Topography Automatic Identification Method for Low-Voltage Stations Based on Line Impedance Analysis

Yan Jiang¹, Xueying Song¹*, Hua Lin², Yupeng Zhao³, Kaiyi Qiu¹, Chao Yang⁵ and Shidanjie Dong⁶

¹ Beijing Fibrlink Communications Co., Ltd, Beijing, China
² State Grid Fujian Electric Power Research Institute, Fuzhou, Fujian, China
³ National Network Siji NetAn Technology (Beijing) Co., Ltd, Beijing, China
⁴ Application of System Integration Platform in Power Enterprise Information System, Hulihot, Inner Mongolia, China
⁵ State Grid Dalian Electric Power Supply Company, Dalian, Liaoning, China
⁶ State Grid Information & Telecommunication Group Co., Ltd., Beijing, China

*Corresponding author’s e-mail: songxueying@sgitg.sgcc.com.cn

Abstract. Aiming at the problems of high cost, low real-time performance of manual identification of low-voltage station topological structure, automatic identification methods such as instantaneous power failure method and voltage correlation analysis method, which affect residents' electricity consumption, and the success rate is limited by application scenarios, this paper proposes a method based on multiple linear regression Line impedance analysis method for automatic identification of the topology of low-voltage stations. Based on the relationship between voltage, current, and impedance, multiple linear regression analysis is used to derive the upstream topology step by step from the bottom up to the top of the topology and then determine the station membership by voltage correlation. The cloud platform based on Docker technology and Hadoop technology is used for data distributed storage calculation, and an identification model in the context of big data is established to effectively improve the efficiency of the algorithm. Finally, the correctness of the algorithm verification is judged based on the results of manual field verification.

1. Introduction

The automatic identification of low-voltage station area topology is a process of calculating the connection relationship of various electrical equipment in the community, and building through the identification method, and storing these connection relationships in a certain way. Topology identification plays a very important role in the line loss calculation and energy-saving analysis of low-voltage distribution systems. When the topological relationship is wrong, it will affect the accuracy of the line loss calculation in the station area, and even the safe operation of the power grid in severe cases [1].

Based on research and application at home and abroad, there are two main types of low-voltage station topology identification: manual field identification and automatic identification. Manual identification mainly relies on manual inspection methods, that is, in accordance with the distribution and direction of the low-voltage station area distribution network, manual surveys are carried out.
section by section using a handheld household change identification instrument. Because the low-voltage distribution network has many meshed distribution lines, cable trenches or overhead lines. In contrast, automatic identification is more efficient, but the success rate is limited by application scenarios. At present, the commonly used automatic identification methods mainly include power failure record matching method\textsuperscript{[2]}, HPLC module identification method\textsuperscript{[3]}. The power outage record matching method requires a differentiated power outage or power-on time in adjacent stations. If the two stations have the same outage and power-on time, they cannot be used for household-transformation relationship verification. At the same time, for a station area with a new meter, because it cannot read earlier power outage records, the station area cannot be judged based on the power outage records. For the HPLC module identification method, during the day (the load is lighter at night, the communication effect is better), some line environments with heavier loads or particularly complex lines in the station area will exist in the same station area because of strong line interference and large signal attenuation. The point-to-point successful power line carrier communication cannot be reached between the outlet of the low-voltage bus and the user, resulting in the false appearance that the metering device is not under the station. In some new communities, due to the fact that there are not many occupants, the load is light, the grid wiring is standardized, the wire diameter is thick, and the impedance is relatively small, especially when the power line carrier communication capability is strong, there may be a test host cross-station area (Even if the signal isolation and attenuation of two transformers exist), it can successfully communicate with the detection terminal, causing the false appearance that the metering device is under the station area.

Based on this, this paper proposes a low-voltage station topology automatic identification method based on line impedance topology analysis, which can not only accurately identify the relationship between households and changes, but also simulate the branch impedance of the station topology to realize the visualization of the station topology. The big data platform realizes the efficient calculation of massive data.

2. Grouping of adjacent stations

In the low-voltage distribution network, the chaos of the topological relationship between stations is mainly manifested in the abnormal relationship between households and changes. The abnormal household change relationship is mainly caused by two reasons. One is that the low-voltage distribution network topology is wrong due to the error of the user's meter line file during the construction of the low-voltage distribution network; the second is that the low-voltage station area distribution network is faulty and is undergoing inspection and repair. When adjusting the wiring, the change of the meter file did not record or the record was wrong, which caused the low-voltage distribution network topology relationship error\textsuperscript{[4]}. Through a large amount of historical data analysis, it can be found that abnormal household change relationships usually occur in adjacent stations. Neighboring stations refer to stations that are relatively close in electrical distance.

In the low-voltage distribution network, the voltage often fluctuates due to the uncertainty of the load everywhere. Loads with relatively close electrical distance have relatively similar voltage fluctuation curves (high correlation), while loads with relatively long electrical distance have low voltage fluctuation curves (low correlation). The similarity of the two curves can be visually compared by drawing, or quantitatively by mathematical methods. The most commonly used Pearson correlation coefficient for correlation analysis. The formula is as follows:

\[
\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}}
\]  \hspace{1cm} (1)

The closer the coefficient is to 1, the higher the correlation, and the closer the coefficient is to 0, the lower the correlation.
Therefore, the voltage curve of the total meter of the station area can be selected as the basis for judging the electrical distance of the station area. Take the station area under the same bus node as a large group, calculate the voltage correlation coefficient of the station area table separately regardless of the phase and take the maximum value, divide the adjacent station area grouping according to the threshold value, then $X_i$ and $Y_i$ in the formula (1) respectively represent the voltage of the two stations, $i$ corresponds to different sampling moments, $\bar{X}$ and $\bar{Y}$ respectively represent the average value of a certain zone voltage.

The automatic identification of low-voltage station topology involved in the article is for the identification within the adjacent station group.

### 3. Principle of Line Impedance Analysis Method

The line impedance analysis method uses the relationship between voltage, current and impedance to derive the upstream topology step by step from bottom to top, until the top of the topology, it is more accurate to judge the subordination of the station area through the voltage correlation \(^5\) than using the voltage correlation analysis method alone.

Assuming that node 1 and node 2 have a common upper node 0, there are:

$$U_0 = U_1 + R_1I_{1,R} + X_1I_{1,X} = U_2 + R_2I_{2,R} + X_2I_{2,X}$$  \(2\)

$$2U_0 = U_1 + R_1I_{1,R} + X_1I_{1,X} + R_2(-I_{2,R}) + X_2(-I_{2,X})$$  \(3\)

By judging the fit of the regression equation (coefficient of determination $R^2$), it can be inferred whether node 1 and node 2 actually have a common upper node.

### 4. Three-phase unbalanced line impedance analysis method

The line impedance analysis method shown in Section 2 only considers the single-phase situation. When the three-phase four-wire power supply is used and the three-phase is unbalanced, the neutral line impedance and the neutral line current need to be considered, and the node is three-phase node.

$$U_0 = \frac{(U_1 + R_1I_{1,R} + X_1I_{1,X}) + (U_2 + R_2I_{2,R} + X_2I_{2,X})}{2}$$  \(4\)

$$I_0 = I_1 + I_2$$  \(5\)
According to Figure 2, for each phase of the node, modify the pressure drop equation as follows:

$$U_{0,L} = U_{1,L} + R_{1,L}I_{1,L,R} + X_{1,L}I_{1,L,X} + R_{1,N}I_{1,N,R} + X_{1,N}I_{1,N,X}$$

(6)

Figure 2: Line impedance branch diagram considering zero sequence impedance and zero sequence current

The article temporarily thinks $R_L = R_N$, $X_L = X_N$ [7]. Then:

$$U_{0,L} = U_{1,L} + R_{1,L}(I_{1,L,R} + I_{1,N,R}) + X_{1,L}(I_{1,L,X} + I_{1,N,X})$$

(7)

Since the voltage phase deviation of the entire network cannot be obtained, the three-phase phase difference of all nodes can only be regarded as just 120° [8]. Take phase A as an example, sum the zero line current through three phase line currents, and take the phase A voltage as the reference phase, then:

$$\begin{align*}
I_a &= 2I_{1,A} \cos \varphi_a + I_{1,B} \cos(\varphi_b + 120°) + I_{1,C} \cos(\varphi_c + 240°) \\
&= 2I_{1,A} \cos \varphi_a - \frac{1}{2}I_{1,B} \cos \varphi_b - \frac{\sqrt{3}}{2}I_{1,B} \sin \varphi_b - \frac{1}{2}I_{1,C} \cos \varphi_c + \frac{\sqrt{3}}{2}I_{1,C} \sin \varphi_c \\
I_{a,x} &= 2I_{1,A} \sin \varphi_a + I_{1,B} \sin(\varphi_b + 120°) + I_{1,C} \sin(\varphi_c + 120°) \\
&= 2I_{1,A} \sin \varphi_a - \frac{1}{2}I_{1,B} \sin \varphi_b + \frac{\sqrt{3}}{2}I_{1,B} \cos \varphi_b - \frac{1}{2}I_{1,C} \sin \varphi_c - \frac{\sqrt{3}}{2}I_{1,C} \cos \varphi_c
\end{align*}$$

(8)

(9)

$\cos \varphi$ is the power factor, the load curve of the smart meter is recorded, $\sin \varphi$ can be converted by $\sqrt{1-\cos^2 \varphi}$. By calculating the correlation between the second-level node and the transformer voltage curve of the station area, the station area membership can be judged, and all downstream nodes, that is, the station area membership of the meter box can be determined. The connection relationship between the nodes is represented by the connection line, and the horizontal length of the connection between the nodes represents the impedance size, and the topological map of the production area can be drawn.

5. **Algorithm steps**

1) Suppose the set of all bottom nodes is $A$, and obtain the $U$, $I_R$, $I_X$ curves of each node according to the phase.

2) Take every two nodes in the same phase in set $A$ as a pair for regression analysis.

3) The pair with the largest average value of the three-phase R2 is regarded as having a common upper node. Calculate the equivalent three-phase voltage and current values $U$, $I_R$, $I_X$ of the upper node $k$. Delete the two nodes of this pair from set $A$, and add their common upper node $k$. 


4) Repeat step 2) until only one node remains in the set A.
5) By calculating the voltage correlation between the next-level node and the concentrator, the subordinate of all nodes downstream can be determined.

6. Algorithm verification

The algorithm verification is implemented on a cloud platform based on a distributed architecture. The data source is the electricity consumption information collection system data of a community. The community contains two residential buildings with a total of 140 users, 12 meter boxes, 135 single-phase meters, and 5 A three-phase meter, powered by two 315kVA transformers. The smart meter collects once every 15 minutes, 96 data points a day. Because the phase voltage difference in the box is too small, it is difficult to calculate the impedance. The article starts from the first level of the meter box. After the regression iteration of the line impedance analysis method, the results shown in Table 1 are obtained.

| Rounds | Add upper node | Lower node 1          | Lower node 2          | R²    |
|--------|----------------|-----------------------|-----------------------|-------|
| 1      | 84             | 354010389385          | 354010389386          | 0.9994|
| 2      | 85             | 354010389381          | 354019022839          | 0.9992|
| 3      | 86             | 354010389383          | 84                    | 0.9988|
| 4      | 87             | 85                    | 86                    | 0.9984|
| 5      | 88             | 354010389382          | 354010389384          | 0.9983|
| 6      | 89             | 354010389388          | 354013332709          | 0.9976|
| 7      | 90             | 354019022841          | 88                    | 0.997 |
| 8      | 91             | 89                    | 90                    | 0.997 |
| 9      | 92             | 354019022843          | 87                    | 0.9964|
| 10     | 93             | 354019022842          | 91                    | 0.9957|
| 11     | 94             | 92                    | 93                    | 0.9952|

By determining the affiliation of the sub-top node, the affiliation of all nodes downstream can be determined. The correlation coefficients between the voltage curves (derived) of nodes 92 and 93 and the true voltage curves of the two concentrators are shown in Table 2:

| Derive node | Concentrator | Phase A | Phase B | Phase C |
|-------------|--------------|---------|---------|---------|
| 92          | 019100255682 (1#) | 0.990   | 0.995   | 0.993   |
| 92          | 019100255684 (2#) | 0.985   | 0.994   | 0.988   |
| 93          | 019100255682 (1#) | 0.980   | 0.991   | 0.987   |
| 93          | 019100255684 (2#) | 0.992   | 0.998   | 0.997   |

Therefore, it is judged that node 92 belongs to transformer 1#, and node 93 belongs to transformer 2#. The final topology is shown in Figure 4.
The horizontal length of the connection between the nodes in the above figure represents the impedance. It can be seen that the meter box 35401332709 is the farthest from the transformer electrically, while the electrical distances 86, 87, 92 and 88, 90, 91, 93 are very close, and it may be the same. Node or on the same bus. Through manual on-site hand-held household variable instrument detection, the identification results of the station area are consistent with the detection results. Compared with the file data, it is found that the file belonging to the meter box file of 354010389388, 354019022841 and 354019022843 is wrong. After on-site inspection, the electrical distance between the meter box and the transformer is consistent with the topological diagram.

7. Summary
Aiming at the problems of inaccurate household-transformation relationship and difficulty in identification involved in low-voltage distribution network topology, this article proposes an automatic identification method for low-voltage station topology based on line impedance analysis. First, use Pearson's voltage correlation analysis to construct adjacent station groupings, and then use multiple linear regression analysis to derive the upstream topology from bottom to top through the relationship between voltage, current and impedance, and then pass the voltage correlation judgment station to the top of the topology. District affiliation. At the same time, it is based on the cloud platform for data distributed storage and calculation, and the identification model in the context of big data is established. The cloud platform is built based on Docker technology and Hadoop technology to ensure the efficient and reliable operation of the algorithm. Finally, the operation and maintenance personnel are arranged to go to the site to confirm and judge the correctness of the results to be verified. The verification method requires low labor cost, fast calculation speed, strong real-time performance, strong operability, and has high application value.

References
[1] Li H, Zhang Z.L, Gao YJ. (2015) Discussion on an Identification Method of Distribution Network Topology. China Power, 48:133-138.
[2] Qian LJ, Lu HX, Huang Y. (2015) Realization and Application of Station Power Outage Monitoring Function Based on Collection Terminal. Electrical Application, S1:212-214.
[3] Zheng J, Zhao JF, Wang Haijiang, Huang XD. (2015) A power phase identification method based on carrier communication. Electrical Measurement and Instrumentation, 52:77-79.
[4] Xiao Y, Zhao Y, Tu ZD, et al. (2019) A verification method of low-voltage distribution network topology based on improved Pearson correlation coefficient. Power System Protection and Control, 11:37-43.
[5] Short, Tom A . (2013) Advanced Metering for Phase Identification, Transformer Identification, and Secondary Modeling. IEEE Transactions on Smart Grid, 4:651-658.
[6] Luan W, Peng J, Maras M, et al. (2013) Distribution network topology error correction using smart meter data analytics. In: Power and Energy Society General Meeting (PES), Vancouver. pp. 1-5.

[7] Cavraro G, Kekatos V, Veeramachaneni S. (2013) Voltage Analytics for Power Distribution Network Topology Verification. IEEE Transactions on Smart Grid, 1:1-1.

[8] Liu C, Jiang Y, Zhao F, et al. (2014) Design and applied research of the distributed real-time database in smart grid. In: International Conference on Advanced Communication Technology. Pyeongchang. pp. 719-724.