Filling Algorithm for the Missing Value of Network Stability Monitoring Results Based on Big Data

Pinpin Lyu  
Department of Information Engineering, Zibo Vocational Institute, Zibo, Shandong 255000 China  
*Corresponding Author Email: 10902@zbvc.edu.cn

Abstract. In the research field of network stability monitoring based on big data, there are some disadvantages. For example, the processing method of the missing value of results influences execution effect, interfere the final filling effect and can only be applied in small data set. Therefore, a missing value filling algorithm for network stability monitoring results based on big data is proposed. The evaluation index of network stability monitoring based on big data is selected, and data missing mode and missing value filling mechanism are analyzed. The filling algorithm for the missing value of results based on big data is designed, the missing value filling parameter framework is constructed, the missing value filling parameters are processed, and the filling of the missing value of network stability detection results is completed. Compared with the traditional algorithm, the experimental results show that the efficiency of the designed missing value filling algorithm is 70.3% higher than that of the traditional filling algorithm, and the accuracy rate is far higher than the traditional method.

1. Introduction  
The guarantee of network security is a security measure that must be assured in the Internet era based on big data [1], but there is more or less the loss and lack of data in the working process of network stability monitoring platform, and the treatment methods of lack of the missing value of network stability monitoring results include deletion of incomplete records, being treated as the special value, or interpolation of the empty value. Obviously, regardless of the quantity or quality, the data processing result of the missing value filling algorithm for network stability monitoring based on big data is accurate and efficient, so it is imperative to design a perfect set of missing value filling algorithm.

2. Selection of Evaluation Indicators of Network Stability Monitoring Based on Big Data  
The multimedia network stability monitoring based on big data is composed of data collection and processing modules. There will be many problems in the design process of the platform for information transmission and monitoring, including ensuring the stability of the data during the transmission, because the rapid and steady data transmission is the key of production development. If some problems happen to the information transmission, such as delayed information delivery or damaged information in the process of delivery, they are not conducive to development. In the traditional technology of remote network communication security monitoring, designers will meet the needs of customers to reduce unnecessary costs, because the previous monitoring system used to focus on the data transmission speed, and care less about the security, so the function of the security monitoring platform should have a clear concept, and designers should have clear design goals in their minds [2].
2.1. Data Missing Mode of Network Stability Monitoring

The filling of the missing value of network stability monitoring results based on big data is a very common problem [3]. With the further development of network and computer, the demand for filling results of data missing value is becoming higher and higher, so the missing model of network monitoring should be fully studied to make a better solving plan. A summary of data missing mode in the network stability monitoring system based on big data is shown in Figure 1. It can be found that the missing model of network monitoring is not single but complex and changeable. After selecting the monitoring samples of target network stability, the missing mechanism of network monitoring should be fully understood and corresponding countermeasures should be adopted to carry out the filling calculation of the missing value.

![Figure 1. Data missing model of network monitoring based on big data](image)

2.2. Filling Mechanism of the Missing Value of Network Stability Monitoring

As the data collection work shifts from manual labor to machine, the rapid expansion of data volume makes the quality problems of all kinds of data mixed, among which the problem of filling the missing value of network stability monitoring results is a very common, and the filling mechanism of data missing value is in urgent demand of being formed. The correct filling methods of the missing value are to measure the missing value method, collect the production condition of the missing value and restrain the data missing value [4]. The specific missing value filling mechanism is shown in Table 1.

| Poli Name | advantage | disadvantage |
|-----------|-----------|--------------|
| Central task scheduling strategy | Can the comprehensive each center scheduling nodes | In a lot of computing nodes, the scheduling node |
| The sender launch strategy | No center scheduling nodes, the system light load | The system is overloaded, the sender is not easy to find |
| The receiver launch strategy | No center scheduling node, heavy load system | Under light load conditions, the receiver starts |

From table 1, it can be seen that in the missing value filling mechanism, the missing value calculation method of processing continuous values has a rich theoretical basis and a lot of practical applications are produced. Under the same background of big data, the network stability monitoring results can also handle the discrete value properties well by the corresponding methods, and the missing value filling method of frequency estimation is used to deal with the discrete value missing value [5-6].
3. Design of Filling Algorithm for the Missing Value of Results Based on Big Data

In view of the specific application premise of the missing value of network stability monitoring results based on big data [7], mining the missing value attributes are adopted to construct a new form of filling algorithm for the missing value. The new filling algorithm for the missing value does not set any target attribute, and the most influential filling attribute of the missing value is regarded as the root. The missing value filling is obtained according to the probability reasoning result. The probability information produced by reasoning can reflect the degree of uncertainty of filling value. This provides a reference for evaluating filling quality. In order to make the algorithm suitable for the mixed attribute set, this paper adds the processing of the continuous attribute in the missing value filling algorithm, and all the attributes are filled under one model. For the big data sets, parallel technology is used to solve efficiency problems. The specific design concept is shown in Figure 2.

Figure 2. Design of filling algorithm for the missing value

3.1. Build the Filling Parameter Framework of the Missing Value of Results

When the missing value of the network stability monitoring results based on big data is filled, it is assumed that the data distribution of the current data set is known, and only some parameters of the probability density function are unknown. This missing value filling parameter framework is a statistical inference of the known category samples and the parameters of the population distribution. For the network stability monitoring data set to be processed, the sample category is known, but the overall probability density function form is unknown, which requires direct inference of the probability density function itself. This inference method is called the missing value filling parameter estimation. After the selection of the missing value filling parameters, the corresponding optimal window width should be chosen. In fact, the selection of window width is more important than that of missing value filling parameters. The existing methods are only applicable to certain fields, and they are more complicated than single missing value filling parameters in form. Therefore, missing values are used to fill the missing value in the parameter algorithm to compute the following formula:

$$K_{\text{max}} = \rho K_{\text{poly}} + (1-\rho)K_{\text{rbf}}$$  \hspace{1cm} (1)

Among them, $K_{\text{poly}} = (\leq x, x_i \geq +1)^q$, $K_{\text{rbf}} = \exp\left(-\frac{(x-x_i)^2}{\sigma^2}\right)$, and the best mixing parameter is $\rho (0 \leq \rho \leq 1)$. $\rho$ is the smoothing parameter of polynomial kernel, and the window width parameter of RBF is $\sigma$. Compared with other methods with rich theories, the missing value
filling kernel function framework theory is more practical in terms of technology, and big in terms of the number of research results.

3.2. Processing of Missing Value Filling Parameters
The processing of missing value filling parameters usually includes two parts: one is to select missing value filling parameters, and the other is to select the optimal window width. Correct selection of kernel function can reduce the calculation time. In the processing of the missing value filling parameters, they are analyzed and arranged according to the missing data characteristics. The overall assumption of the missing value filling parameters is small, it has extensive applicability, and all types of missing value filling parameter data can be processed. After the reliable probability density is obtained, a classifier or filler is established to fill the missing value. Before filling missing value data of the network stability monitoring results based on big data is completed, the network monitoring missing data can be initially arranged according to the calculation method of the parameter formula. Supposed that \( \lambda \) represents the format coefficient of the characteristic data of the network stability monitoring, \( t \) represents the basic transformation period of the data. Using \( \lambda \) and \( t \), the optimized systematic parameter feature extraction operator can be expressed as follows:

\[
q = \int \frac{h(\lambda^2 + y)}{3p + e}
\]

(2)

Among them, \( q \) represents systematic parameter feature extraction operator, \( y \) represents the storage coefficient in the network stability monitoring database, \( p \) represents the parameter transformation of the network stability monitoring, and \( e \) represents the total execution operation of the remote operation code. With the increase of the operation time of the system, the total amount of running parameters matching the stability monitoring data of the large data network will also increase continuously. Under the influence of the feature extraction operator, the total amount of the operation parameter that is matched with network stability monitoring data based on big data will continuously increase. Under the influence of the feature extraction operator, \( i \) is supposed to represent the network stability monitoring feature data and the basic matching parameter of the system database. Combined with and the formula (2), the improved systematic configuration parameter can be expressed as follows:

\[
R = \prod_{i} i + \frac{(\sigma + 1)\mu}{f - l}
\]

(3)

Among them, \( R \) represents the improved operating configuration parameter of the system, \( \sigma \) represents the reciprocal of the storage limit of the system database, \( \mu \) represents the fixed retrieval index of the parameters, \( f \) represents the remote independent control factor of the system, and \( l \) represents the base amount of the characteristic data. This is essentially an interpolation process. In this way, the interval can be in any form (decided by the selection of kernel function). It does not stick to histograms or cubes. If there are enough samples, no matter which kernel function form is used, a reliable estimation result converges to the density can be obtained theoretically.

4. Experimental Results and Analysis
In order to verify the practical value of the missing data filling algorithm based on big data, the following comparative experiments are designed. Two computers with the identical configurations are used as the experimental targets. The computer, with the common missing value calculation method, was used as the control group; the computer, with the missing value calculation method in this paper, was used as the experimental group. In the case of keeping the other conditions unchanged, the accuracy of analysis of the network stability monitoring based on big data is recorded after the application of experimental and control groups.
4.1. Setting of Experimental Parameters

In order to ensure the authenticity of the experimental results, the relevant experimental parameters can be set according to the following table. The parameters in the above table respectively represent the experimental time, the computer network monitoring, the basic analysis parameters and the upper limit. In order to ensure the authenticity of the experimental results, the experimental parameters of the experimental and control groups are always consistent.

Table 2. Experimental parameter setting table

| Mixing | Poly  | RBF  |
|--------|-------|------|
| EMT/(min) | 70    | 70   |
| BBF/(%) | 83.2  | 83.2 |
| MRC/(t)  | 0.94  | 0.94 |
| BAP      | 0.45  | 0.45 |
| ULA/(%)  | 83.7  | 83.7 |

4.2. Calculation Accuracy Comparison of the Missing Value of Network Stability Test Results

Under the condition of the basic analysis parameter being 0.45, 70min was used as the experimental time to respectively record the changes in the accuracy of the basic analysis result of the economic situation after the application of the experimental and control group models. The comparison of the specific experiment is shown in Table 3.

Table 3. Repeat filling times and missing value complement

| Normal | Experimental accuracy |
|--------|-----------------------|
| 80.0   | 5 69.2                |
| 80.0   | 10 69.0               |
| 80.0   | 15 67.8               |
| 83.5   | 20 68.5               |
| 83.5   | 25 67.9               |
| 83.5   | 30 67.6               |
| 86.7   | 35 67.1               |
| 85.4   | 40 67.8               |

Figure 3. Curve contrast graph
After analyzing the data in Table 3, the drawing curve graph is shown in Figure 3. It can be seen in Figure 3 that with the increase of experimental time, the calculation result accuracy of the missing value of network stability monitoring based on big data presents a step up and a step downward trend after the application of experimental group model. When the experimental time is between 30-35min, the calculation result accuracy of the missing value of network stability monitoring based on big data reaches the maximum value of 80.6%, exceeding the target upper limit of 73.5%. After the application of the control group model, the calculation result accuracy of the missing value of network stability monitoring based on big data presents a step downward and a step up trend. When the experimental time is 5min, the calculation result accuracy of the missing value of network stability monitoring based on big data reaches the maximum value of 70.3%, far lower than that of experimental group. The usual method of continuous value is better in the number of repeated filling and the mean of convergence. The Normal method is only designed for the conditional attribute of the continuous value. When the condition attribute of the discrete value are dealt with, the discrete value can only be regarded as a continuous value. Comparing the four hybrid kernel methods, the algorithm is obviously the best.

5. Conclusion
On the basis of the existing technical means, the filling algorithm for the missing value of network stability monitoring results based on big data can reach the expected level by upgrading the algorithm. Through the analysis, it can be seen that the application of new means is very effective for the filling calculation of the missing value of network stability monitoring results based on big data. It can break the drawbacks of the traditional methods brought to the processing method of missing value results, and has a strong practicability.

6. Acknowledgements
The study was supported by the key R & D projects of Zibo City, Shandong Province in 2018( Grant No.2018kj010132); Zibo city in 2019 focus on research and development program (school city fusion development class)(Grant No. 2019ZBXC207); and the 1st development projects of school city fusion in Zhoucun district (Grant No. 2020ZCXCZH04).

7. References
[1] He Xiaoyong, Hou Dongjin. Simulation and optimization of resource information characteristics of big data repository [J]. Computer simulation, 2018, 35(06):380-383+455.
[2] Liu Xiangdong, Chen Jinan. A Differentiated Strategy of Tourism E-Commerce Product Promotion Based on Provincial Tourism Development Difference: Based on Large Data and Data Visualization Method [J]. Statistics & Information Forum, 2017, 32(8):31-38.
[3] Liu Yi, Bao Jigang, Chen Kaiqi. Sentimental Features of Chinese Outbound Tourists in Australia: Big-data Based Content Analysis [J]. Tourism Tribune, 2017, 32(5):46-58.
[4] Guo Jinyu, Li Enping. Research on the Impact on Innovation Ability of Small and Medium-sized Technology-based Enterprises under the Background of Big Data — Based on the fuzzy comprehensive evaluation model [J]. Mathematics in Practice and Theory, 2016, 46(10):77-84.
[5] Bi Hexia. Construction of confirmation model for poor college students under big data -- An Empirical Study Based on the integration of "fuzzy comprehensive evaluation" and "fuzzy analytic hierarchy process" [J]. Higher Education Exploration, 2016, 16(8):105-114.
[6] Cai Liyuan, Zhang Jinhai. Negative entropy: the construction of the five stage model of TPWKKR enterprise marketing in the era of big data -- Taking the five stage model of purchase as the object of analysis [J]. Modern Communication, 2016, 38(5):127-130.
[7] Tang Jing, Li Ruixuan, Huang Yuhang, er al. Research on the accurate forecasting of monthly electricity demand based on the multi-feature analysis [J]. Power System Protection and Control, 2017, 45(16):145-150.