Study on the Identification of the Differential Lasso in the Low-voltage Transformer Area

Yi Wang¹, *, Zhitong Liu², Yuanyi Li², Bo Pang³, Nan Zhao⁴, Jia Liu⁵, Diancai Shi⁵, Gang Hou⁵, Qi Wang⁵, Chuanzong Zhao⁵, Qinghao Wang⁵

¹Measuring Center, Liaoning Electric Power Company Limited, State Grid, China
²Shenyang Power Supply Company, Liaoning Electric Power Company Limited, State Grid, China
³Anshan Power Supply Company, Liaoning Electric Power Company Limited, State Grid, China
⁴Yingkou Power Supply Company, Liaoning Electric Power Company Limited, State Grid, China
⁵Fushun Power Supply Company, Liaoning Electric Power Company Limited, State Grid, China

*Corresponding author e-mail: 627102275@qq.com

Abstract. The validity of the transformer area’s file in the Power information acquisition system is very important, the man-made way is time-consuming, and the cost of relying on the equipment is high. In order to provide efficient, convenient and economical identification technology of the transformer area archives, this paper proposes a low voltage topology recognition method based on differential Lasso, which can complete automatic topology recognition only by using the power consumption information of the transformer area. First of all, the statistical model of Lasso is established for the data of electricity consumption. According to the characteristics of the meter in the meter box belonging to the same area, the differential Lasso model is further constructed. The regression coefficient based on the differential lasso can be calculated for each transformer area in turn by polling, and select the meter according to the threshold. The method of this paper is used to analyze the field collected electricity data. The results of topology recognition show the effectiveness of the proposed method.

1. Introduction
The electricity information system is composed of main station, concentrator, collector and watt-hour meter, the power company sends the transformer area (transformer) file to the concentrator through the main transformer, as the basis for collecting electricity consumption information of collection meter. Therefore, whether the file corresponds to the topological relation of actual transformer and electric meter. [1]In recent years, in order to respond to the national efforts to promote the call of the smart grid, which will directly affect the marketing management level. Various localities gradually carry out reconstruction Project of centralized Meter Reading in low Voltage transformer area. However, due to the lack of strict control over the construction quality, the fault of the line connection often occurs,
which leads to an error in the file of the area. In addition, some old lines are complicated, the file
maintenance in the transformer area is not perfect, and the information of changing meters is not
updated, which also makes the topological relationship difficult to identify. The above problems make
it impossible to accurately obtain the one-to-one correspondence between watt-hour meter and
transformer, which will lead to meter reading error, reduce the management level of line loss in
transformer area, and seriously damage the image of power supply company.

The early topology identification of transformer area was mainly completed by manual, this
requires a large number of power staff to check on the spot, and comb the corresponding relationship
between transformer and meter, the efficiency is very low. This method of relying on manual
troubleshooting has not been able to meet the actual needs in the face of a large-scale area. Therefore,
there is an urgent need for an efficient and convenient method to identify the topological relationship
of the transformer area. In order to save labor cost, there are many kinds of equipment in the market,
which greatly improve the speed and efficiency of topological relationship combing. The existing
transformer identification instrument generally uses power line carrier (PLC) technology or current
pulse technology to judge the corresponding relationship between transformer and electric meter by
sending and receiving signals. However, in the actual application, this way of signal judgment can
be misjudged due to the inaccurate signal. For PLC technology, when the transformer side can not
completely isolate the signal leakage, the carrier signal will be coupled to other areas, resulting in low
reliability of the signal. For the pulse current technology, it is necessary to send the pulse current
signal from the slave machine on the user side, and the host computer located on the low voltage side
of the transformer receives the signal, and the user ownership is judged by harmonic analysis. But the
mode of installing the current transformer on the outlet side has certain risk. In addition, although the
equipment type of checking method can save labor costs, but need to arrange a large number of on-site
equipment or operators to check door-to-door, the workload is also very huge. Therefore, it is
necessary to study a safe, reliable and economical low-voltage topology identification method.

In recent years, with the comprehensive promotion of intelligent electricity meter and power
consumption information acquisition system, a large amount of electricity data can be collected and
stored, such as user voltage, current, electricity and so on. Through the analysis of the data, the
topological relation in the transformer area becomes the current research focus, and has the advantages
of low cost and safety and reliability. In reference [3], the similarity between user side voltage and
transformer voltage is compared by grey correlation degree analysis method, so as to judge the
belonging of transformer area. In reference [4], In this paper, a topological recognition method based
on space-time correlation is proposed, and the corresponding relation between the user and the
transformer is given according to the voltage similarity between the time and the space. However, the
similarity of voltage data is easy to produce approximate results statistically, which affects the
judgment. Considering that the supply and sale of electricity has an equality relationship in theory, the
power reading of check meter on the transformer side check meter will be equal to the sum of all user
meters. Therefore, the problem of topology identification is transformed into: Select a partial meters
from a certain scale of user meters so that the sum of its electricity consumption is equal to the
electricity consumption of the transformer check meters. This can be abstracted as a variable selection
problem in the mathematical sense. In the field of statistics, how to select variables to obtain a realistic
model is a very important subject. To this end, a number of methods have been proposed, which are
representative of the Lasso approach proposed by Tibhirani [5].In this method, the coefficients of
some variables in the linear regression problem are directly compressed to zero, and the variable
selection is facilitated. Efron [8] proposed to use the LASRS method to solve the Lasso problem, the
method has gained more attention, and it is applied in many fields such as finance, biological gene
sequence, signal processing and so on. Based on the massive historical daily frozen electricity data in
residential power acquisition system, a low voltage topology recognition method based on differential
Lasso is proposed in this paper. Once the total meter data of a certain sample and the electricity data of
all the users under a certain sample are given, a linear regression model based on Lasso should be
established first. Further, considering that users in meter boxes must belong to the same transformer
area, therefore, a constraint term of coefficient difference is applied to the Lasso model. So that the nearest neighbor user meter has approximate coefficients, and the difference Lasso model is constructed. Through the statistical analysis of the power consumption data, the ownership relationship between the user and the transformer in transformer area can be obtained, which has important theoretical research significance and practical application value.

2. Topology recognition based on differential Lasso

Principle of user recognition in transformer area

In general, within the normal range of line loss in the transformer area, the power supply of the meter will be approximately equal to the sum of the electricity consumption of the users under the transformer area. Described as a linear regression problem, that is: Given the data \( X \in \mathbb{R}^{n \times m} \), where \( n \) represents the number of sample points, \( m \) represents the number of meters, the following relationship is satisfied for the \( i \) transformer area.

\[
\mathbf{y}_c = \sum_{j=1}^{m_c} \beta_j^{(c)} \mathbf{x}_j^{(c)} + \mathbf{e}_c
\]

(1)

\( m_c \) represents the number of tables in the \( c \) area, \( \mathbf{e}_c \) represents some error information. For a typical topological meter assignment problem, the coefficient \( \hat{\beta}_j^{(c)} \) in the above formula will be strictly equal to (1), but considering the actual influence of noise, the value should float up and down 1. Therefore, for transformer area \( \mathcal{Y}_c \), the following optimization objectives are constructed:

\[
\hat{\beta}_j^{(c)} = \arg \min_{\beta} \| \mathbf{y}_c - X\beta \|_2
\]

(2)

It can be seen that the upper formula belongs to the least square method (OLS) problem. By solving the above formula, the \( \hat{\mathbf{x}} \) with large coefficients are selected as the meter of the transformer area.

Topological solution based on differential Lasso

In the previous section, the general idea of solving topological problems is described, but the least square is easy to be affected by noise, which makes the coefficient fluctuate violently and is not convenient to judge the relative size between the coefficients. On the other hand, the coefficient of non-local transformer area should be strictly equal to zero in practical problems, so it is necessary to further compress the coefficient of least square. In this paper, the solution of Lasso is introduced. The Lasso method is to impose a norm penalty term on the basis of OLS, so that the coefficients of some variables are directly zero, and their optimization objectives are as follows:

\[
\hat{\beta}_j^{(c)} = \arg \min_{\beta} \| \mathbf{y}_c - X\beta \|_2^2 + \lambda \| \beta \|_1
\]

(3)

Among them, \( \lambda \) is a non-negative penalty coefficient.

The Lasso method can obtain a series of non-zero coefficients, which is convenient for variable selection. However, it is not use that that meter should all belong to this information of the same transformer area, so that part of the meter may be omitted in some cases. Therefore, this paper further constructs a Diff-Lasso method, by restricting the difference of the meter coefficient, the coefficient of the nearest neighbor is approximately equal, and the problem of omission in the selection can be solved to a certain extent. The specific optimization objectives are as follows:

\[
\hat{\beta}_j^{(c)} = \arg \min_{\beta} \| \mathbf{y}_c - X\beta \|_2^2 + \lambda \| \beta \|_1 + \lambda_2 \| \mathbf{L}\beta \|_2^2
\]

(4)
Among them, \( L \in R^{(m-1)\times m} \) is a coefficient difference matrix, that is,

\[
L = \begin{bmatrix}
-1 & 1 & \cdots & 0 \\
0 & -1 & 1 & 0 \\
0 & \cdots & -1 & 1
\end{bmatrix}
\]  (5)

According to reference [9], formula (4) can be further simplified to standard Lasso format. If \( \lambda = \lambda_1 / \sqrt{1 + \lambda_2} \), \( \beta_{arg}^{(c)} = \sqrt{1 + \lambda_2} \beta^{(c)} \), the following formula can be obtained:

\[
\hat{\beta}_{arg}^{(c)} = \arg \min \| \tilde{y}_c - \tilde{X} \hat{\beta}_{arg}^{(c)} \|^2 + \lambda \| \beta_{arg}^{(c)} \|
\]

\[
\hat{\beta}^{(c)} = \frac{1}{\sqrt{1 + \lambda_2}} \hat{\beta}_{arg}^{(c)}
\]  (6)

Among them, \( \tilde{X} = \frac{1}{\sqrt{1 + \lambda_2}} \left( X \sqrt{L} \right) \in R^{(n-m)\times m} \), \( \tilde{y}_c = \left( y_c, 0 \right) \in R^{n-m+1} \).

When transformed into standard Lasso, it can be solved according to LARS method, and the specific process can be referred to documentation[8]

The judgment process of topology attribution

For a total of \( C \) transformer areas, the corresponding meter can be selected for each area according to the following steps.

1. Collect the data of meter electricity consumption \( X \in R^{m\times n} \) and the total meter data \( Y \in R^{n\times C} \) of transformer area, and set the number of iterative steps \( c=1 \) (\( 1 \leq c \leq C-1 \)).

2. For transformer area \( c \), it is required to solve the Diff-Lasso coefficient, the meter with greater coefficient is selected as the user in the local area, and the corresponding data is set to zero.

3. Repeat step 2 and determine the final remaining meter to the last transformer area.

3. Example analysis

In order to verify the effectiveness of the proposed method, 70 days historical data from 3 transformer areas in a certain area are collected for simulation analysis. There are 3 meter boxes in each area, and 6 electric energy meters in each meter box. The specific topological relationship is shown in figure 1, among them "Transformer" represents the transformer area, "Box" represents the meter box and the outermost label represents the meter serial number. The penalty parameter is set to 20 and the filter parameter is selected to 0.5 in the Diff-Lasso method.
**Fig. 1** Topological structure diagram of transformer area

**Tab. 1** The home results of meter based on Lasso

| Transformer Area | Lasso Topology Results | misjudgment | erroneous judgement |
|------------------|------------------------|-------------|---------------------|
| 1                | 1,2,3,4,5,6,9,16,18,20,21,22,24,38,39,40,41,42 | 19,23,37 | 9,16,18 |
| 2                | 7,8,10,11,12,25,26,27,28,29,30,44,45,47,48 | 9,43,46 | |
| 3                | 13,14,15,17,19,23,31,32,33,34,35,36,37,43,46,49,50,51,52,53,54 | 16,18 | 19,23,37,43,46 |

**Tab. 2** The home results of meter based on Diff-Lasso

| Transformer Area | Diff-Lasso Topology Results | misjudgment | erroneous judgement |
|------------------|-----------------------------|-------------|---------------------|
| 1                | 1,2,3,4,5,6,19,20,21,22,23,24,38,39,40,41,42 | 37 | |
| 2                | 7,8,9,10,11,12,25,26,27,28,29,30,44,45,46,47,48 | 43 | |
| 3                | 13,14,15,16,17,18,31,32,33,34,35,36,37,43,49,50,51,52,53,54 | 37,43 | |

The home table information of each transformer area is obtained by using Lasso and Diff-Lasso methods, respectively, which are listed in tables 1 and 2. According to the form, the Lasso-based method can accurately identify the attribution of most of the meters, and the method is effective, but there is also missing or misjudgment. After using the difference constraint, the overall accuracy is greatly improved, and only two meters have errors in judgment.
Fig. 2 Comparison diagram of electric quantity fitting

Fig. 2 shows the comparison between the electric quantity curve and the actual curve fit by two methods. It can be seen that the two methods can basically follow the actual power curve trend, but it is clear that the Diff-Lasso method fits higher, especially for the latter two transformer areas. In order to quantitatively evaluate the performance of the two methods, the fitting degree of the curve is calculated by the following formula. It is recorded as \( \delta \).

\[
\delta = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right|
\]  

(7)

Among them, \( \hat{y}_i \) is obtained from the sum of the values of the electric energy meter at the same time, which represents the fitting value of \( y_i \). The closer the result of the above formula is to zero, the better the fitting effect is. It should be noted that even if the meter is completely correct in its ownership, but under some systematic errors, such as theoretical line loss, the index will still be slightly larger than zero. The fitting evaluation indexes of each transformer area are listed in Table 3, which shows that the Diff-Lasso method has an evaluation index which is closer to the real value.
Tab 3 The evaluating indicator of Meter ownership

| Area  | Area 2 | Area 3 |
|-------|--------|--------|
| The fact | 0.0329 | 0.0291 | 0.0300 |
| Lasso  | 0.0514 | 0.0658 | 0.0395 |
| Diff-Lasso | 0.0315 | 0.0316 | 0.0320 |

It is pointed out that the overall performance of Meter ownership based on statistical analysis method, it is closely related to the scale of the sample point. In theory, the more the sample is, the more accurate it is. Figure 3 shows the overall accuracy of two methods at different sample points. It can be seen that the accuracy of Lasso method is improved obviously with the increase of the number of samples; and the overall accuracy of the Diff-Lasso method is relatively high and relatively stable, and after the number of samples exceeded 160, it reached 100%. Therefore, this paper proposes to select more than twice the scale of the meter sample to establish a statistical model.

Fig. 3 Accuracy of different sample points

4. Conclusion
In this paper, the electric quantity data in the transformer area is analyzed in connection with the demand of power management, and a topology identification method for low-voltage area is proposed based on differential Lasso. The method only needs to construct the Diff-Lasso statistical model for the collected electric quantity data, and the attribution judgment of the area-meter can be completed without additional equipment assistance or a large amount of personnel workload. The simulation results of the field data illustrate the validity of the method, and can accurately identify the topology attribution of the low-voltage transformer area. In addition, it is worth further study to construct an accurate statistical model under the condition of small samples.

References
[1] Li Jian, Zhao Hanchang. The development and manufacture of multi-function Equipment for low voltage area identified [J]. Electrical Measurement & Instrumentation, 2014, 51 (13): 107-111.
[2] Chang Wenjing, Bai Jian. User’s transformer determining in the mode of electricity information collection system [J]. An Hui Electric Power Company, 2013, 30 (1): 56-58.
[3] Fan Di, Li Shaoshan, Li Haitao, et al. Study on the application in transformer-user identifying instrument. North china electric power, 2010, (7): 27-30.
[4] LI Ya, FAN Rusen, JIANG Wei, et al. Research on the intelligent transformer area recognition method based on BP neural network [J]. Electrical measurement & instrumentation, 2017, 54 (3): 25-30.
[5] YE Aliaosha, GU Jun, ZHANG Xiaqiu. Intelligent transformer room recognition technology based on temporal and spatial data correlation [J]. Modern Electronics Technique, 2018, 41 (4): 30-35.