

# Research and Application of Smart Operation Supporting System Based on Centralized Operation and Maintenance

Bai Yang<sup>1</sup> & Lang Yan-sheng<sup>1</sup> & Zhang Yin<sup>1</sup> & Xing Ying<sup>1</sup> & Li Qiang<sup>1</sup>

Zhu Cheng-zhi<sup>2</sup> Dou Cheng-long<sup>3</sup> & Han Feng<sup>3</sup>

<sup>1</sup>China Electric Power Research Institute, Beijing, China

<sup>2</sup>State Grid Zhejiang Electric Power Company, Hangzhou, China

<sup>3</sup>China Electric Power Research Institute, Beijing, China

**KEYWORD:** Smart Grid; Dispatching automation system; Centralized operation and maintenance; Data mining

**ABSTRACT:** The development and application of the SG-OSS (the Smart Grid Operation Supporting System) puts forward higher requirements towards the system operation and maintenance, and the operation model of SG-OSS has changed from traditional decentralized operation and maintenance mode to the centralized operation and maintenance mode. Focusing on the problem of urgent needs to improve the data mining analysis capability for the SG-OSS centralized operation and maintenance mode, this paper will conduct the study in the modeling and application of data mining technology in the SG-OSS centralized operation and maintenance mode, establish the data mining analysis model that is suitable for the centralized operation and maintenance mode, determine the function structure and scope of data mining, and apply this in plenty aspects in SG-OSS like comprehensive analysis of power grid faults, power grid trend warning, SG-OSS operation status and so on. Data mining technology in centralized operation and maintenance has achieved remarkable effects in the real application, which effectively enhance the reliability and satiability of SG-OSS and have strong practical value.

## Introduction

In accordance with the working idea of the State Grid, which is to “promote the overall construction of three centralizing and five large, and speed up the development of the unified strong smart grid”, enhancing the diagnosis and maintenance ability of the technical support system is the basic requirement of intelligent dispatching construction. The smart grid operation supporting system (SG-OSS) has been in operation at all levels, resulting in the new situations like the business has been processed centrally, application data show rapid growth, system laterally coupled closely, and a large number of new technologies have been developed and applied. The rapid development of SG-OSS puts forward new demands on the comprehensiveness and timeliness of the operation and maintenance [1-2], and the original mode of operation cannot meet the new demands, so the SG-OSS has changed from the decentralized operation mode to the centralized operation model. Centralized operation mode is to get the SG-OSS of provincial level and above under the State Grid Corporation to access to the centralized operation center of SG-OSS, and the operation center will provide comprehensive technical operation service, including centralized operation and maintenance, centralized monitoring, analysis and so on. Along with the development of SG-OSS in centralized operation, the demand of centralized information analysis become more and more strong, which puts forward higher requirements toward comprehensive information analysis and data mining ability.

The domestic and foreign research institutions have studied the operation, maintenance and data analysis on the equipment or system from substation, computer system monitoring center and so on.[3] offers a data analysis solution based on algorithm of Aprioris after the fault happening in SCADA system;[4] studies the application of data mining in power system;[5] introduced substation remote monitoring and data analysis under unattended situation. [6] introduced remote centralized data analysis and alarm scheme of the electronic equipment in smart grid;[7] conducted specific study on the overall frame work model, logic structure model, function structure model, information model, system network and system software structure of the decision support

system in the power enterprise based on the warehouse and data mining technology; Until now, there is no study on the data mining technology in the SG-OSS centralized operation and maintenance model.

In this paper, a data mining analysis solution is proposed for the centralized operation and maintenance of the SG-OSS. It analyzes the operation condition and changing status of the power grid information and SG-OSS through a variety of data mining methods, so as to realize the data mining analysis on SG-OSS operation condition, failure information, power grid data and so on, which provides data support towards the system alarm, trend analysis, software defect and system optimization, and improve the stability of system operation and automation level.

## Data mining features and scope for centralized operation and maintenance

### Data mining features for centralized operation and maintenance

The data analysis technology of centralized operation and maintenance conducts data mining analysis towards system operation condition (network status, disk status, CPU and memory workload, system key process, environment and equipment running status, etc.), the power grid running state (power system, frequency, the calculation amount of total monitoring, etc.), and running status of the safety protection system, which will analyze the stability of SG-OSS through system resources, availability, information security, network status and abnormal status; it will analyze the co-relations of fault information through whole network fault information; Through the analysis of multiple alarm sources of the fault (including network anomaly, full workload of server and etc.) of SG-OSS, and weighting the alarm source, it can diagnose the location of the fault system and determine the alarm location and fault reasons; combining the changing trend of the power grid information in nearby factory when grid fault occurs, it can determine the influence scope of the fault, give signs when fault shows, and can warn and control the risks that may exist in the grid; it will dig the potential risks of mining on the SG-OSS through the analysis towards common problems of the whole network. The function can be seen in Figure 1.

### Data mining scope for centralized operation and maintenance

The object of data analysis for centralized operation and maintenance is divided into three categories according to information collection types of SG-OSS: information analysis of power grid operation, SG-OSS resource information analysis and comprehensive fault analysis

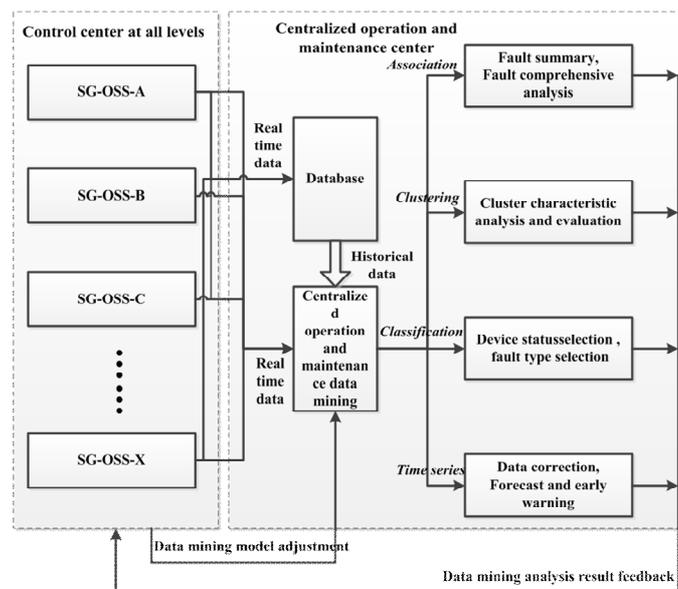


Figure 1 Centralized operation and maintenance data analysis function structure

The first category is the power grid operation information analysis, which is a summary analysis of the operation of the power grid:

1) The analysis will mainly measure the data over-limit, data changing, and the causes and trend of abnormal volatilities including the abnormal data such as frequency, the total data, the deviation from the plan data, the stable section and so on.

2) Analyzing the important statistical assessment index data, including the qualification rate of status estimation, control performance standard (CPS) and etc.

3) Analyzing the abnormal status of the network topology which mainly contains the topological state of the important power grid.

The second category is the resource information analysis of SG-OSS, mainly analyzing the system resource usage of SG-OSS:

1) Analyzing the resource occupancy rate of SG-OSS, including the CPU of server and workstation, workload, memory usage, disk partition usage, database performance, resource growth rate, etc.

2) Analyzing the stability of communication channels, including the running status of the dedicated line (through the terminal services or the serial link) and the quality of the link data.

3) Analyzing the running status of the application and the process, including applications, services and the important process status, such as the status of automatic voltage control (AVC) / automatic generation control (AGC).

The third category is the comprehensive analysis of the fault information. Aiming the occurring time of the fault, it analyzes the effect area and trend of the fault according to the equipment information status that is nearby the fault.

1) Analyzing the power grid fault information comprehensively. Before and after the fault of the power grid happens, it analyzes the relations of different kinds of equipment information within the fault influence area and provide data analysis for the power grid fault early warning and controlling through the changing status of the equipment information in the factory station that is nearby the power grid fault location.

2) When the SG-OSS fails, it analyzes the system's internal resources and changing trend of performance, and provide data support in checking the fault of SG-OSS and finding hidden trouble.

## Centralized operation and maintenance data mining technology research

According to the data characteristics of power grid information, fault information and system operation information that get from centralized operation and maintenance center, centralized operation and maintenance data mining uses classification analysis, clustering analysis, correlation analysis, time series analysis to conduct multi angle analysis, and get more accurate data analysis by adjusting the object and target in the analysis model according to the actual needs.

### Classification analysis

Bayesian Model Naive (NBM) is used to analyze the operating information of the SG-OSS. According to the simple Bias classification model:

$$V_{map} = \arg \max P(V_j | a_1, a_2 \dots a_n) \quad (1)$$

Where  $V_j$  belongs to  $V$  collection,  $V_{map}$  is the most likely value for a given example,  $a_1, a_2 \dots a_n$  is the property of this example.

Centralized operation and maintenance of data mining uses historical data of SG-OSS operation information to train this mode. Through the statistics on the probability of all kinds resources in each interval, it does analysis on the current system situation, gives assessment on the stability of the current running system, gives early warning towards the possible risks, and calculates the probability of risk that may arise in the future of power grid, which effectively improve the stability of power system. It establishes the modeling on the history abnormal faults of the existing SG-OSS, and classifies the abnormality in the current situation, which can detect the abnormal fault in time.

### Time series analysis

Centralized operation and maintenance of data mining uses ARMA model in the time series to establish the model by using the historical data of power grid and SG-OSS resources. Considering the data serial that gets from the collect information and forms with the time passage into a set of random serial, this set of random variables with dependency rations show the continuity of original

data in time. On the one hand, it is affected by the influence factors, on the other hand, has its own variation rule, assuming the influence factors were  $x_1, x_2, \dots, x_K$ , and by regression analysis;

$$Y_t = b_1 x_1 + b_2 x_2 + \dots + b_p x_p + Z \quad (2)$$

Where  $Y$  is the observation data of the collection information,  $Z$  is the error. As the prediction object  $Y_t$  is influenced by the change of itself, the rule can be shown in the following:

$$Y_t = b_1 Y_{t-1} + b_2 Y_{t-2} + \dots + b_p Y_{t-p} + Z_t \quad (3)$$

The error has dependency relations in the different time, and the error is expressed by the following formula,

$$Z_t = e_t + a_1 e_{t-1} + a_2 e_{t-2} + \dots + a_q e_{t-q} \quad (4)$$

Thus, the ARMA model of the collection information is:

$$Y_t = b_0 + b_1 Y_{t-1} + b_2 Y_{t-2} + \dots + b_p Y_{t-p} + e_t + a_1 e_{t-1} + a_2 e_{t-2} + \dots + a_q e_{t-q} \quad (5)$$

Centralized operation and maintenance data mining uses the formula (4) to emend the information data of the SG-OSS resources so as to reduce the information collection error and predict the future trend of information.

### Clustering analysis

The centralized operation and maintenance data mining technology uses cluster analysis to dig the running situation of the power grid and system resources to find the clustering center to cluster them into groups, and does in-depth analysis towards the characteristics of the clustered group. Clustering analysis uses the K-meas algorithm, after initializing the data, uses the expression matrix as a set of objects  $x$ , inputs the specified clustered number  $N$ , and selects  $N$  objects from  $X$  randomly to be the initial clustering center to do the iteration, assigns the data objects to the nearest cluster centers according to similarity, updates clustering center according to the results of evaluation formula, and then takes the average vector of each class as the new clustering center to assign the data objects until all clustering is completed. The evaluation formula is as the follows:

$$J(c, m) = \sum_{i=1}^m \|X^i - m_c^i\|^2 \quad (6)$$

Where  $X$  is the data point,  $M$  is the number of data sources,  $C$  is the number of cluster centers,  $\mu_c$  is the cluster center.

We complete the clustering analysis when  $J$  (the sum of square of the difference between the data of each class and each cluster center) comes to minimum.

### Association analysis

Centralized operation and maintenance data mining uses association analysis to analyze the relation of the cascading failure of power grid and SG-OSS system multi-stage operation fault, and uses Apriori algorithm to do the correlation analysis on the fault information. The first is to find all the all the frequency set and then generate strong association rules from the frequency set so as to generate the desired rules, and retains the rule that is greater than the minimum reliability. It will analyze the correlation degree between the fault information, and between the fault information and collecting information through this model, and it will also analyze the correlations between each faults, and between the collecting information and fault information, calculate the correlation factor between the data, and show the internal logic relationship of each data directly through the adjust of the correlation rules and feedback from the data validation on the real-time data.

## Application of data mining in centralized operation and maintenance

The data mining technology has been officially put into use in the centralized operation and maintenance center.

Through the mining analysis towards the power grid data of national SG-OSS and project information, and combining with the actual needs on the operation to classify the warning information using association analysis, the daily warning has been reduced from more than 20000 to 150, and when the technical support system and power grid fault occurs, the time of dealing with the faulty by the operation and maintenance people has been shortened from several hours to half an hour, which greatly improve the efficiency of fault handling; a great many abnormalities situations that caused by not discovering in time was avoided through the time serials analysis on the usage situation of SG-OSS resources and doing the forecast towards the system resource usage and the abnormal frequency was reduced from 200 times to 10 times; it realizes the key monitoring on the equipment group that affect the system usage after the equipment and application were clustered, which increased the available rate from 95% to 99.9% through the clustering analysis on the system equipment and operation status of the application; it strengthens the management towards the link that has been screened to be unstable and increases the stability of the monthly link from 97% to 99% by using the Bias classification model to screen the communication link. After nearly one year operation, the data mining technology in SG-OSS centralized operation and maintenance has significantly improved the stability and reliability of the system, and greatly improved the automation level of the controlling center.

Table 1 data mining technology application effect on the centralized operation and maintenance of smart grid

	Before	after
Daily Warning Number	20000	150
Average Time of Dealing with the Warning (Hour)	2	0.5
Abnormal Number of Monthly System Resources	200	10
Monthly Resources Availabilities	95%	99.9%
Monthly Link Stability	97%	99%

## Conclusion

Data mining analysis technology is proposed in this paper, which has the practical uses in fault synthetic analysis, early warning trend analysis and statistical analysis of the SG-OSS status according to the requirement of SG-OSS centralized operation and maintenance. Through the clustering analysis, correlation analysis and other aspects of analysis towards the power grid data, system resources information, fault information, this data mining analysis realizes the centralized statistics and analysis on all power grid information of the SG-OSS and the project information. It provides horizontal comparison and vertical analysis towards the operation condition of all SG-OSS, which effectively improve the data analysis level of the centralized operation and maintenance, and the operation ability of the SG-OSS, reduces the maintenance cost for the SG-OSS operation, and provide data and technical support for the following centralized operation and maintenance, system upgrades, defect elimination, selection, early warning and preventive control of SG-OSS.

## Acknowledgements

This work was financially supported by by the Science and Technology Project of State Grid (Project No. DZ71-14-046).

## REFERENCES

- [1] MA Tao-tao, GUO Chuang-xin, CAO Yi-jia, et al. 2010. Current status and development trend of intelligent dispatching automation system of power system[J]. *Automation of Electric Power Systems* 34(9): 7-11.

- [2] WANG Ji-feng, SHEN Guo-rong. 2012. Some key technical issues on intelligent power dispatching of bulk power grid[J]. *Automation of Electric Power Systems* 36(1): 10-16.
- [3] HUANG Xiao-hong, CHEN Li-hua, WANG Qian. 2008. Data analysis after fault based on improved association rules Apriori in SCADA system[J]. *Journal of North China Electric Power University* (4):27-32.
- [4] YU Zhi-hong, GUO Zhi-zhong. 2001. DATA MINING AND POWER SYSTEM [J]. *Power System Technology* 25(8): 58-62.
- [5] ZHAO Jia-qing. 2003. Computer centralized SCADA system on unmanned 220 kV substation[J]. *Automation of Electric Power Systems* 27(8): 78-80.
- [6] MA Hao, HAN Si-liang. 2005. Design of remote monitoring and fault diagnosis system for power electronic equipment[J]. *Automation of Electric Power System* 29(2): 50-55.
- [7] ZHANG Lei, LI Xiao-ping, ZHANG Xiao-hua. 2005. Design of Decision Supporting System Based on Data Warehouse and Data Mining in Power Enterprises[J].*Electric Power Science and Engineering* (4):75-77.