Research on Real-time Automatic Acquisition Technology of Topology Data in Distribution Network

Jing Chen*, Chao Lin, Yingwei Xue, Jiandi Hu and Jiaying Li

State Grid Info-Telecom Great Power Science and Technology Co., LTD, Fuzhou Fujian 350003, China

*Corresponding author email: jingchen@sgcc.com.cn

Abstract. The current distribution network equipment renovation plan arrangement is mainly based on the equipment health level, operating years and other indicators. Coordination and optimization between renovation projects are still lacking. Scattered and disorderly transformation will cause potential safety hazards in the distribution network operation. The paper proposes to establish a unified real-time data collection platform for distribution network based on real-time/historical database, and gives the construction idea and implementation plan of the platform. The designed platform has data interfaces with mainstream production control systems, user meters and electrical related systems, and integrates a wealth of common software tools, which can manage and analyze the real-time/quasi-real-time data of the distribution network in a unified manner and can facilitate coordinated with the implementation of advanced applications of smart distribution networks, it has laid the foundation for the analysis and control of the optimized operation of the entire grid/local area grid and the construction of an integrated model of smart grid operation and distribution. At the same time, the paper proposes a method for checking the topology of low-voltage distribution networks based on outlier detection, calculates the correlation coefficient between the voltage sequence data to measure the similarity of the voltage curves of different users, and finally verifies the users and stations in the grid GIS system The correctness of the zone transformer topology connection relationship.

Keywords: Real-time data, unified acquisition platform, eDNA database, network topology.

1. Introduction

The distribution network is an important foundation for the energy internet and a key foothold for serving the society and the people's livelihood. State Grid Corporation strives to build a first-class modern power distribution network with high reliability, friendly interaction, and cost-effectiveness. Investment continues to increase and management requirements continue to increase. The distribution network equipment is multi-faceted, the project is small, and the construction period is short. Restricted by traditional management ideas and methods, the phenomenon of "heavy construction and light planning" is more common. Scientific planning is the key to leading the high-quality construction and development of the distribution network. In the era of big data, starting with data, in-depth analysis of
distribution network problems based on massive data in the distribution network, accurate positioning of weak links in the distribution network, and orderly arrangement of distribution Grid projects have become an important part of scientific planning and research. Due to the wide coverage of the power grid, the complex structure of the power system, and the huge number of services involved, the distribution and deployment of the distribution network relies heavily on manual on-site verification, lack of effective means for verification and identification, and a lack of a complete and accurate source-network-load full topology Graphics, intelligence and automation need to be improved. Therefore, under the new technical conditions, in-depth research on the automatic integration of grid planning data based on ICT new technologies and intelligent management and control technology for planning execution, and real-time automatic collection of distribution network topology information are important for the improvement of distribution network management efficiency and development quality. Meaning.

Analyze various new technologies of ICT, design the basic technical framework of real-time data collection, and complete the collection of the topological relationship of the distribution network, realize the automatic access to the professional data of the distribution network such as operation inspection, marketing, dispatching, etc. A professional, multi-topic, and multi-application unified distribution network database provides the basis for distribution network diagnosis and big data analysis, and solves the problem of "manually looking for data and problems; project professional report, landing is unknown; the effect is unclear, right "All standards are good" and other problems in distribution network planning and management, promoting online operations, automatic identification, online analysis, and equipment management and control of distribution network planning services [1].

2. Real-time automatic gathering demand analysis of distribution network topology data

2.1. Topological structure analysis of distribution network
The grid of the distribution network is a tree structure, with the substation as the starting source, and the "station-line-transformation-household" hierarchy is a subordinate relationship, which can be directly supplied after distribution through the distribution line and after the distribution transformer is stepped down User use. Through the topological connection relationship of each level of the distribution network, combined with the operation data of the distribution equipment, it can provide a more comprehensive feedback on the healthy operation level of the distribution network [2].

2.2. Monitoring requirements of distribution network operation equipment
Substation: monitor the interval switches and outgoing lines of each unit of the substation; distribution line: monitor the interval switches and outgoing lines of each unit such as the ring cabinet, switchgear and branch box on the distribution line; switchgear: monitor the units of the switchgear Interval switch and outgoing line; ring network cabinet: monitor the interval switch and outgoing line of each unit of the ring network cabinet; branch box: monitor the interval switch and outgoing line of each unit of the branch box; distribution transformer: monitor the distribution line relationship of the distribution transformer; User power meter: monitor the current user power supply point.

2.3. Distribution network topology data collection requirements
Station-line relationship: In the distribution network structure, the relationship between the substation and the distribution line; Line change relationship: In the distribution network structure, the relationship between the distribution line and the distribution transformer; Platform relationship: In the distribution network structure, the distribution The relationship between the transformer and the user; branch: the relationship between the backbone and the branch under the distribution line in the distribution network structure.

2.4. Distribution network topology data collection, monitoring and early warning requirements
The topological data collection range of the distribution network is the substation as the starting point, and it penetrates through layers, involving distribution lines, ring cabinets, switchgears, branch boxes,
distribution transformers, user meters, etc., according to the distribution network optimization and upgrade construction and synthesis Business development analysis requires real-time monitoring of the equipment below the substation, early warning to ensure the reliable operation of the distribution network, and auxiliary support for the high-quality and efficient development of the distribution network. The substation needs to be combined with infrastructure projects and line transformation, etc. Monitor and guarantee the basic data quality of the camp. As shown in Table 1.

### Table 1. Distribution network topology data collection, monitoring and early warning requirements

| Topological relationship | Distribution network equipment       | Acquisition frequency |
|--------------------------|-------------------------------------|-----------------------|
| Station line relationship| Substation                          | 24h                   |
|                          | Distribution line                   | real time             |
| Linear relationship      | Distribution transformer            | real time             |
| Change relationship      | User meter                          | real time             |
| Branch                   | Ring network cabinet                | real time             |
|                          | Opening and closing                 | real time             |
|                          | Branch box                          | real time             |

3. **Overall architecture of real-time automatic collection of topology data of distribution network**

Integrated the influence factors of the distribution network topology relationship, the overall architecture of the real-time automatic collection technology of distribution network topology data is divided into a perception layer, a data layer and an application layer. Deploy automatic collection terminals at the perception layer for real-time monitoring, early warning of risks, and release device operating data to the data layer. After the data layer cleans and converts the data, it is released. After the application layer receives the data results, it displays for related businesses, such as Figure 1.

![Figure 1. Overall structure of real-time automatic collection of topology data of distribution network](image-url)
4. Real-time automatic collection technology of distribution network topology data

4.1. Communication Technology Research

4.1.1. Two-way power frequency communication technology. The core idea of the two-way power frequency communication technology (TWACS) is to use the small distortion of the grid voltage and current waveforms to carry the required transmission information. Its modulation signal is very special, with small distortion of the grid voltage and current as the modulation signal, and each modulation signal can be attenuated in a fundamental cycle, without affecting the modulation signal of the next cycle. Based on this characteristic of the modulated signal, two-way power frequency communication is very suitable for differential Manchester encoding, that is, every two adjacent fundamental wave periods change (one carries a modulated signal and the other has no modulated signal). Define "0", no change (two both carry modulated signals or both have no modulated signals) Definition "1".

The transformers in the substations of the distribution network all have high leakage inductance to prevent overcurrent, and the system energy tends to be minimum near the zero-crossing point of the voltage waveform. Therefore, the energy requirement for modulating the signal at the zero-crossing point is small, and the particularity of the zero-crossing point of the voltage provides conditions for the modulation and detection of the signal positioning. Therefore, the TWACS communication system can use the fundamental zero-crossing point of the voltage waveform as the modulation position. Figure 2 below shows voltage modulation and current adjustment.

![Figure 2. Is the voltage modulation and current adjustment](image)

The encoding method of TWACS input signal is more than this one. The prescribed 8 modulation positions modulate only 4 of them, and two positive and two negatives, so that the resulting coding method has a total of 36 groups of 18 pairs. In these codes, 5 orthogonal code sets can be excluded, and each orthogonal code set contains 6 codes and are orthogonal to each other. According to such coding characteristics, 4 adjacent current cycles can simultaneously carry 6 channels to improve the throughput of communication transmission information [3].

There are generally two methods for the detection of modulated signals: time difference method and amplitude difference method. Relatively speaking, the latter is more reliable. After the PT\CT of the
fundamental wave signal of the power grid is amplified, the bias level is added to make it suitable for the input requirements of AD. The difference between the quantized value of the previous period after quantization and the background signal is eliminated, and the modulated signal can be extracted. Due to the change of the grid frequency and the timing error of the single-chip microcomputer, the angles of the sampling points of the two periods before and after cannot be guaranteed to be exactly the same. The characteristics of the signals of the Nth and N+1th cycles are:

$$u_N = U \sin(\omega t + \alpha)$$

$$u_{N+1} = U \sin(\omega t + \beta)$$

$$u_{N+1} - u_N = U \sin(\omega t + \beta) - U \sin(\omega t + \alpha)$$

$$= 2U \cos \left(\frac{\omega t + \beta - \omega t - \alpha}{2}\right) \sin \left(\frac{\omega t + \beta - \omega t - \alpha}{2}\right)$$

In the formula, $\alpha \approx \beta, \alpha, \beta \in [0, 2\pi]$. It can be seen from the above formula that after the difference between two adjacent cycles, the resulting signal retains the characteristics of the original signal, and has a large reduction in amplitude, and a phase difference of approximately 90°, which is equivalent to a derivative operation. When the period of two adjacent fundamental waves changes, the modulation signal can emerge through the difference method.

4.1.2. Comparison with traditional HPLC communication. The two-way power frequency automatic communication method is a new technology developed in recent years. Its application range is the power distribution network. It is the same as the traditional PLC transmission. The modulation signal is on the power line. The difference is that the modulation frequency of this method is very low. The method is special.

Compared with the traditional carrier wave method, this two-way industrial communication method (TWACS) based on power distribution network has many advantages: the method also uses the existing power network as a transmission carrier, saving a lot of equipment costs; the signal is in the transmission process Medium, no leakage and bypass, small attenuation, no need of filter and wave stopper, using two-way power frequency communication technology, can achieve cross-regional communication, without adding equipment on the transformer, reducing regional restrictions; can be achieved Complete two-way communication, the upstream and downstream channels do not interfere with each other, and can carry out multi-channel communication; the visit to the remote equipment or instrument is direct without the need for relay links; there is no interference with the power grid itself, which is completely within the allowable range; Its own frequency and amplitude changes are not sensitive, and its anti-interference ability is strong; the signal is modulated near the zero crossing point, the required modulation power is small, and it is easy to implement. The disadvantage is that in comparison, the transmission rate of the TWACS method is much lower than that of the carrier method. Although a certain algorithm and combined modulation can be used for multi-channel communication, the transmission rate is improved to a certain extent, but it still belongs to low-frequency transmission. Therefore, its main application is when the communication rate is not high, such as: automatic meter reading system, remote load control, etc [4].

4.2. Distribution network topology identification scheme
In line with the principle of “allowing to simplify the setting of the safety zone based on the actual situation of the application system and meeting the overall safety requirements, but should avoid the formation of vertical cross-connections of different safety zones through the WAN”, this paper will Variable monitoring systems, such as systems with both real-time data collection and control functions, are grouped into the control area. Systems with electricity information collection systems that have real-time data collection functions but no control functions are grouped into the management information
The unified real-time data collection platform can simplify the distinction between the production control area and the management information area [5]. The overall network topology is shown in Figure 3.

5. Introduction to automatic algorithms for data collection

5.1. Algorithm introduction

At present, the commonly used clustering algorithms have two methods based on division and layering. Traditional clustering algorithms generally have the disadvantages that it is difficult to determine the clustering centre, there is an irrelevant dimension in the high-dimensional space to mask effective clusters, and it is easy to produce meaningless clusters. In this paper, the inner square distance method is used to observe the change of the data square error in the clustering process, accurately reflect the clustering degree of clusters, quickly determine the initial clustering centre, and cluster the data from the bottom up. It can effectively avoid some shortcomings of common clustering algorithms. For example, adjacent objects may be separated as the centre points of different classes. The data cluster is regarded as a circular set containing n similar data. The inner square distance refers to the total square of the Euclidean distance of the data in two data clusters as the distance between the two data clusters.

The distance between the data must be Euclidean distance, the basic idea comes from analysis of variance, and the formula is expressed as

$$d(v, w) = \sqrt{\frac{2n_v n_w}{n_v + n_w}} \| x_v - x_w \|_2$$

Where: $v$, $w$ are data clusters containing different data; $n_v$, $n_w$ is the number of data contained in data clusters $v$ and $w$; $x_v$, $x_w$ is the centre of data clusters $v$ and $w$; $\|x_v - x_w\|_2$ is the Euclidean distance between data clusters. It is applicable to the 2 and 3 dimensional spaces and represents the shortest distance between two points. For a data cluster, the Euclidean distance is the shortest distance between the data phasors in the centre of the data cluster, $d_{vw} = (x_v - x_w)^T(x_v - x_w)$.

The correlation coefficient (cophenetic) is an index used to verify the clustering results. For actual objects and specific applications, different distance calculation methods will produce multiple classification results with different bifurcations for decision-making reference. The correlation coefficient can check the degree of consistency between the clustering result generated by a certain
algorithm and the actual situation, that is, the correlation between the distance between each element in
the cluster and the actual distance generated by the calculation. The correlation coefficient of phasor
\( x_i, x_j \) is expressed as

\[
c_{ij}(2) = \frac{\sum_{k=1}^{n} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{\left( \sum_{k=1}^{n} (x_{ik} - \bar{x}_i)^2 \right)^{0.5} \left( \sum_{k=1}^{n} (x_{jk} - \bar{x}_j)^2 \right)^{0.5}}
\]

The size of the correlation coefficient reflects the clustering effect. The closer the correlation
coefficient is to 1, the better the clustering effect. Correlation coefficients can be used to verify and
compare the clustering effects of different distance calculation methods and different system clusters,
and select reasonable clustering results [6].

5.2. Example verification

The 10 kV power grid has the characteristics of complex structure and large number of equipment.
Taking the 10 kV power grid connected by the T3 substation in Figure 4 as an example, the typical
multi-segment moderate connection network topology is intercepted as shown in Figure 5. The network
includes DT1- There are 5 DT5 distribution transformers, 6 transmission lines from L1 to L6, and 13
switchgears from CB1 to CB13. The calculation example takes the power supply area of the 10 kV
busbar on the low-voltage side of the 110 kV substation as the basic transformation unit, calculates the
equipment operation years and health levels in the area, and obtains the weighted average of various
equipment index scores as the operating state of the basic transformation unit. Index. Take the operation
status index scores of various devices in the D3 transformation unit D3 of the low voltage side of the T3
substation in Figure 4 as an example, as shown in Table 2. In the same way, the operating state scores
of D1-D5 transformation units D1-D5 on the low voltage side of T1-T5 substation can be obtained, as
shown in Table 3.

![Figure 4. Topological structure diagram of 10 kV power grid](image-url)
Figure 5. Schematic diagram of clustering division based on transformation unit status indicators

Table 2. Main equipment specifications of 10 kV distribution network transformation unit on the low voltage side of T3 substation

| Indicator type                                      | Single device index score | Index score weighted mean |
|-----------------------------------------------------|---------------------------|----------------------------|
| Index of operating years of distribution transformers | DT1=5/30; DT2=8/30; DT3=13/30; DT4=21/30; DT5=7/30; | 0.36                       |
| Distribution Transformer Health Level Index         | DT1=0.21; DT2=0.43; DT3=0.61; DT4=0.72; DT5=0.26; | 0.45                       |
| Circuit breaker operating years index               | CB1=2/30; CB2=6/30; CB3=12/30; CB4=9/30; CB5=11/30; | 0.32                       |
| Whether it is an oil circuit breaker indicator      | CB1=0; CB2=0; CB3=1; CB4=0; CB5=0; | 0.21                       |
| Line operation life index                           | L1=2/30; L2=5/30; L3=12/30; L4=6/30; L5=14/30; | 0.28                       |

Table 3. Main equipment specifications of the 10 kV distribution network transformation unit on the low voltage side of each 110 kV substation

| project                                           | Operating years of distribution transformers | Distribution transformer health level | Circuit breaker operating years | Proportion of oil circuit breaker | Line operation years |
|---------------------------------------------------|-----------------------------------------------|-------------------------------------|--------------------------------|----------------------------------|----------------------|
| D1: 10 kV distribution network in T1 substation    | 0.28                                          | 0.25                                | 0.23                          | 0.02                             | 0.27                 |
| D2: 10 kV distribution network in T2 substation    | 0.80                                          | 0.72                                | 0.63                          | 0.33                             | 0.57                 |
| D3: 10 kV distribution network in T3 substation    | 0.36                                          | 0.45                                | 0.32                          | 0.15                             | 0.28                 |
| D4: 10 kV distribution network in T4 substation    | 0.55                                          | 0.48                                | 0.46                          | 0.23                             | 0.49                 |
| D5: 10 kV distribution network in T5 substation    | 0.72                                          | 0.63                                | 0.54                          | 0.32                             | 0.55                 |
| D6: 10 kV distribution network in T6 substation    | 0.24                                          | 0.21                                | 0.19                          | 0.01                             | 0.18                 |
The T3 and T4 substations are in the ring network connection structure. Any modification of one of the substations will not affect the power supply of other substations in the ring network. However, if the lines L2 and L3 connecting the substation T3 are simultaneously modified, the T3 substation will lose power supply if the power supply is lost, the same lines L3 and L4 should also be avoided at the same time. Comprehensive analysis of the equipment operating status indicators shows that the operating status of lines L2 and L4 is relatively poor, while the operating status of line L3 is relatively good. Therefore, it is considered to divide T3, T4, L3 into a transformation zone, and at the same time arbitrarily choose to include L2 or L4 in the partition. Reconstructing the partition can not only prevent the total amount of equipment in the partition from being too large and affect the reliability of power supply, but also prevent the partition from being too scattered. This article chooses to divide T3, T4, L2, L3 into the same transformation zone, and carry out unified transformation planning and arrangement, according to the priority of equipment transformation needs, rationally arrange the transformation sequence of lines L2, L3.

6. Summary

In order to realize the automatic identification of the distribution network topology, especially the real-time identification of the occurrence of the change of the relationship between the households and the users, this paper first intends to use the two-way power frequency communication technology (TWACS) as the implementation method, and analyses and determines the principle and its advantages over traditional carrier technology. After that, the topology of the distribution network is divided into two parts: the medium-voltage side and the low-voltage side. The design topology automatic identification scheme. Due to the difference in voltage level and identification requirements, the medium-voltage side and the low-voltage side have certain differences in the implementation of the scheme. In the end, it provides a more feasible solution for the automatic identification of the distribution network topology, which lays the foundation for the pilot implementation verification.

Acknowledgments

This paper is funded by the Technology project of State Grid Corporation of China, "Research and Application of Power grid planning Data Automatic Integration and Planning Execution Intelligent Control Technology Based on NEW ICT"

References

[1] Tong, F. He, S., & Pan, J. Modeling and analysis for data collection in duty-cycled linear sensor networks with pipelined-forwarding feature. IEEE Internet of Things Journal, 6 (6) (2019) 9489-9502.

[2] Chen, L. W. Peng, Y. H., Tseng, Y. C., & Tsai, M. F. Cooperative sensing data collection and distribution with packet collision avoidance in mobile long-thin networks. Sensors, 18 (10) (2018) 15-20.

[3] Gabr, H. Todor, A., Dobra, A., & Kahveci, T. Reachability analysis in probabilistic biological networks. Computational Biology & Bioinformatics IEEE/ACM Transactions on, 12 (1) (2015) 53-66.

[4] Cavraro, G. & Arghandeh, R. Power distribution network topology detection with time-series signature verification method. IEEE Transactions on Power Systems, 33 (4) (2018) 3500-3509.

[5] Luan, W. Peng, J., Maras, M., Lo, J., & Harapnuk, B. Smart meter data analytics for distribution network connectivity verification. IEEE Transactions on Smart Grid, 6 (4) (2015) 1964-1971.

[6] Zan, G. Wang, Z., Li, Q., Wang, L., & Li, W. High voltage distribution network reliability evaluation based on state space truncation and isolation scope derivation. Dianli Xitong Zidonghua/automation of Electric Power Systems, 41 (13) (2017) 79-85.