Information System Operation and Maintenance Analysis Method Based on Five-dimensional Model

Xusheng Liu¹, Wei Zhao¹, Zhimin Li¹, Qing Zhu¹ and Ziqian Li¹,
Runhao Qiang²*  
¹State Grid Customer Service Center.
²Nanjing Data Foundation Technology Co, Ltd.
*Corresponding author e-mail: wanghe6363@126.com

Abstract. Considering the complexity of the current power industry information system and the severity of the existing data islands, a five-dimensional analysis method for business information system is proposed. By dividing the data collected by the system into different dimensions and then combining them for unified analysis, the index tolerance interval model is constructed to find hidden dangers and anomalies in the information system and assist the operation and maintenance personnel in root cause positioning and impact range assessment.

1. Introduction
With the increasing scale and complexity of information systems in the power industry and the application of artificial intelligence and machine learning technology, the focus of operation and maintenance work is also to focus on a single operation and maintenance object or operation index to develop in the direction of system overall state evaluation and multi-index joint analysis and prediction. However, most of the current popular operation and maintenance tools only focus on one direction of operation and maintenance work and lack comprehensive and complete overall evaluation and analysis capabilities. The reasons can be summarized into the following two points: First, the data is missing, and there are many kinds of operation and maintenance objects, from slave ring equipment, host equipment, network equipment, storage equipment to various business system software and tool software, so the data cannot be effectively and uniformly collected. Second, the indicators are isolated and lack the ability to carry out joint analysis of different types of indicators, which cannot form a unified analysis model.

2. Overall Conception of Five-dimensional Model
Build a unified information system analysis model. In order to ensure the effectiveness and integrity of the system model, all relevant indicators from hardware to software should be considered when building the model. The specific indicator classification is shown in the figure below.
On the basis of this data classification, we abstracted five dimensions to build a system analysis model, as shown in the following figure. These five dimensions are interrelated and interact with each other, closely linking the hardware, software and functions of the system and mapping out the overall situation of the system.

Figure 1. Five-dimensional model data classification.

(1) Infrastructure layer
The infrastructure layer corresponds to the physical equipment and dynamic environment of the machine room, and the main indexes include equipment ledger, machine room temperature, equipment temperature, abnormal alarm of equipment, operation time of equipment, host power, cpu fan speed,
machine room water leakage, air conditioning, smoke and other indexes. It mainly reflects whether the system hardware state is normal from the external physical level.

(2) Physical link layer

The physical link layer corresponds to the actual connection relationship of IT equipment in the computer room, equipment operation indexes and equipment operation logs. The connection relation mainly includes the connection relation between the host and the switch, the connection relation between the host and the load balancing, the connection relation between the firewall and the switch, the connection relation between the security equipment and the switch, the connection relation between the storage and the switch, etc. Indicators mainly include CPU, memory, IO utilization rate, network interface delay packet loss rate, storage utilization rate and throughput, etc. The logs mainly correspond to the operation logs of the host, network, storage and security equipment. The connection relationship forms the network topology between the equipment in the computer room, and the indexes reflect the running state of hardware equipment and the utilization rate of resources. At the same time, the log records the abnormal and internal error information of the equipment. Through this layer, full monitoring of the operating state of system equipment can be realized, and alarms can be issued for equipment with abnormal operating state.

(3) Logical topology layer

The logical topology layer realizes the recording of software connection relation, software operation index and log. The software connection relation mainly includes the connection relation between middleware of business system, the connection relation between load balancing and application, the connection relation between application system and database, and the connection relation between software cluster, database cluster or RAC(oracle). The indicators are mainly application system work order quantity, business volume, concurrent reading and writing of database connection, middleware queue and other indicators. Logs mainly include logs of industry middleware and databases themselves, as well as logs of operations printed by users. The logical topology layer mainly records the running status and current traffic volume of the software.

Due to the large amount of data contained in the service topology layer and the complex relationship between each module software and middleware, the service topology layer can be logically divided into different slices according to clusters or functional services. Each slice contains only a single device and software related to system functions. Slices have the characteristics of high cohesion and low coupling, and can be seen as an independent system when analyzing, thus reducing the amount of data and complexity of operations.

(4) Transaction analysis layer

The transaction analysis layer mainly records the call relationship between systems and the Apdex index, which is different from the relationship between software in the logical topology layer determined during the construction period. When the system is running, the call relationship between different modules and services is more complex, which is difficult to be counted by manual sorting. Therefore, the transaction analysis layer mainly records the request call relationship in the software when the system is running. It mainly includes the origin information, destination information and the time of the whole request, as well as the call stack information of the code and the execution information of SQL. This layer of data can only be retrieved by tools such as APM when the system is running. Each retrieved call information is called a transaction, and the name of the transaction is the URL path from which the call originated. The Apdex index of the system can be calculated to reflect the availability of the system by counting the usage time of multiple calls and the usage tolerance of users.

(5) Business process layer

The business process layer mainly records the function items corresponding to the system business and the dependency relationship between business processes. Each function corresponds to a menu button of the system, which in turn corresponds to the URL path of the background accessed by the button.
3. Correlation of Five-dimensional Models
The five-dimensional model divides our operation and maintenance work from the lowest level of physical hardware equipment to the highest level of functional menus into five levels, and the capabilities of each level step by step, finally building the overall model of the system. There is a specific connection between each layer, and the infrastructure layer is associated with the physical link layer by device code, MAC address or IP. The physical link layer is associated with the logical topology layer through IP or user-defined slice ID; The logical topology layer is associated with the transaction analysis layer through IP. The transaction analysis layer is associated with the business process layer by the transaction name. All levels of the five-dimensional analysis model are linked through these correlation relationships. According to the progressive relationship between layers, down can be specific to specific operation and maintenance objects and operation indicators, can be used for root cause analysis and prediction of faults, up can be traced back to specific functions and processes, and can be used for range impact judgment of faults.

4. Analysis Method of Five-dimensional Model
The five-dimensional model makes independent analysis of the data contained in each layer, and at the same time combines the data of each layer to make analysis. The data in each level can be divided into baseline index analysis and expert model analysis. Baseline index analysis is to set a fixed or dynamic threshold range for indexes collected in each layer. When the index data falls outside the tolerance intervals, an alarm action is triggered. Expert model analysis is generally based on multi-index joint construction of slices divided in logical topology layer. It can be analyzed and calculated based on weighted statistics or machine learning. When the statistical score or machine learning calculation result falls outside the tolerance intervals, an alarm will be triggered.

When calculating the tolerance interval, we describe the user's feelings through three intervals: satisfaction, tolerance, and irritability. We sample the indicators or scores, then classify the sampled values, divide them according to the user's feelings, and divide them into three categories: satisfaction, tolerance, and irritability. Tolerance Index = (1 × satisfactory sample + 0.5 × tolerance sample) ÷ total number of samples. In this way, our sampling result is quantified into a value range of 0 ~ 1, and the current availability of the system is comprehensively evaluated according to the range value. When the system availability is low, the system issues an alarm, and then locates the cause or estimates the scope of the alarm by linking up or down different levels.

According to the calculation of tolerance score, we can find that the alarm can come from an object or the combination of multiple objects. When the alarm comes from an object, we can comprehensively analyze the indicators of the object in different levels. For example, check the object's temperature, fan speed, device status indicator and other options at the infrastructure layer to determine whether the problem occurs at the infrastructure layer. Check whether the network connectivity, port packet loss, cpu, IO, memory and other indexes of the device are normal at the physical link layer, and judge the root cause of the problem by combining the log information of the device. For the problem of a single device, the root cause of the fault can be analyzed through these two layers, and then upward analysis is to analyze the possible impact range of the fault, which can be directly associated to the transaction layer through IP and to the business process through transaction ID, thus counting the actual affected functions. For the joint problem of multiple objects, the selection of multiple objects is to select objects with direct or indirect connection in the logical topology layer, and generally it is to select objects of a slice. When it is found that the score of the multi-object model falls within the fretful area, the object index in the slice is analyzed. The analysis method is similar to that of single equipment. The analysis is made by finding the scoring bottleneck devices in the slice, but the difference between the analysis method and the single device is that the dependency of multiple devices in the slice is analyzed, such as load balancing, shared queue and other factors. This is a downward root cause analysis through indicators, logs, etc., and an upward impact range analysis.
5. Summary
In the five-dimensional analysis method, various operation and maintenance indexes originally mixed together in the system are divided according to different scopes and distributed in five different dimensions. Each dimension focuses on different issues, and the types of data collected are also different. By analyzing the data in the dimension, anomalies occurring in different levels can be analyzed more effectively. At the same time, the root causes of anomalies can be located downward, and the impact range of anomalies can be evaluated upward.

References
[1] Cai Liang, Li Fen. Application of APDEX-based Application Performance Analysis in Meteorological Services [J]. Agricultural Network Information, 2008 (10): 24-27 + 30.
[2] Yu Huixiong. Active operation and maintenance analysis of distribution network under big data analysis [J]. Telecommunication Power Technology, 2019, 36 (12): 119-120.
[3] Li Feng, Yang Lin. Analysis and Design of Operation and Maintenance Monitoring Management System [J]. Network Security Technology and Application, 2019 (05): 41-43.
[4] Zhao Wei. Analysis of the key points of the integrated construction of smart grid power distribution operation and maintenance [J]. Integrated Circuit Applications, 2019, 36 (04): 59-60.
[5] Gao Jingbin, Peng Ziqiao, Wang Gang, Dang Yongqiang, Kang Yixu. Application Analysis of Digital Intelligent Remote Control Operation and Maintenance System for Nuclear Power Plant [J]. Instrumentation Users, 2019, 26 (08): 51-53 + 58.