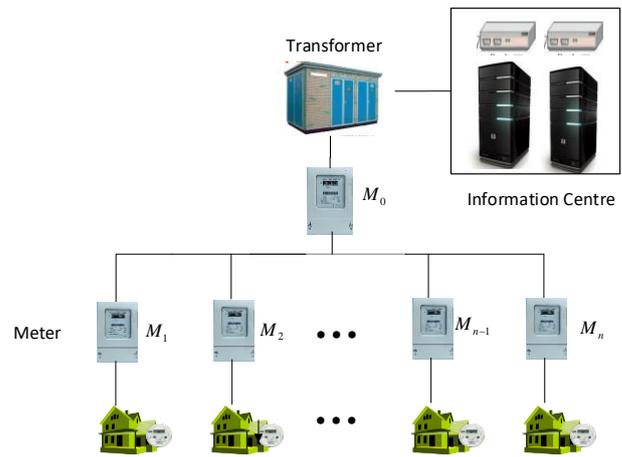


Figure 1. Low-voltage network.

2. ONLINE CALIBRATION

2.1. Low-voltage network

Smart meters are deployed in a low voltage network. As shown in Figure 1, there is a transformer connecting the information centre and the smart meters. The energy meter installed under the low-voltage side records the consumption of the energy user. All data collected by the meter are transmitted to the information centre through the communication system. Wireless, power line carrier, and RS485 are typical power information transmission systems in a power grid. The information centre collects meter data at a specific frequency, such as every fifteen minutes or one hour. The collected data contains information about the current, voltage, power, time, meter number, address, load rate, and average power factor of the energy grid.

2.2. Online calibration model

Smart meters are topologically deployed in the low-voltage system, which means a master meter is installed in front of some submeters. The master meter has higher accuracy than the submeter and is considered as the standard meter in the system. We assume the master meter has been calibrated and the reading is accurate. Electric energy passes through the master meter and then to all the power users, whose consumption is recorded by the submeters. The consumption value that the master meter records equals the sum usage value of the submeters in an ideal way. Then, we have the equation:

$$\sum_{j=1}^n E_j(i) = E_0(i) \quad (1)$$

where $E_j(i)$ is the true energy consumption on branch j recorded by submeter j , and $E_0(i)$ is the true energy consumption on all branches recorded by the master meter at the i th measurement.

However, the meter has an error, and the meter reading is not the true energy consumption. Furthermore, there are energy losses in the system due to the meter's own consumption, connection, cabling, etc. To take this into account, Equation (1) can be rewritten as:

$$\sum_{j=1}^n \frac{\Psi_j(i)}{1 + \xi_j} + E(i) = \frac{\Psi_0(i)}{1 + \xi_0} \quad (2)$$

where ξ_j is the relative error of meter j , and $\Psi_j(i)$ is the reading of submeter j . $\Psi_0(i)$ is the reading of the master meter. $E(i)$ is the

energy losses in the system. After n measurements, we can write a matrix equation:

$$\Psi^T \Theta = Y \quad (3)$$

Where:

$$\Psi = \begin{bmatrix} \psi_1(1) & \psi_2(1) & \cdots & \psi_i(1) \\ \psi_1(2) & \psi_2(2) & \cdots & \psi_i(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(n) & \psi_2(n) & \cdots & \psi_n(n) \end{bmatrix}$$

$$\Theta = \begin{bmatrix} \frac{1}{1+\varepsilon_1} \\ \frac{1}{1+\varepsilon_2} \\ \vdots \\ \frac{1}{1+\varepsilon_n} \end{bmatrix} = \begin{bmatrix} \theta(1) \\ \theta(2) \\ \vdots \\ \theta(n) \end{bmatrix}$$

$$Y = \begin{bmatrix} \frac{\psi_0(1) - E(1)}{1+\varepsilon_0} \\ \frac{\psi_0(2) - E(2)}{1+\varepsilon_0} \\ \vdots \\ \frac{\psi_0(n) - E(n)}{1+\varepsilon_0} \end{bmatrix} = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix} \quad (4)$$

The measurement number should not be less than the submeter number n , such that the relative error can be computed by solving the linear equation.

The energy losses mainly consist of self-consumption and power line losses. The mathematic formulation of E_i is as follows [21]:

$$E_i(t) = \sum_{j=1}^n \bar{p}_j(t)t + \frac{E(t)\Delta U(t)}{100} \times K(t) \quad (5)$$

where \bar{p}_j is the rated power of the smart meter. $E(t)$ is the master meter reading, and ΔU is the loss rate of the voltage. K_j is a factor related to the voltage between the submeters and master meter.

2.3. Recursive meter error estimation

The advanced metering infrastructure enables the information centre to collect and store a large amount of data. The effective analysis of those meter data could enhance estimation accuracy. Solving the $n \times n$ linear equation can result in an error estimation, but there is a limitation. It is an off-line method and only uses the n times measurement without considering other measurement data. The data quality of the meter reading strongly affects the solution's accuracy because one abnormal data could change the matrix property and then lead to an unfavourable result – even *no* result. The recursive algorithm is an effective method of estimating the meter error by updating new data.

The least squares method seeks to minimize the squares of the difference between the measured data and the fitted data, and the cost function is defined as follows:

$$J = \sum_{i=1}^n (y(i) - \psi(i)\theta(i))^2 \quad (6)$$

J is minimised by calculating partial derivatives to each θ :

$$\frac{\partial J}{\partial \theta_j} = 2 \sum_{i=1}^n (y(i) - \theta(i)\psi(i))(-\psi_j(i)) = 0 \quad (7)$$

Then, θ in the classic least squares model is solved, and the error parameter is estimated as:

$$\hat{\Theta} = (\Psi^T \Psi)^{-1} \Psi^T Y \quad (8)$$

To impose weights on the measurements with suitable weights ω and the weighted least squares estimation for θ is:

$$\hat{\Theta} = (\Psi^T W(n) \Psi(n))^{-1} \Psi^T W(n) Y(n) \quad (9)$$

where

$$W = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_n \end{bmatrix} \quad (10)$$

In the recursive weight least squares model, the new model $\theta(n+1)$ is calculated as the sum of the previous estimates. The recursive equation is written as follows:

$$\hat{\Theta}(n+1) = \hat{\Theta}(n) + \zeta(n)(y(n+1) - \Psi(n+1)^T \hat{\Theta}(n)) \quad (11)$$

where $\hat{\Theta}(n+1)$ is the new estimation parameter, and $\zeta(n)$ is the update factor. The $n+1$ measurement is

$$\hat{\Theta}(n+1) = (\Psi(n+1)^T W(n+1) \hat{\Theta}(n+1))^{-1} \Psi(n+1)^T W(n+1) Y(n+1) \quad (12)$$

Using the exponential forgetting factor in the weight matrix and the estimation equation results in:

$$\hat{\Theta}(n+1) = \hat{\Theta}(n) + P(n+1)\psi(n+1)(y(n+1) - \Psi(n+1)^T \hat{\Theta}(n)) \quad (13)$$

Where

$$P(n+1) = (\Psi(n+1)^T W(n+1) \hat{\Theta}(n+1))^{-1} \quad (14)$$

Calculating P and the updating factor, the recursive weight least squares are:

$$\zeta(n) = P(n)\psi(n+1)(\Psi^T(n+1)P(n)\Psi(n+1) + \lambda)^{-1} \quad (15)$$

$$P(n+1) = (I - \zeta(n)\psi^T(n+1))P(n)\frac{1}{\lambda} \quad (16)$$

2.4. The framework of error estimation

In order to estimate meter error, abnormal data should be filtered at first. Then, data with similar features should be processed and clustered. Next, the linear equation should be constructed using the clustered data. Finally, the equation should be solved in order to obtain the meter error estimator. Figure 2 shows the framework of online error estimation:

- 1) Loading and screening data: Given the smart meter data and check each measurement $E_j(j)$. The measurement $\sum_{j=1}^n E_j(i) > E_0(i)$, which means the sum of the submeters exceeds the sum of the master meters, will be classified as abnormal data. Those data will be removed from the data set.
- 2) The meter self-consumption and online energy losses are calculated by Equation (5). Energy losses are subtracted from the master meter reading in order to reduce the energy influence on the meter error estimation.
- 3) Data pre-processing: Data is clustered using the K-means algorithm in order to collect data with similar features [21]-[23]. A submeter matrix can affect the solution of the linear equation. The data sets with high correlation will lead to an ill-condition matrix that is difficult to solve. Clustering data with similar characters into a group and pick up the data from different groups to build the matrix. This matrix is a low ill-condition matrix.

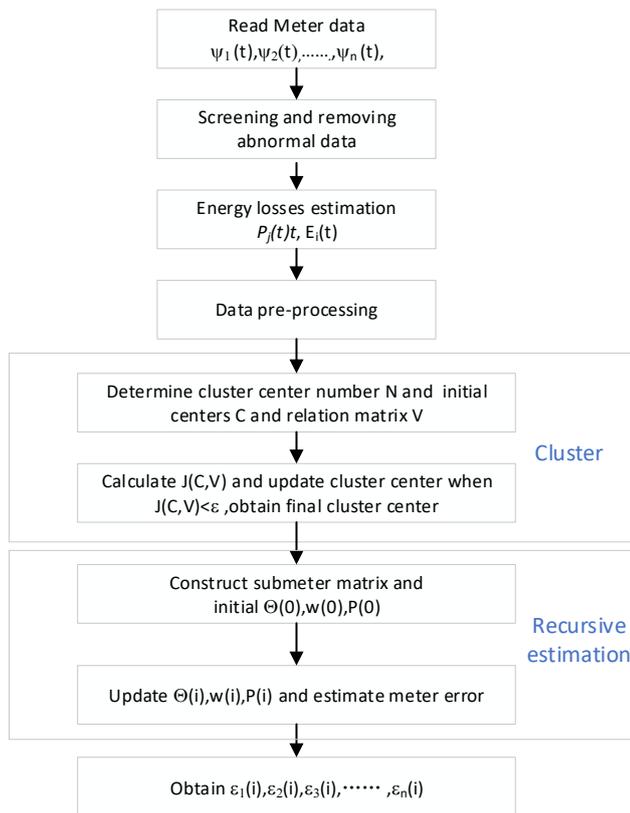


Figure 2. Framework of online error estimation.

- 4) Meter error estimation: The linear equation with the recursive algorithm in Section 2.3 is solved. Initialising $\Theta(0) = 1$, $\sum w(0) = 0$, $P(0) = \kappa I$. κ is a small constant. For an n submeters system, $P(0)$ is an $n \times n$ matrix and $\Theta(0)$ is a $n \times 1$ vector. The meter error can be obtained through the solution.

Root Mean Square Error (RMSE) is a method of verifying the effectiveness of a calculation algorithm. The formula of RMSE is:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{\varepsilon}_i - \varepsilon_i)^2} \quad (17)$$

A smaller RMSE value indicates a better estimation.

3. ERROR ESTIMATION

In order to verify the validity of the algorithm, we selected a dataset from a typical low voltage system in an urban area as the testing case. There are 123 electricity meters with one master meter and 122 submeters deployed at this site. Meter measurement data were collected for three years, from 2014 to 2016. The master meter measured all the consumption at the transformer secondary, and all consumption data were recorded and sent to the energy centre at 0:00 every day. The dataset has no information about meter error, which we simulate. We randomly assign large meter errors to the submeters. According to the official smart meter regulation, a meter error above 2% is classified as an abnormal meter. We start by screening and filtering the abnormal data, then we cluster the data in order to construct the meter data equation.

The meter errors estimated by the recursive algorithm are shown in Figure 3. Meter errors are calculated once a day – at the

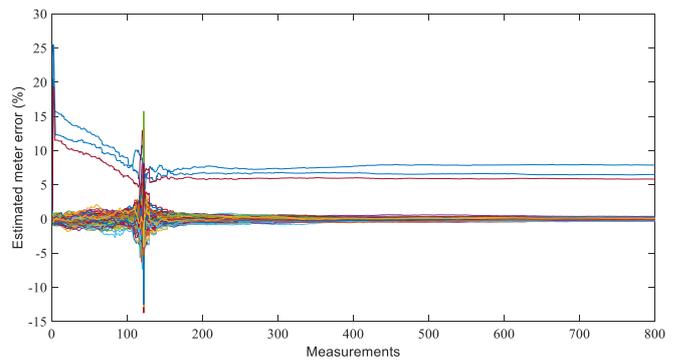


Figure 3. Estimate error from different measurements

same frequency as the data acquisition. The x -axis represents the number of measurements and the y -axis shows the meter error in percentage form.

The error estimator varies greatly in the initial 200 measurements, then it seems to be a stable value after 500 measurements. It can be shown that three meters estimated errors exceeding 2%.

All the submeters' estimated results are gathered to verify the effectiveness of the method, as shown in Figure 4. It can be seen that most meter errors fall into an acceptable range, with an absolute value of error below 2%. Meters 36, 78, and 91 were abnormal meters, with errors 7.969%, 6.537%, and 5.871%, respectively. The RMSE of this estimation is 0.22%, which means the deviation between the estimator and the true meter error is very small. This is an effective approach to the calculation of meter error.

4. DISCUSSION

The basic idea of online calibration has been introduced in previous sections. It is an effective method in the energy system with a master deployed in front of submeters. However, the real low-voltage system is a complex energy network, and the meter's data quality is not as good as we expected. There are several issues in online smart meter calibration application.

- 1) Data collection

In order to construct and solve the equation (3) successfully, energy consumption data from all meters should be collected. The more data collected, the more accurate the meter error estimation result will be. In some energy systems, the data are transmitted once a day, while meter data are collected every 15 minutes in other systems. In a 200-submeter system, the matrix requires at least 200 measurements. It takes a long time for a system to store the data if the data are collected only once a day.

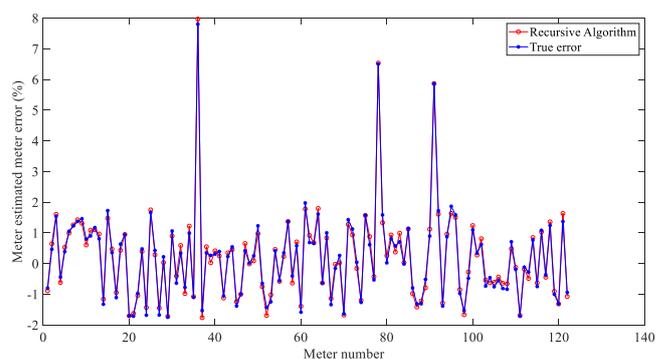


Figure 4. Meter error estimates.

Data consistency is also an important consideration. In the framework of online calibration, the data from the master meter and submeters should be collected at the same time. In real data collection, some of the submeters fail to receive the data transmission signal on time and transmit the delayed data to the control centre. Those nonsynchronous data adversely affect the validity of the online calibration model.

Data collection is an important foundation in remote calibration. This is an issue that analysts can hardly solve in a mathematical way. It is necessary to confirm the synchronization of measurement data collection.

2) Energy system

In a normal low-voltage system, a master meter is deployed in front of the submeters. The master meter is three-phase energy meter, while the submeters are single-phase energy meters. In the low-voltage system model, all submeters should be installed in the same phase lines so that the master meter energy consumption reading is equal to the sum of submeters. However, the real energy system may mix with the submeters and place them in different phases. Two more submeters connected with each other in different phase branches will affect the energy conservation relationship. This will further significantly influence the equation solution and lead to a biased error estimator.

3) Abnormal data

Abnormal data always exists in a power system. Power theft and environmental changes will generate abnormal data. The stolen electricity in the system is only recorded by the master meter. So, those abnormal data will break the equality relationship and damage the meter error solution. Electricity theft is a worldwide problem, and researchers have attempted to address the problem by analysing smart meter data. Supervised learning methods, such as support vector machines (SVM), are effective approaches to detect the abnormal data. They first train an abnormal data detector and then input the real data to classify and determine the bad data. A pre-processing of smart meter data will be a great help in the completion of the online error computing.

4) Energy losses

Energy losses always exist in the energy network. Most energy losses happen on the power line and the equipment self-consumption. In the low voltage system, the smart meter self-consumption is fixed. The accuracy of meter error estimation declines with increasing energy losses online. The energy losses are difficult to estimate due to the lack of system parameters such as the length of power line from each master meter to submeter. The deviation of energy loss estimation will bring some issues to the remoting meter calibration.

5) Analytical methods

The solution to the equation and estimation of the meter error is one of the most effective methods of online calibration. However, the imperfect equal relationship in the equation will influence the equation solution. Many algorithms have been proposed to improve the accuracy of meter error estimation on estimating meter error. New technologies involving computer science, particularly machine learning, will be powerful tools in meter data analysis. Classification, regression, and deep learning methods are popular in smart grid data analytics. Transfer learning and incremental learning have also been used in smart meter analytics. However, when introducing one algorithm to solve a problem, the meaning and physical implications thereof should be taken into consideration carefully.

5. CONCLUSIONS

This article explained the online calibration principle and proposed a recursive algorithm approach to estimate meter error. The case analysis showed the excellent performance of this method. Online calibration is a promising metrological method, as are the development of big data acquisition systems and advanced data analytics techniques. Online meter error estimation and calibration is an emerging research area, which has a broad application and great economic value. We hope this article can provide readers with a picture of and insights into the usage of meter data and online meter calibration.

ACKNOWLEDGEMENT

This work is supported by the National Key R&D Program of China under Grant No.2016YFF0201201.

REFERENCES

- [1] D. Alahakoon, X. Yu, Smart electricity meter data intelligence for future energy systems: a survey, *IEEE Transactions on Industrial Informatics* 12 (2016) pp. 425-436.
- [2] J. Hu, A. V. Vasilakos, Energy big data analytics and security: challenges and opportunities, *IEEE Trans. Smart Grid* 7 (2016), pp. 2423–2436
- [3] Y. Kabalci, A survey on smart metering and smart grid communication, *Renewable and Sustainable Energy Reviews* 57 (2016) pp. 302-318.
- [4] Y. Wang, Q. Chen, T. Hong, C. Kang, Review of smart meter data analytics: Applications, methodologies, and challenges, *IEEE Transactions on Smart Grid* 10(2018), 3125-3148.
- [5] M. Huang, Y. Shi, N. Lin, Z. M. Lin, Y. Liu, Y. S. He, Common faults and treatment methods in the verification process of electric meter verification device, *Electr. Eng.* 2 (2018) pp. 89-91.
- [6] F. Benzi, N. Anglani, E. Bassi, L. Frosini, Electricity smart meters interfacing the households, *IEEE Transactions on Industrial Electronics* 58 (2011) pp. 4487-4494.
- [7] A. Bidram, A. Davoudi, Hierarchical structure of microgrids control system, *IEEE Transactions on Smart Grid* 3 (2012) pp. 1963-1976.
- [8] F. L. Quilumba, W. Lee, H. Huang, D. Y. Wang, R. L. Szabados, Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities, *IEEE Transactions on Smart Grid* 6 (2015) pp. 911-918.
- [9] Y. Wang, Q. Chen, D. Gan, J. Yang, D. S. Kirschen, C. Kang, Deep Learning-Based Socio-Demographic Information Identification From Smart Meter Data. *IEEE Transactions on Smart Grid* 10(2019) pp. 2593-2602.
- [10] A. K. Marnerides, D. P. Pezaros, H. Kim, D. Hutchison, Internet traffic classification using energy time-frequency distributions, *Proc. of the IEEE International Conference on Communications IEEE, Budapest, Hungary*, 9-13 June 2013.
- [11] X. Tong, R. Li, F. R. Li, C. Q. Kang, Cross-domain feature selection and coding for household energy behavior, *Energy* 107 (2016) pp. 9-16.
- [12] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, S. Mishra, Decision tree and SVM-based data analytics for theft detection in smart grid, *IEEE Transactions on Industrial Informatics* (2016) pp. 1005-1016.
- [13] J. Bedi, D. Toshniwal, Empirical mode decomposition based deep learning for electricity demand forecasting, *IEEE Access* 6 (2018) pp. 49144-49156.
- [14] W. Kong, Z. Y. Dong, D. J. Hill, F. Luo, Y. Xu, Short-term residential load forecasting based on resident behaviour learning, *IEEE Transactions on Power Systems* 33 (2018) pp. 1087-1088.
- [15] H. Shi, M. Xu, R. Li, Deep learning for household load forecasting: a novel pooling deep RNN, *IEEE Transactions on Smart Grid* 9 (2018) pp. 5271-5280.

- [16] H. Janetzko, F. Stoffel, S. Mittelstadt, D. A. Keim, Anomaly detection for visual analytics of power consumption data, *Computers & Graphics* 38 (2014) pp. 27-37.
- [17] A. Korhonen, Verification of energy meters using automatic meter reading data, Master's Thesis, AALTO University, Espoo, Finland, 2012.
- [18] X. Kong, Y. Y. Ma, X. Zhao, Y. Li, Y. X. Teng, A Recursive Least Squares Method with Double-Parameter for Online Estimation of Electric Meter Errors. *Energies*, 12(805) (2019) pp.1-12.
- [19] F. Liu, Q. He, S. Hu, L. Wang, Z. Jia, Estimation of smart meters errors using meter reading data, Proc. of the Conference on Precision Electromagnetic Measurements (CPEM), Paris, France, 8-13 July 2018.
- [20] X. J. Peng, L. Wang, Z. S. Jia, X. W. Wang, H. T. Huang, L. J. Liu, Evaluation of digital energy meter error by Monte Carlo method. Conference on Precision Electromagnetic Measurements (CPEM 2016), Ottawa, Canada, 10-15 July 2016.
- [21] H. Tang, Research on line loss analysis and loss reduction measures in low voltage substation. Master's Thesis, Southeast University, Nanjing, China, 2017.
- [22] P. Arora, S. Varshney, Analysis of k-means and k-medoids algorithm for big data, *Procedia Computer Science*, 78 (2016) pp. 507-512.
- [23] A. K. Jain, Data clustering: 50 years beyond K-means, *Pattern Recognition Letters* 31(2010) pp. 651-666.