Deep Perception of Power Failure of Substation Area in Smart Grid

Liyan Kang¹, Ying Shang¹, Lina Sun¹, Zhongcheng Li¹ and Muxin Zhang¹

¹ Electric Power Research Institute of Liaoning Electric Power Co., Ltd., Shenyang, Liaoning, 110006, China

Abstract: Power failure of substation areas is a regional power failure event. It is difficult to judge whether or not the failure occurs in the substation area from the status of a single equipment. Moreover, the local network is interrupted after the power failure of substation area, and the power grid equipment is disconnected, so the information of the substation area cannot be perceived. Build high-speed communication network of substation area in smart grid, increase emergency power supply function of equipment of substation area, strengthen the perception ability of equipment status of substation area, introduce data mining analysis method to deal with massive perception data, and strengthen the processing ability of perception information to form deep perception of power failure in substation areas, which make it possible to fast, accurately and completely judge the power failure of substation area. Unexpected power failure or manual-drive power failure of substation can be used to support power grid construction and disaster prevention.

1. Introduction

It is necessary to accurately judge the power failure events of substation area and deal with the faults in time to avoid serious impacts such as chain power failure[1]. Deep perception of power failure of substation area is based on the perception ability and information processing ability of equipment. It not only needs to be able to directly perceive each point, but also aggregates perception points to form perception aggregation, which has deeper and wider perception ability. After the power failure of substation area, the network of the area is interrupted, the equipment is out of service, and the basic perception is lost. Moreover, the state perception of a certain device cannot be used to determine a regional power failure event[2].

In order to deeply perceive the power failure events of substation area, this paper proposes to optimize the local communication network and enhance the data transmission capability; update the equipment in the substation area with the terminal equipment with disaster preparedness power supply to ensure the failure information perception and timely transmission. In addition to improving the basic perception ability, this paper also focuses on the use of big data analysis method to process perception information. The XGboost ensemble algorithm is used to establish multiple decision tree models to study the power failure events in the substation area, so as to form the perception aggregation of power failure of substation area and realize the deep perception of power failure of substation area.

2. Analysis of the perceptive ability of current power failure of substation area

At present, the State Grid Corporation has completed the construction of the power information acquisition system. The power information acquisition system is an important part of the smart grid.
The system mainly relies on the narrowband power line carrier technology to communicate within the scope of the substation area. At the same time, it uses RS-485 serial bus communication, micro-power wireless communication, NB-IoT narrowband Internet of Things communication as necessary supplements. Smart meter, collector and concentrator are used as basic equipment to form perception layer, which can realize terminal perception of user's power consumption information and equipment status information. The communication between the substation network and the main station of the system is mainly based on the 4G wireless communication network. The main station of the system can receive a large number of users' electricity data and equipment status data every day.

Although the power information acquisition system has the basic perception ability, this perception is still in the initial stage. It cannot meet the needs of rapid, comprehensive and accurate perception of power failure events in the substation area. So the analysis ability of big data is introduced for deep perception.

3. The formation of deep perception ability of power failure of substation area

To form the deep perception ability of power failure of substation area in smart grid, it is necessary to enhance the basic perception ability of substation area in smart grid and strengthen the perception information processing ability of substation area in smart grid. This paper presents a comprehensive solution including hardware equipment selection and software technology application, and carries out pilot tests at the level of substation area to verify the feasibility of this scheme.

3.1. Improve the basic perception ability of substation area in smart grid

In order to improve the basic perception ability of smart grid, it is necessary to build perception components, terminal equipment, communication network and perception strategy as a whole. For this reason, the perception layer of the substation area has been upgraded and transformed. A new generation of smart terminals are used in the pilot station area to improve the acquisition frequency; the use of high-speed power line carrier communication enhances the anti-electromagnetic interference capability, which improves the basic perception ability in an all-round way.

(1) In the pilot test, the terminal of the substation strengthens the perception frequency of voltage and current data representing the quality of power supply, and realizes the high frequency monitoring of 24 to 96 points of the whole network equipment every day.

(2) In the pilot test, the selected intelligent terminal sensing module has multi-power supply capability. It can work continuously for a period of time after power failure in the substation area, realize the sense of power failure, and complete the report of power failure events.

(3) The local communication network has been reformed in the pilot substation area. The high-speed power line carrier communication equipment has been used to replace the original narrow-band power line carrier, RS485 communication, special frequency wireless communication
and NB-IoT narrowband Internet of Things communication equipment. The application of high-speed power line carrier technology has greatly improved the network communication capacity of the substation area.

3.2. Enhance the perceptive information processing ability of power failure of substation area in smart grid

In order to strengthen the information processing ability of power failure of substation area in smart grid, statistics and probability theory are introduced to make full use of the data characteristics of power failure of substation area to study and judge the power failure. According to the judgment rules of power failure experts, the characteristics of terminal power failure events, meter power failure events, file data, equipment voltage and current data are selected and extracted, and the feature library of outage is constructed, and the diagnosis model of power failure of substation area is determined according to the data characteristics[3]. Training and testing the model of power failure of substation area, and combining with the actual power failure events in the field, continuously optimize and improve the diagnosis model of power failure of substation area, and realize the reinforcement learning of the model. The flow chart of the treatment scheme is as follows:

![Flowchart](image)

Figure 2. Power failure model construction diagram of substation area.

This paper mainly carries on the data analysis to the power failure of substation area, extracts the characteristics to construct the power failure model, which is divided into four parts:(1) extracting the data based on the power failure process and simulating the power failure data manually, making sample enhancement;(2) simplifying the sample data set, selecting and extracting the main characteristic of power failure model to construct the feature library of power failure of substation area;(3) based on the feature library of power failure of substation area, the XGBoost ensemble algorithm is used to construct the diagnosis model of power failure of substation area;(4) training the diagnosis model of power failure of substation area; if the accuracy of the model is not high, the feature library can be reconstructed;(5) re-selecting the data of power failure and verifying the judgment results on the spot.
3.2.1. Generating sample data set of power failure of substation area. Analyze and collect reverse data on terminal power failure events, meter power failure events, voltage and current data, file data and other data, and establish sample data set of power failure of substation area[4]. It is mainly divided into two parts : (1) filter and screen power failure sample data based on power failure process of substation area; (2) collect the power failure data based on the power failure process of substation area for reverse analysis, and enhance the sample set.

(1) Filtering and screening sample data of power failure of substation area
Firstly, according to the power failure process of substation area, the indicators influencing the power failure process of substation area are selected from the power acquisition system. Through the judgment of terminal and meter power failure, we can determine the power failure process of substation area. The terminal voltage is correlated by the terminal power failure event. If there is no return value of voltage, or the value is 0, or the voltage value falls between 0 and 132 during the time period when power failure occurs in the terminal, it is considered a terminal power failure event is incurred. The following figure shows the voltage distribution after the terminal power failure.

![Voltage diagram of power failure terminal](image)

The meter's judgment is similar to that of the terminal. It can judge whether the equipment is out of power by reporting events and voltages.

(2) Collect data of power failure of substation area for reverse analysis
It is also common to collect planned power failure data or conduct manual-drive power failure testing to enhance the sample set of power failure of substation area. Select suitable substations for power failure simulation, and collect a batch of manual-drive power failure data to increase the sample set of power failure of substation area.

3.2.2. Constructing feature library of power failure of substation area. Based on the sample data set of power failure of substation area in power consumption information acquisition system, the importance of characteristics to classification results is measured by characteristic correlation, and the characteristics of power failure are selected to construct the feature library of power failure of substation area.

According to the correlation theory and the actual scenes of power failure of substation area, this power failure model belongs to the classification problem. Given the sample set of power failure $X = \{x_1, x_2, \ldots, x_i, \ldots, x_n\}$, $i = (1, 2, \ldots, n)$, where $x_i = \{x_{i1}, x_{i2}, \ldots, x_{ij}, \ldots, x_{ik}\}$, $j = (1, 2, \ldots, k)$, and $j$ represents the characteristic number of the sample set[5]. Category set $Y = \{y_1, y_2\}$, where $y_1$ is the category of power failure of substation area, $y_2$ is the category of the non-power failure of substation area.
Using the separability criterion to distinguish the separability between classes of category set $Y = \{y_1, y_2\}$ of power failure of substation area, and using KL divergence to measure the importance of characteristics to power failure of substation area, the likelihood of the category of power failure of substation area is $p(X|y_1)$, the likelihood of the category of non-power failure of substation area is $p(X|y_2)$, and its logarithmic likelihood ratio is as follows[6]:

$$l(x) = \ln \frac{p(X|y_1)}{p(X|y_2)}$$  \hspace{1cm} (1)

Divergence $J_p$ is defined to judge the separability of categories. The larger the divergence, the greater the separability[7].

$$J_p = I_{12} + I_{21} = \int [p(X|y_1) - p(X|y_2)] \ln \frac{p(X|y_1)}{p(X|y_2)} \, dx$$  \hspace{1cm} (2)

KL divergence is used to select and extract the characteristics of power failure of substation area, and a feature library of power failure of substation area is established to provide data support for the training of diagnosis model of power failure.

3.2.3. Determining the diagnosis model of power failure of substation area. The supervised classification algorithm can be used to study and judge the power failure events after building the feature library of power failure of substation area. With the deepening of business scenarios, simple classification algorithm has been unable to effectively research, judge and predict power failure of substation area. XGBoost can integrate parallel processing of multiple decision tree models, introduce regularization while optimizing loss function, and reduce generalization error[8]. The specific process is as follows:

Step 1: Based on the feature library of power failure of substation area, data are randomly selected in equal proportion to form training set and test set data.

Step 2: Define the training data set of power failure of substation area, and the weight value of the characteristic is $\theta = \{\theta_1, \theta_2, ..., \theta_i, ..., \theta_n\}$, then the objective function is $y_i = \varphi(X, \theta) = \sum_{i=1}^{k} f_i(X, \theta)$. Select XGBoost model to learn and find the optimal solution iteratively.

Step 3: Increase the weight of the wrong classification, reduce the weight of the correct classification, optimize the decision tree in the iteration process, and use pruning method to prevent the model from over-fitting when constructing the decision tree.

Step 4: Dynamically adjust the XGBoost’s hyperparameters, and set the root mean square error as the evaluation index of the evaluation of objective function according to the characteristics of the decision tree model.

Step 5: Based on the objective function, a diagnosis model of power failure of substation area is constructed to train the data of power failure. The model with higher probability is selected as the final model, and a test set is added to the final model to determine whether the diagnosis of power failure of substation area conforms to the model. If not, step 1 is returned to reconstruct the feature library of power failure of substation area. If satisfied, proceed to step 6.

Step 6: Add new test sets, conduct on-site diagnosis verification, continuously expand the sample set, adjust parameters, and optimize the model.

3.2.4. Training of diagnosis model of power failure of substation area. In order to verify the accuracy and validity of the model of power failure of substation area, the data of the information acquisition system of Liaoning province in the past six months are analyzed and verified. as shown in the following table:
### Table 1. Power failure data sheet of substation area.

| Data Sheet | Related Data Field |
|------------|--------------------|
| E_ERC14    | EVENT_ID,POWER_OFF_TIME,POWER_ON_TIME,VALID_FLAG,ORG_NO,TERMINAL_ID,TERMINAL_ADDR,DATA_DATE |
| E_METER_EVENT_NO_POWER | METER_ID,ORG_NO,EVENT_TYPE,NO_POWER_SD,NO_POWER_ED,TERMINAL_ID,METER_ADDR,CONS_NO,DATA_DATE |
| E_MP_VOL_CURVE | VALUES,DATA_DATE,ORG_NO,PHASE_FLAG,TERMINAL_ID,PN,CONS_NO,CONS_TYPE,U1-U96_VOLTAGE |
| E_MP_CUR_CURVE | VALUES,DATA_DATE,ORG_NO,PHASE_FLAG,TERMINAL_ID,PN,CONS_NO,CONS_TYPE,I1-I96_CURRENT |
| C_CONS | CONS_NO,CONS_SORT_CODE,ELEC_ADDR,TRADE_CODE,ELEC_TYPE_CODE,STATUS_CODE,RRIO_CODE |
| R_MP | TERMINAL_ID,RMP_ID,COLL_TYPE,LINE_NO |

The importance of the characteristics to the classification results is measured by the characteristic correlation, and the characteristics of the power failure of substation area are selected to build the feature library of power failure of substation area. In order to illustrate the feasibility of the theoretical method, based on the data of the feature library, the method of combining real data with simulated data is selected to train the diagnosis model of power failure of substation area. The specific scheme is as follows:

- **Scheme 1:** Select 50% of the real and 50% of the simulated power failure data to form the sample set of the model to train the model.
- **Scheme 2:** Select 60% of the real and 40% of the simulated power failure data to form the sample set of the model to train the model.
- **Scheme 3:** Select 70% of the real and 30% of the simulated power failure data to form the sample set of the model to train the model.

The final model training results are shown in the following table:

| Options | Data Rate | Training Time | Training Result |
|---------|-----------|---------------|-----------------|
| Option 1 | 50% Real Data, 50% Simulated Data | 1h34min | Precision 74.11%, Recall 82.50%, F1-Score 78.08% |
| Option 2 | 60% Real Data, 40% Simulated Data | 1h38min | Precision 78.97%, Recall 85.63%, F1-Score 82.16% |
| Option 3 | 70% Real Data, 30% Simulated Data | 1h29min | Precision 92.54%, Recall 95.28%, F1-Score 93.90% |

As can be seen from the table above, only the composition of data has been changed in the three schemes. However, the model has different performance in the new test set. Generally speaking, the model in scheme 3 has better adaptability and is more in line with the expectation of power failure of substation area.

3.2.5. **On-site verification and reinforcement learning.** Re-select a batch of field data of power failure of substation area and observe the generalization ability of the model. The data of power failure in 20 substation areas are verified on the spot. The accuracy of the model reaches 96.74%. It can be used to study and judge the power failure in time.

4. **Conclusion**
Perception is the basis of power Internet of Things. Perception layer provides necessary data for the intellectualization of power grid, improves the perception ability of smart grid, strengthens the ability of perception information processing, and makes smart grid more intellectualized, which is the purpose of this paper. In this paper, the transformation of basic equipment of substation area in smart grid is proposed, and the perception data of smart grid are analyzed by big data method. A deep
perception scheme of power failure of substation area in smart grid is formed. The feasibility of the scheme is verified by the pilot test of power failure of substation area in smart grid. Accurate and rapid judgment and research on power failure of substation area can be taken as the basic condition to provide strong support for user-transformer relations, line losses calculation, power failure maintenance, disaster prevention and other business.

Acknowledgment
Science and Technology Project Support of State Grid co., LTD. No. SGLNJL00ZHJS1900017

References
[1] Dong Pan,Gang Mei. (2017)The analysis of fault and solution of electricity information system acquisition terminal. Telecom Power Technologies., 34(05): 180-181.
[2] Tao Lu,Jie Liu,Zhangquan Chou,ZhiGang Liu,Qian Zhang. (2019) Judgment of repeated power failure in the court based on big data technology. Power & Energy, 40(01): 46-48.
[3] Yuliang Shi,Yiping Song,Weiyi Zhu. (2018) The identification method of stealing electricity based on analysis of electricity characteristics.Journal of Computer Research and Development., 55(08): 1599-1608.
[4] ZhiMin Guo,Yonghao Zhang,Yinghua Zhou,Juan Su,Bo Wu,Juncheng Geng,Jie Ning. (2018) Failure analysis of distribution network based on multi-source data fusion strategy. Power System and Clean Energy., 34(01): 84-88.
[5] Luwa Hu. (2018) The Quest for Machine Learning. In: Geyue Zhu, People Post Press. PTPRSS.,Beijing.
[6] Ethen Alphaydin. (2014) Introduction to Machine Learning. Tsinghua University Press.,Beijing, pp. 62-84.
[7] Xuegong Zhang. (2010)Pattern Recognition. Tsinghua University Press.,Beijing, pp. 178-203.
[8] Tianqi Chen and Carlos Guestrin. (2006) XGBoost: A Scalable Tree Boosting System. https://arxiv.org/pdf/1603.02754.