Research on Intelligent Operation and Maintenance Model of Large Database for Electric Power Industry

Chaoyi Yang¹, Ye Liu¹, Jianyi Xiao², Xiaobo Huang² and Jiahua Chen³,*

¹ Information Center, Guangdong Power Grid Co., Ltd., Guangdong, China
² Guangdong Electric Power Technology Development Co., Ltd, Guangdong, China
³ School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing, China

*Corresponding author email: 1871184489@qq.com

Abstract. Existing research on automated database operation and maintenance for the electrical industry mainly focuses on distributed and cloud platforms, and there is a lack of traditional large-scale database intelligent operation and maintenance research. This paper designs an overall operation and maintenance model framework of "intelligent perception-intelligent decision-intelligent execution", and proposes feasible implementation plans, including: (1) Use the prophet time series forecasting model to perceive and predict important database performance indicators, and dynamically adjust the threshold of each performance indicator according to the predicted value; (2) Perform correlation analysis on abnormal indicators through the association rule model to construct "Indicators" ->"Operation" optimized combination operation strategy library for intelligent decision making; (3) According to the intelligent decision library, automatically restrict the associated operations under abnormal conditions to ensure the normal operation of the service and realize intelligent execution.

Keywords: Zabbix; prophet; Intelligent operation and maintenance; Service monitoring; Association rules.

1. Introduction

Database operation and maintenance refer to a series of services such as software installation, configuration optimization, backup strategy selection and implementation, data recovery, data migration, troubleshooting, and preventive inspections for user databases. Jia Zhu¹ pointed out that the important basis for the reliability and cost-saving of the entire system is to use scientific methods to gather information, comprehensively and correctly evaluate equipment operating status and changing trends, and reasonably arrange operation and maintenance plans. With the digital transformation of the power grid, enterprises include both traditional steady-state services and new sensitive-state services, generating a large number of different types of data, and increasing the difficulty of database operation and maintenance management. Therefore, any fluctuations in the server will have a significant impact on the business environment, and the operation and maintenance system is increasingly being valued by major enterprises. According to Lucy's research at the "Eleventh China Database Technology Conference", the evolution of the database operation and maintenance system has gone through four stages:

Scripting stage: At this stage, service traffic and clusters are relatively small, so manual and scripting are sufficient to support the entire service.
Platformization stage: With the rapid development of the company’s business, only a few dozen database administrators cannot adapt to the company’s development. Therefore, opening and authorizing some operation permissions for database administrators can greatly reduce the daily workload of database administrators, but at this stage, the passive post-rescue method is still used to solve the problem. Failure to find and deal with the problem in time may cause great losses.

In the automation stage, it is based on the generation of a large number of automated scripts, and the automated scripts are used to monitor the system and resources. And carry out a series of set procedures for processing. Automated operation and maintenance can effectively reduce losses, but there are still many bad cases that cannot be taken care of. After the failure occurs, only fixed scripts can be executed, and corresponding real-time processing strategies cannot be generated.

In the intelligent phase, Gartner proposed the concept of intelligent operation and maintenance in 2016, which aims to use big data, machine learning and other methods to improve operation and maintenance capabilities. Humanized and fully automated. The emergence of intelligent operation and maintenance has effectively solved the shortcomings of traditional and automated operation and maintenance, allowing operation and maintenance personnel to transition from troubleshooting and then solving problems to predicting problems and stifling problems in the bud[3].

For the field of database intelligent operation and maintenance systems, related research is gradually developing, and the processing tasks are concentrated on the operation and maintenance system and key indicators positioning. In terms of operation and maintenance system research, Feng Zhanli[4] pointed out that many nodes in large systems will generate a large amount of discrete and related data. In the process of data analysis, time series databases can display these discrete data in a visual manner, making it easier to track faults. Li Ming[5] proposed a new automated operation and maintenance system based on a time series database, using Redis technology to improve the storage capacity of the database, and using pre-aggregation and retention strategies to improve the computing power of the system. Rong Xueqin[6] and others proposed to build a service Support the platform to ensure the smooth operation of the system. The front-end platform of the system client is built using HTML+bootstrap technology, and the operation and maintenance mode is in the form of Reset API to establish shared services in node.js and time series databases. Li Jun[7] proposed a network operation and maintenance platform based on a big data analysis system, using the MongoDB DB data set to classify and store data in different data sets, and collective nouns are unique. Based on deep learning technology, Zengrui[8] proposed an intelligent operation and maintenance system that can efficiently process data. This system uses the isolated forest algorithm to identify faults, and the model evaluation is completed using GAN to integrate abnormal data from various scenarios. The function of identifying faults and guiding maintenance.

In terms of key indicators positioning, Tian Yongming, Ma Qiang[9] and others used a normalized spectral clustering algorithm to analyze normal data and abnormal data from a multi-dimensional perspective, and depict historical data in the form of curves. Identify the characteristics of database failures and obtain important indicators parameters. Because the system data is unstructured and difficult to analyze and utilize, Xiaozhou[10] and others proposed to use convolutional neural methods to extract data features from it, and then build a decision tree to automatically classify faults. Danúbia[11] and others developed a visualization tool: users can receive information from the operation and maintenance system through related visualization equipment, and can visualize the information they want through voice commands and other methods. Aiming at the problem of strong subjectivity of classification methods, Tieqiang[12] and others proposed the use of deep neural networks (DNN) for indicators classification. On the basis of summarizing the running status, pre-processing the important indicators, and then introducing deep neural networks and deep learning historical data to complete more accurate and effective classification. Huang Xiaoming[13] proposed a method that can analyze the impact factors of operation and maintenance indicators. According to the impact factors, a tree model is established for each indicator, and then the degree of influence of each factor is calculated. In the process of operation and maintenance, the degree of influence can be Real-time monitoring of higher factors and immediate rectification. Wu Muyang[14] proposed a method of screening indicators by comparing the correlation between indicator data. The whole method divides the indicators into different categories according to the strength of the correlation, and then selects representative
indicators in each category to construct an indicator system, and processes the data according to the weight of the indicator. Establish an operation and maintenance evaluation model.

Through combing, analyzing and summarizing existing research, it can be known that many scholars have carried out tasks such as system design and key indicators classification for database operation and maintenance and have achieved many results. However, there are still the following shortcomings: in the research on database intelligent operation and maintenance services, most of the research is aimed at the research of distributed databases and cloud platforms, and the research on the intelligent operation and maintenance of traditional relational databases is very lacking. This paper proposes a large-scale database intelligent operation and maintenance model based on zabbix, designs the "intelligent perception->intelligent decision->intelligent execution" overall operation and maintenance model framework, and proposes a feasible implementation plan.

2. Overall Framework

2.1. Operation and Maintenance Framework

The intelligent operation and maintenance model is an important means to realize effective database management and monitoring, and dynamic resource allocation and optimization. In the process of database operation and maintenance, it is necessary to monitor the status and readiness of database resources in real time, provide service quality assurance according to the negotiated service level agreement, and then adjust the operation as needed to provide the best possible service. Through the effective and efficient operation of the database, it meets the needs of rapidly changing users, strives to ensure the realization of the concept of database support network capabilities, and enhances the intelligence and autonomy of operation and maintenance management. Through analysis and refinement of the above-mentioned intelligent operation and maintenance capabilities, as shown in figure 1, the database intelligent operation and maintenance model is designed and implemented from three aspects: intelligent perception, intelligent decision-making, and intelligent execution.

![Figure 1. Intelligent operation and maintenance framework.](image)

Intelligent perception covers three levels of perception, understanding and prediction, mainly including future prediction, anomaly detection, and pattern discovery. It will automatically summarize the operation and maintenance knowledge based on the platform, be reliable, measurable and controllable, and provide timely warnings and warnings so that information center personnel can Tracking and tracing, regularly reporting the analysis trend of database intelligent operation and maintenance, effectively supporting management decision-making, and maintaining the stability of database operation and maintenance.

Intelligent decision-making formulates strategies based on the current operating status information of the database, transforms the perceived multi-dimensional comprehensive situation information into decision-making information, formulates important decisions, mission plans and resource control plans, and carries out task-based automatic planning, perception data analysis, strategic decision-making, Independent collaborative decision-making, resource mapping decision-making, situation assessment and optimization strategy generation. Intelligent decision-making is the output of automatic perception. Intelligent execution must first build a strategy database. The data grid contains fault descriptions, indicators, operations, and statuses, and displays the attributes of abnormal indicators and compliance thresholds. The user can add, modify and delete the policies in the policy library. When the system is abnormal, it will search for and execute the strategy in the strategy database.

2.2. Steps of Implement

The entire intelligent operation and maintenance service process is divided into 9 steps: collecting KPI data, reading data, updating thresholds, reading recent data, judging whether to alarm, selecting
strategies and sending to execution, executing strategies, recording data, reading and displaying. The intelligent operation and maintenance execution architecture is shown in figure 2:

![Intelligent operation and maintenance execution architecture](image)

Figure 2. Intelligent operation and maintenance execution architecture.

- Read the values in the corresponding table of the corresponding performance indicators under different operations, record the variance ranking of the indicators, provide a basis for judgment in the later intelligent decision-making, and use the RSR algorithm to comprehensively rank them.
- Read the KPI data, and analyze the abnormal indicators of the business system in terms of trend, seasonality and holiday factors through the prophet model.
- Predict the known data, get the range of prediction in the future, and use the relevant interface to dynamically adjust the threshold of each KPI, so that the entire service will continue to improve and ensure stability.
- Read the latest data and judge whether the indicators value exceeds the threshold.
- If the value of an indicator exceeds the threshold, an alarm will be issued.
- Through the association model, the association rules between the indicators are obtained, and the optimized combination operation strategy library of "indicator->operation" is constructed, combined with abnormal indicators to realize the decision of which operations are restricted, and the decision is sent to the execution.
- The administrator can add, modify and delete strategies, and the system can select and execute strategies based on fault descriptions or abnormal indicators. The decision-making operation status can be viewed in the decision-making library.
- Send the result of the execution decision to the zabbix server database for the next operation.
- The data of the zabbix database is displayed to the user through the front-end interface, so that the maintenance personnel can easily understand the status of the entire operation and maintenance system.

3. Implementation Plan

3.1. Collect KPI Data
Step1: Deploy zabbix monitoring suite in the database system to monitor all performance indicators of the database;
Step2: Operate the database in four situations, including: 100% read operation; 100% write operation; 50% write 50% read operation; idle state. The operation interval is 1 time/second, and each operation is executed 10,000 times;
Step 3: Use zabbix to monitor the above operations separately, and read the values in the corresponding table of performance indicators. The indicator values are organized into the following form: each row represents the value of N indicators for one operation, and the column names include: timestamp, opt (operation), itemid (indicator ID), etc;

Step 4: Find the variance of each column of indicators and arrange them in descending order according to the size of the variance, that is, the indicators with large fluctuations are ranked higher;

Step 5: Obtain the ranking of each indicator under various operations, and use the RSR (Rank-sum ratio) sorting algorithm to obtain the final ranking of each indicator;

Step 6: Select the top M indicators as KPI (Key Performance Indicator); respectively record the ranking of KPIs under various operations to provide an inference basis for which operation is affected when abnormalities occur later.

3.2. Abnormal Prediction Based on Prophet Time Series Data

Simulate multiple forms of random operations on the database in the morning, afternoon, and evening by time, and record the time and the value of M KPIs. After obtaining a large amount of simulated data (simulating 1 month), use the time series model prophet to predict each time the normal range of the stage KPI.

The Prophet model is a model specially used to analyze time series data\cite{15}, as shown in figure 3. Its operation process is to obtain the optimal result through the iteration of modeling and prediction evaluation cycles.

![Figure 3. Prophet model running diagram.](image)

The Prophet model can be divided into a trend term function, a seasonal term function and a holiday factor function. \( \varepsilon(t) \) is an error term, which represents a special situation. The function expression is as follows:

\[
y(t) = g(t) + s(t) + h(t) + \varepsilon(t)
\]  

(1)

The trend function \( g(t) \) represents the value of aperiodic changes in the time series. The function expression is as follows:

\[
g(t) = \frac{C}{1 + \exp[-k(t - m)]}
\]  

(2)

\( C \) represents the model capacity, \( k \) represents the growth rate, and \( m \) represents the offset.

The seasonal function \( s(t) \) represents the value of the cyclical change of the time series. The function expression is as follows:

\[
s(t) = \sum_{n=-N}^{N} C_n \exp\left(\frac{j2\pi nt}{P}\right)
\]  

(3)

\( P \) represents the sequence period, \( C_n \) represents the coefficient parameter, \( C_n \sim N(0, \sigma) \).

The holiday factor function \( h(t) \) represents the influence of special factors such as holidays on the system data. The function expression is as follows:

\[
h(t) = \sum_{i=1}^{L} K_i \mathbb{1}(t \in DL)
\]
\[ Z(t) = [1(t \in D_1), \ldots, 1(t \in D_L)] \]  

\[ h(t) = Z(t)k \]

Di represents the date of holiday i. The \( Z(t) \) function judges whether \( t \) is during the holiday. If \( t \) is within the holiday, the value of \( Z(t) \) is 1, otherwise, it is 0. parameter \( k \sim N(0,v) \).

Through the prophet model to analyze and predict the abnormal data of the business system, realize the use of related interfaces to dynamically adjust the threshold of each KPI.

3.3. Association Rule Algorithm

First, perform a normal test or a stress test on the database according to a certain strategy to generate a training data set containing abnormal indicators values. Each fault is called an item, which contains a different number of abnormal indicators, and all the faults are combined into an item set. Then conduct association rule analysis on abnormal indicators. Assuming that A and B are two indicators, the association rule is an implicit formula that forms \( (A \rightarrow B) \). The association rule contains two subtasks: frequent indicators mining and association rule generation.

Frequent item mining: Select the frequently occurring indicators from the abnormal indicators training data set, and calculate the frequency of each failure, namely the support degree \( S(A \rightarrow B) \), the function is as follows:

\[ S(A \rightarrow B) = \frac{\text{The number of failures contain both A and B}}{\text{Total number of failures}} \]  

Select the min support threshold. When the indicator frequency exceeds the threshold, this indicator is a frequent indicator.

The specific operation is shown in figure 4:

- Scan the data set, count the number of occurrences of each indicator, form a candidate 1-indicators set \( X_1 \), count them, and select the frequent 1-indicators set \( Y_1 \) according to the threshold;
- Combine the indicators in each fault in pairs, use the hash function to map to the bucket, and count the count in the bucket. If the count is higher than the threshold, the bucket is a frequent bucket, and the corresponding bitmap is marked as 1, if the count is lower than the threshold, it is recorded as 0;
- Scan the data set for the second time, generate a 2-indicator combination based on the frequent 1-indicators set \( Y_1 \), combine each indicator in the combination in pairs, and use the hash function to get the bucket number. If the number corresponds to the bitmap value of 1, The candidate 2-indices are retained, and the candidate 2-item set \( X_2 \) is finally obtained, which is counted and the frequent 2-item set \( Y_2 \) is filtered according to the threshold;
- Repeat the above process continuously, and stop if the candidate indicators set is empty.

Generating association rules: According to the generated frequent indicator set, the confidence \( h(A \rightarrow B) \) is the probability that indicator A and indicator B appear at the same time based on the existence of indicator A. The function expression is as follows:
\[ S(A \rightarrow B) = \frac{\text{The number of failures contain both } A \text{ and } B}{\text{The number of failures contain } A} \]  

Finally, the corresponding management rules are generated by the confidence between the two indicators.

### 3.4. Intelligent Execution Based on Decision Library

According to the association rules between the indicators, combined with the KPI rankings of the four database operations in intelligent perception, the decision on which operations to restrict can be realized. Specific steps are as follows:

- When the system fails, check the abnormal indicators, and find the indicators associated with them in the association rules, and mark them as pseudo-abnormal indicators;
- If a problematic indicator ranks high in a certain operation, it means that the operation needs to be restricted. Therefore, according to the ranking of abnormal indicators and pseudo-abnormal indicators in various operations, it is possible to obtain those operations that need to be restricted. Then analyze the specific operations that need to be performed. For example, if the operations that need to be restricted are all writing operations and 80% write 20% read operations, the operation performed is to limit write operations to less than 80%.
- The priority of setting anomaly indicators is higher than that of pseudo-anomaly indicators, that is, when the operations corresponding to the anomaly indicators and pseudo-anomaly indicators conflict, the operations corresponding to the anomaly indicators will be executed first.
- Build a decision-making database, the data grid contains: 1. Fault description: the specific situation of the system failure; 2. Indicators: abnormal indicators and pseudo-abnormal indicators; 3. Operation: specific execution operations to respond to the failure; 4. Status: Whether the strategy is active.
- The decision database can display the attributes and compliance thresholds of abnormal indicators.
- The user can add, modify and delete the strategy library, and can sort the strategies, that is, adjust the strategy priority, or set the default strategy, that is, set the strategy to the highest priority.
- The system can search for strategies in the decision database based on abnormal indicators or fault symptoms and execute them.

### 4. Conclusion and Prospection

This article introduces a zabbix-based data intelligent operation and maintenance model, using zabbix database for data storage, using the prophet model to perceive and predict the abnormal indicator's value of the service, and analyze the abnormal indicators through the association rule model to obtain the association rules. Finally, combined with the operation and maintenance decision-making library, intelligent execution decision-making is realized. This research can effectively improve the operating efficiency of the database in theory, and can also help operation and maintenance personnel save a lot of time and energy in practice. Since intelligent operation and maintenance is a newly proposed concept in recent years, and there is a lack of theoretical research in traditional large-scale databases, the intelligent operation and maintenance scheme proposed in this paper still needs to be applied and tested in practice.

### References

[1] Xu Chao. *Design and implementation of network operator enterprise operation and maintenance system*[D]. Zhejiang University of Technology, 2019.

[2] *Research on Intelligent Grid Data Center Operation and Maintenance Based on Artificial Intelligence Technology*[C]// CIPAE 2020: 2020 International Conference on Computers, Information Processing and Advanced Education. 2020.
[3] Leon Wu, Gail Kaiser, Cynthia Rudin, Roger Anderson. Data quality assurance and performance measurement of data mining for preventive maintenance of power grid[P]. Data Mining for Service and Maintenance, 2011.

[4] Feng Zhanli, An Jinchao. Design and implementation of a service monitoring system based on time series database[J]. Computer and Network, 2021, 47(06): 60-63.

[5] Li Ming, Zhang Jing, Wu Shang, Xu Daolei. Research on automated operation and maintenance technology based on time series database[J]. Information Technology, 2021(06): 102-107.

[6] Rong Xueqin, Liu Yong, Liu Hao, Bu Shupo. Research on key technologies of power operation and maintenance system based on time series database[J]. Electrical Measurement and Instrumentation, 2018, 55(09): 48-51+83.

[7] Li Jun, Wang Tao. Application and development of network operation and maintenance platform based on big data analysis technology[J]. Computer Programming Skills and Maintenance, 2021(07): 112-114.

[8] Research on Intelligent Grid Data Center Operation and Maintenance Based on Artificial Intelligence Technology[C]// CIPAE 2020: 2020 International Conference on Computers, Information Processing and Advanced Education. 2020.

[9] Tian Yongming, Mei Chenglei, Ma Qiang, Zhao Haiyang. Intelligent operation and maintenance management mode of distribution network based on Internet platform[J]. Computer System Applications, 2021, 30(06): 100-106.

[10] Xiaozhou Du, Yang Yu, Pei Wang, Zhenyu Huang, Bin Wu. Unstructured Log Oriented Fault Diagnosis for Operation and Maintenance Management[P]. Computer Science and Application Engineering, 2019.

[11] Danúbia B. Espíndola, Carlos E. Pereira, Márcio Pinho. Visualization of intelligent maintenance systems through mixed reality[P]. Virtual-Reality Continuum and Its Applications in Industry, 2008.

[12] Tieqiang Wang, Peng Lu, Xin Cao, Xiaodong Yang, Wei Wang, Hao Lv, Chunxian Feng, Pushi Wang. Classification of Power Grid Operation States Based on DNN[P]. Information Technologies and Electrical Engineering, 2019.

[13] Huang Xiaoming, Ling Wanshui, Wu Dongqiu, Yang Tao. Research on practical operation and maintenance indexes of distribution automation based on fault tree analysis method[J]. Power System Protection and Control, 2017, 45(24): 92-98.

[14] Wu Muyang, Liu Zheng, Wang Yang, Li Yun, Li Tao. Network operation and maintenance quality evaluation model based on index correlation[J]. Computer Applications, 2018, 38(09): 2535-2542.

[15] Li Ming, Zhang Jing, Wu Shang, Xu Daolei. Research on automated operation and maintenance technology based on time series database[J]. Information Technology, 2021(06): 102-107.