

PAPER • OPEN ACCESS

Adaptive threshold modeling algorithm for monitoring indicators of power network server based on Chebyshev inequality

To cite this article: Benran Hu *et al* 2020 *J. Phys.: Conf. Ser.* **1693** 012089

View the [article online](#) for updates and enhancements.

You may also like

- [Quantum processing by remote quantum control](#)
Xiaogang Qiang, Xiaoqi Zhou, Kanin Aungskunsiri et al.
- [Data center growth in the United States: decoupling the demand for services from electricity use](#)
Arman Shehabi, Sarah J Smith, Eric Masanet et al.
- [Mediated semi-quantum key distribution with improved efficiency](#)
Julia Guskind and Walter O Krawec



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Adaptive threshold modeling algorithm for monitoring indicators of power network server based on Chebyshev inequality

Benran Hu¹, Yanjun Li¹, Jiyuan Ren², Yiqun Li¹

¹Information & Communication Branch, Heilongjiang Electric Power Company (LTD), SGCC, Harbin 150090, China;

²Power Dispatching & Control Center, SGCC Northeast China Branch, Shenyang 110180, China

lyq300@126.com

Abstract. The business system server under the IT automation operation and maintenance platform generates massive data samples, based on which the threshold can be set to realize the allocation and management of hardware resources. The traditional threshold selection method is to determine an appropriate threshold based on human experience. If the threshold is too high, it will not play its due role. But if the threshold is too low, it will frequently produce false positives. To solve this problem, an adaptive threshold method based on Chebyshev inequality theory combined with kernel density estimation is proposed to determine monitoring indexes, and a new dynamic implicit threshold model is established to analyse the data generated by the business system server for real-time monitoring and alarm processing. Through the experimental study on the CPU utilization data of the power grid server, lower missing and false positive rate are obtained, which verifies the feasibility and effectiveness of the proposed method.

1. Introduction

The power grid company strives to promote high-quality development of power grid, and strives to promote scientific and technological innovation, which puts forward newer and higher requirements for the technical support capacity of information and communication operation. However, it is difficult for power grid company to copy the automation operation and maintenance experience of Internet enterprises. The power grid company proposed a new "Internet+ traditional" operation and maintenance technology architecture, IT automation operation and maintenance system, which can support the operation and maintenance work to advance to the automatic and intelligent mode, and improve the operation guarantee capability of information and communication.

The monitoring center is the monitoring strategy and configuration center of IT automation operation and maintenance platform. Through the construction of the monitoring center, the configuration and maintenance of monitoring strategies can be realized, and the original monitoring data and events can be converted into alarms according to the established strategy. At the same time, it also supports the transformation of upper-level business requirements into landing strategies, and provides monitoring and alarm data services for various upper-level business presentation modules [1].

A large number of business system servers under the IT automation operation and maintenance platform will generate massive data business every day, which makes the performance of the hardware



resources CPU, memory and hard disk of the business system server face a serious test. CPU, memory and hard disk are the core parts of hardware resources of the business system server equipment [2], and are the most easily overloaded parts of the system under large data. Therefore, taking CPU, memory and hard disk as objects, it is of certain practical significance to study the monitoring and early warning adaptive evaluation system of business system server [3].

The threshold of CPU, memory and hard disk utilization of traditional business system server equipment is set by operation and maintenance personnel based on experience. Once the threshold is set, it will remain unchanged for a long time. It cannot meet the massive real-time data needs of the monitoring center, and cannot predict the potential risks, which will easily cause economic losses. Therefore, it is of great significance to study the monitoring and early warning methods applicable to the actual operating environment of the business system server equipment of the State Grid IT automation operation and maintenance platform, to adaptively set dynamic thresholds for the CPU, memory, and hard disk of the business system server, to predict potential risks or hidden dangers, and to trace the source of attack events.

At present, scholars have done some research on the extension and application of Chebyshev inequality, which has been applied in various fields [4]. Yang [5] puts forward a calculation method based on Chebyshev inequality for the reasonable interval of power grid engineering cost, which reduces the dependence on experience. Shan et al. [6] proposed an adaptive threshold algorithm for image segmentation and background extraction and update. By using Chebyshev inequality to determine the optional range of thresholds, the background and target of the differential image can be accurately segmented. Xiang et al. [7] calculated the segmentation threshold adaptively through the image dark channel histogram distribution, so that the sudden change of depth of field at image interface were relatively smooth.

The kernel density estimation method has also been widely used in various fields [8,9,10]. He et al. [11] constructed a non-parametric multivariable kernel density estimation load model to achieve approximate estimation of joint probability density of node loads, and effectively incorporated it into the power grid reliability evaluation model. Based on year-by-year reliability index samples, Zhao et al. [12] used the non-parametric kernel density estimation theory to realize probability density estimation of reliability indexes, and overcome the shortcoming of traditional expected value indexes that only measure system reliability from the perspective of probability average meaning. Therefore, the kernel density estimation method can provide a new idea for power grid server monitoring.

To sum up, in view of the lack of mathematical theory guidance and engineering practice analysis in the current power grid server monitoring methods, based on the theory of Chebyshev inequality and the kernel density estimation method, this paper uses the historical data in recent years to determine the adaptive threshold of each monitoring index. A new dynamic implicit threshold model is established to analyze the data generated by the business system server and realize real-time monitoring and alarm of the relevant business system server.

2. Chebyshev Principle

In the study of statistical laws by the Russian mathematician Chebyshev, he demonstrated and expressed an inequality with the standard deviation. This inequality has universal significance. Chebyshev inequality has a very high status in the limit theory of probability theory. It is the theoretical basis for proving the law of large numbers and a powerful tool for studying the central limit theorem. Chebyshev inequality is determined by the distribution of random variables, and can describe the characteristics of a certain aspect of random variables, among which the most important numerical characteristics are mathematical expectation $E(X)$ and variance $D(X)$. It gives an estimation method for the probability of the event ex when the distribution of random variables is unknown and only $E(X)$ and $D(X)$ are known. Suppose a random variable has mathematical expectation $E(X)=\mu$ and variance $D(X)=\sigma^2$, then for any positive number ε , there are inequalities:

$$P\{|X - \mu| \geq \varepsilon\} \leq \frac{\sigma^2}{\varepsilon^2} \quad (1)$$

$$P\{|X - \mu| < \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2} \quad (2)$$

Eq. (1) or (2) holds. This inequality is called Chebyshev inequality. It can be seen from Chebyshev inequality that for a given ε value, the smaller $D(X)$ is, the larger $P\{|X - \mu| < \varepsilon\}$ is. At this point, the value of random variable X is basically concentrated around $E(X)$. When $E(X)$ and $D(X)$ are known, Chebyshev inequality gives a lower bound of probability $P\{|X - \mu| < \varepsilon\}$. The lower bound does not involve the specific probability distribution of the random variable X , but is only related to its variance $D(X)$ and ε .

In the server CPU utilization time series, random variables are introduced to represent the value of each data point in the time series. The probability distribution of the random variable is unknown, but the population mean and variance of the pixel probability distribution can be estimated by calculating the sample mean and sample variance. Through Chebyshev inequality calculation formula, the probability estimation of the event $\{|X - E(X)| < \varepsilon\}$ is obtained, and this event exactly reflects the change of the CPU utilization value corresponding to a certain point in a series of time series data.

3. Kernel Density Estimation

In statistics, it is often necessary to infer the population distribution, namely the density function, from the sample data. If the method of parameter estimation is adopted, the specific form of the population distribution should be assumed first. For example, the population follows the normal distribution $N(\mu, \sigma^2)$, and then the parameter (μ, σ^2) is estimated by using the sample data to obtain the population density function. However, if the real population is far from the assumed distribution, the statistical inference obtained by the parameter estimation method may have a large deviation. The non-parameter estimation method can estimate the density function without assuming the population distribution, thus reducing the error. Kernel density estimation is derived from the idea of histogram density estimation by Rosenblatt and Parzen. When using histogram for density estimation, even if the random variable is continuous, the histogram is always a discrete step function. Histogram method is a non-parametric method of density estimation. Although histogram method is intuitive and simple, the probability density obtained is discontinuous due to the limited sample data. The kernel density estimation can solve this shortcoming and obtain a smooth estimation of density function. The core of kernel density estimation method is to use a smooth and differentiable kernel function. In order to make the probability density continuous, the influence of each observed sample on the density should also be continuous, and its influence on the density should decrease smoothly as the distance increase. The kernel density estimator is:

$$\hat{f}(x_0) = \frac{1}{nh} \sum_{i=1}^N K\left(\frac{x_i - x}{h}\right) \quad (3)$$

Where, the function $K(\cdot)$ is called the kernel function and h is called the bandwidth. Eq. (3) satisfies the following properties:

- (i) $K(\cdot)$ is continuous and symmetric about the origin.
- (ii) $\int K(z)dz = 1$, $\int zK(z)dz = 0$, $\int |K(z)|dz$ is less than infinity.
- (iii) $Z_0 > 0$, when $|z| \geq z_0$, $K(z) = 0$.
- (iv) $\int z^2 K(z)dz = k$, where k is a constant.

Condition (iii) requires the area under the curve of the kernel function to be 1, standardizes the function, and satisfies some bounded conditions. Under condition (iii), ① is stronger than ②. In practice, condition ① is often used, that is, if the weight beyond a certain neighbourhood scope $[-z_0, z_0]$ becomes 0, $[-z_0, z_0]$ is usually normalized to $[-1, 1]$. Condition (iv) is also a bounded condition.

The "variance" of the selected kernel function is determined by h , and the kernel function estimation results under different bandwidths vary greatly. If the bandwidth is not fixed and its variation depends on the estimated location or sample point, a very powerful method called adaptive or variable bandwidth kernel density estimation can be generated.

4. Adaptive Threshold Modeling Based on Chebyshev Inequality

An adaptive threshold modeling algorithm combining Chebyshev inequality and kernel density estimation is designed. An adaptive classification threshold method is proposed to quickly identify data points with significant faults or normal characteristics by using Chebyshev inequality. The kernel density estimation algorithm is used to refine the data points that are difficult to classify. The specific algorithm is described as follows: random variable X is introduced to represent the data of a certain server, and X_k is set to represent the utilization rate of the server at point k , and the population probability distribution of random variable X is unknown.

(1) Read the adjacent N utilization data points, and get the sample $X_1, X_2, \dots, X_N$ and its observed values $x_1, x_2, \dots, x_N$.

(2) Calculate the sample mean of population: $\bar{X} = (1/N) \sum_{k=1}^N x_k$, and second order center distance of sample $S^2 = (1/N) \sum_{k=1}^N (x_k - \bar{X})^2$. Since $\bar{X}$ and S^2 are the maximum likelihood estimators of the mathematical expectation and variance of random variable X , the estimated $\hat{\mu} = \bar{X}$, $\hat{\sigma}^2 = S^2$ of the mathematical expectation μ and variance σ^2 of the random variable X is obtained.

(3) Calculate the Chebyshev inequality corresponding to the K th data point $X_k (k=1, 2, \dots, N)$ as:

$$P_C \{ |X_k - \hat{\mu}| < \varepsilon \} \geq 1 - \frac{\sigma_k^2}{\varepsilon^2} \quad (4)$$

Where, $\sigma_k^2 = (X_k - \hat{\mu})^2$ represents the square of the difference between the utilization rate of data point X_k and the mathematical expected estimator $\hat{\mu}$. When $\sigma_k^2 < \hat{\sigma}^2$, the change degree of data point X_k is smaller than mean square error, and this point is more likely to be a normal point, then $1 - \sigma_k^2/\varepsilon^2 > 1 - \hat{\sigma}^2/\varepsilon^2$. If $\sigma_k^2 > \hat{\sigma}^2$, the data point is more likely to be an outlier, then $1 - \sigma_k^2/\varepsilon^2 < 1 - \hat{\sigma}^2/\varepsilon^2$. In Chebyshev inequality, only ε is required to be a positive number. ε plays an important role in the probability $P\{|X - \mu| < \varepsilon\}$. Taking different values of ε , the calculation results of $1 - \hat{\sigma}^2/\varepsilon^2$ are listed in Table 1.

Table 1 Statistical table of probability estimates

ε	σ^2/ε^2	$1 - \sigma^2/\varepsilon^2$
$\sqrt{3/2}\sigma$	2/3	1/3
$\sqrt{2}\sigma$	1/2	1/2
$\sqrt{5/2}\sigma$	2/5	3/5
$\sqrt{3}\sigma$	1/3	2/3
$\sqrt{7/2}\sigma$	2/7	5/7
2σ	1/4	3/4

(4) Set the adaptive classification threshold, the setting algorithm of the adjustment coefficient θ_1 , θ_2 and the discrimination threshold T_1, T_2 is [13]:

$$\begin{cases} T_1 = ((1 - \hat{\sigma}^2)/\varepsilon_1^2 + 0.5)/2, \varepsilon_1 = 1.414\hat{\sigma} + \theta_1\hat{\sigma} \\ T_2 = ((1 - \hat{\sigma}^2)/\varepsilon_2^2 + 0.5)/2, \varepsilon_2 = 1.414\hat{\sigma} - \theta_2\hat{\sigma} \end{cases} \quad (5)$$

The setting of thresholds T_1 and T_2 varies with the variance of the sample mean square error, as shown in Fig.1.

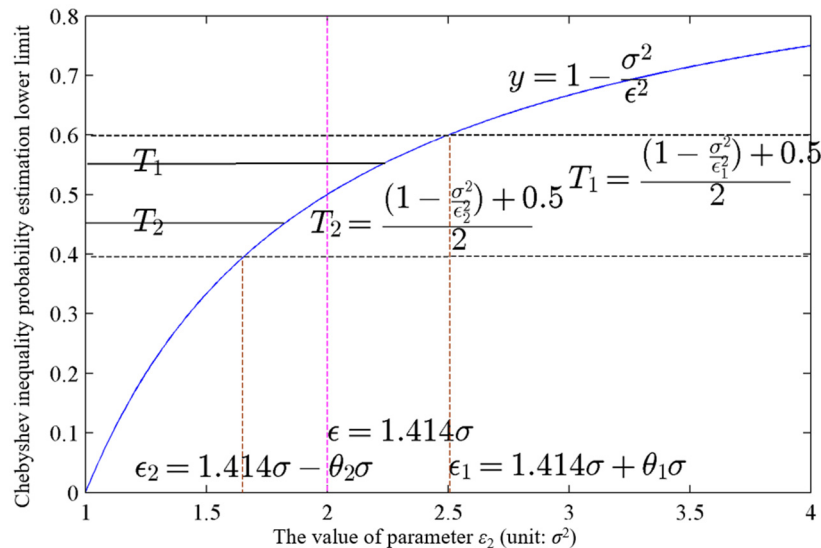


Fig.1 Pixel classification threshold setting

For different samples $X_1, X_2, \dots, X_N$, the algorithm can be adaptive to adjust the threshold, and can effectively perform fault discrimination under different server selection conditions. Adjust the setting of the coefficients θ_1 and θ_2 so that $0 < |\varepsilon_1 - \varepsilon_2| < 2\hat{\sigma}$. The adjustment function can be symmetric ($\theta_1 = \theta_2$) and asymmetric ($\theta_1 \neq \theta_2$). According to the definition of T_1 in Figure 1, the greater θ_1 is, the greater T_1 value is, and the higher the probability of data points being judged as abnormal points is. Similarly, the higher the θ_2 is, the smaller the T_2 value is, and the higher the probability of the data points being judged as normal points is.

(5) Set the piecewise function $M_k^1(i)$, and classify the normal points, abnormal points and suspicious points of the data points as:

$$M_k^1(i) = \begin{cases} 1, (1 - \frac{\sigma_k^2(i)}{\varepsilon^2}) \geq T_1 \\ \varepsilon = 1.414\hat{\sigma} \\ 0, (1 - \frac{\sigma_k^2(i)}{\varepsilon^2}) \leq T_2 \end{cases} \quad (6)$$

When $M_k^1(i) = 1$, the data points are divided into abnormal points. When $M_k^1(i) = 0$, the data points are judged as normal points. And when $T_2 < (1 - \sigma_k^2(i))/\varepsilon^2 < T_1$, the data points are suspicious points. We will use the kernel density estimation algorithm to make the second discrimination.

(6) For suspicious points, use the kernel density estimation algorithm to further distinguish abnormal points and normal points. Set the probability threshold T_3 , and define the following binary function:

$$M_k^2(i) = \begin{cases} 1, P_r(x_k) \geq T_3 \\ 0, P_r(x_k) < T_3 \end{cases} \quad (7)$$

The data point with $M_k^1(i) = 1$ is judged as an abnormal point, and the point corresponding to $M_k^1(i) = 0$ is judged as a normal point. If the probability threshold T_3 is selected too large, the missing rate of the fault point will increase. If the selection is too small, the false positive rate will be greater. T_3 in the kernel density estimation method is an empirical value, which is generally selected through specific experimental data.

When selecting the kernel function, the more commonly used Gaussian kernel function can be selected, and the kernel density estimation formula is:

$$P_r(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(x-x_i)^2}{\sigma^2}} \quad (8)$$

5. Experimental Verification

The experimental data were provided by a state power grid monitoring center, and the data of three servers, server 314, server 357, and server 504, were experimentally verified. The experimental data is the CPU utilization data collected by the monitoring equipment every 5 minutes. The CPU utilization data of server 314 is from May 1 to May 31, a total of 7048 monitoring points. The CPU utilization data of server 314 is shown in Fig.2. The CPU utilization data of server 357 is from May 1 to May 31, a total of 8544 monitoring points. The CPU utilization data of server 357 is shown in Fig.3. The CPU utilization data of server 504 is from April 1 to May 24, a total of 15152 monitoring points. The CPU utilization data of server 504 is shown in Fig.4.

The data monitoring results of 3 servers were checked by random sampling. Taking server 314 as an example, the experimental steps are as follows:

(1) Read 7048 adjacent utilization data points to obtain sample $X_1, X_2, \dots, X_{7048}$ and its observed values $x_1, x_2, \dots, x_{7048}$.

(2) Calculate the sample mean of population: $\bar{X} = (1/N) \sum_{k=1}^N x_k$, and second order center distance of sample $S^2 = (1/N) \sum_{k=1}^N (x_k - \bar{X})^2$. Since $\bar{X}$ and S^2 are the maximum likelihood estimators of the mathematical expectation and variance of random variable X , the estimated $\hat{\mu} = \bar{X}$, $\hat{\sigma}^2 = S^2$ of the mathematical expectation μ and variance σ^2 of the random variable X is obtained.

(3) Find the σ_k^2 corresponding to the K th data point X_k .

(4) According to Eq. (5), take $\theta_1 = \theta_2 = 0.5$ to determine the thresholds T_1 and T_2 . According to Eq. (4), the probability P_c of Chebyshev inequality estimation is calculated, and compared with T_1 and T_2 , the abnormal point and the normal point are preliminarily determined.

(5) According to Eqs. (7) and (8), the remaining suspicious points are further judged.

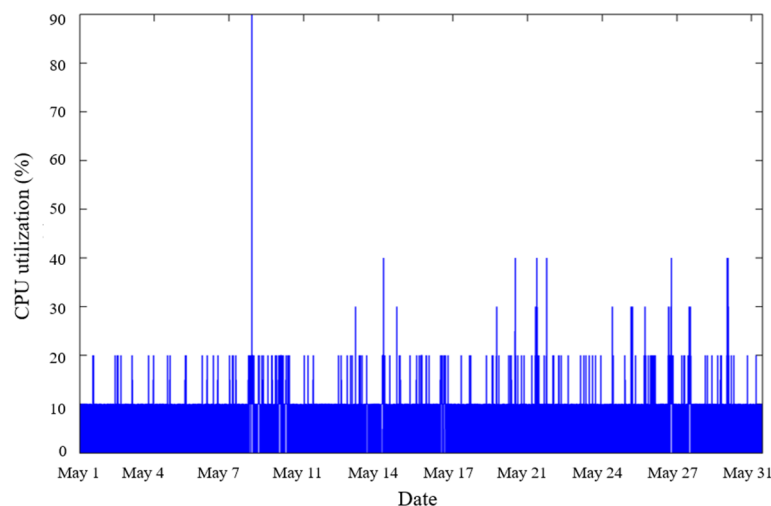


Fig.2 CPU 314 monitoring data

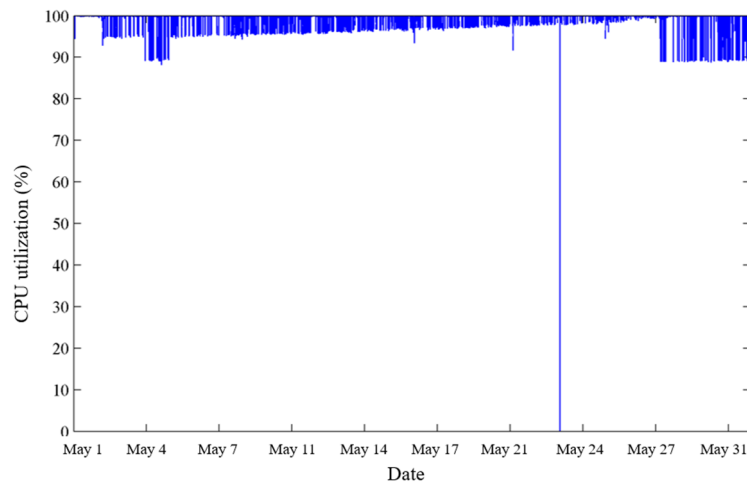


Fig.3 CPU 357 monitoring data

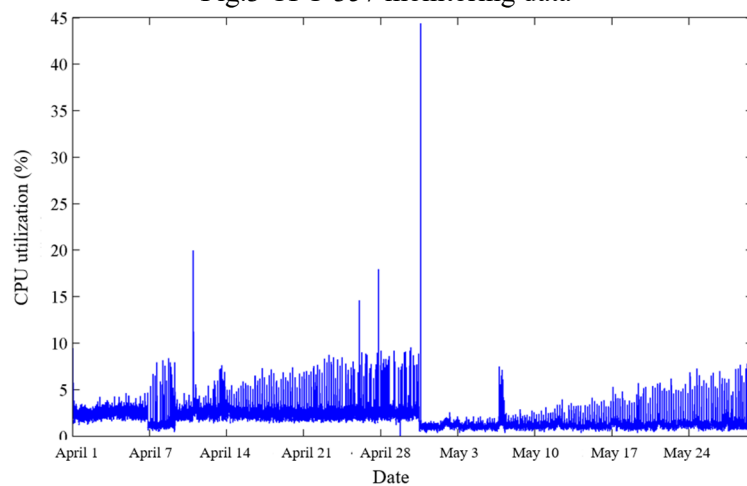


Fig.4 CPU 504 monitoring data

The CPU utilization data processing process of server 357 and server 504 is consistent with the above steps and will not be repeated.

In order to further prove the effectiveness of this method, this paper introduces two indexes: missing rate and false positive rate. The missing rate $R_{missing}$ is the ratio of the number of undetected abnormal points FP to the number of all normal points P . The false positive rate R_{false} is the ratio between the number of normal points in all detected abnormal points FN and the number of detected abnormal points N . The defined distribution is shown in Eqs. (9) and (10).

$$R_{missing} = \frac{FP}{P} \quad (9)$$

$$R_{false} = \frac{FN}{N} \quad (10)$$

The missing rate of server 314 is 0.03%, and the false positive rate is 0.05%. The missing rate of server 357 is 0.012%, and the false positive rate is 0.26%. The missing rate of server 504 is 0.06%, and the false positive rate is 0.31%.

6. Conclusion

This paper mainly studies the problem of adaptive threshold modeling in power grid server monitoring system, and proposes an adaptive threshold background modeling algorithm based on Chebyshev inequality combined with kernel density estimation. At present, there is a lack of monitoring methods for hardware resource allocation of power grid server, and the artificial setting of threshold requires a

lot of experience and is not accurate. The proposed method realizes intelligent monitoring of power grid server CPU data, and experiments are carried out in three situations where the CPU utilization value distribution is different. The missing rate and false positive rate of this method are at a low level for each server. The experimental results show that the method is feasible and the setting of adaptive threshold is effective. In summary, this method has good real-time performance and effectiveness, and is suitable for real-time monitoring of power grid server system.

The business system server under the IT automation operation and maintenance platform will generate massive data in the work. Therefore, using deep learning to mine the deep information of the data and monitoring real-time issues based on this will be the focus of the next research work.

Acknowledgments

This work was financially supported by State Grid Heilongjiang Electric Power Company (LTD) project funding.

References

- [1] Hua Aibing. Research and implementation of intelligent operation and maintenance management system based on big data[J]. Information & communication, 2017(11): 239-240(in Chinese).
- [2] Huang Jinglian. Research on monitoring system of data center server operation state[J]. Software Guide, 2016, 15(09): 140-142(in Chinese).
- [3] Wen Jing. Research on the method of dynamic deployment of virtual machine based on prediction of CPU utilization[D]. Northeastern University, 2014: 5-11(in Chinese).
- [4] Li Chunxiang, Zhao Guihua, et al. Generalizations and applications of Chebyshev's inequality based on variance[J]. Statistics & Decision, 2017(02): 68-70(in Chinese).
- [5] Yang Jing. Study on Chebyshev inequality based calculation method for reasonable range of power grid engineering cost[J]. Modern Electronics Technique, 2017, 40(22): 48-50+53(in Chinese).
- [6] Shan Yugang, Wang Jiabao, et al. Research on Surendra background extraction method based on adaptive threshold[J]. Computer Technology and Development, 2017, 27(07): 91-95(in Chinese).
- [7] Xiang Wending, Yang Ping, et al. Image dehazing algorithm based on adaptive threshold segmentation and transmittance fusion[J]. Computer Engineering, 2019, 045(010): 239-245(in Chinese).
- [8] Sheng Kui, Ma Jian. The internet of things clustering analysis model based on kernel density estimation[J]. Control Engineering of China, 2018, 25(06): 1098-1102(in Chinese).
- [9] Wang Xiaotong, Cai Zhiming. Moving target detection in reverberating background based on kernel density estimation[J]. Journal of Harbin Engineering University, 2019, 40(04): 813-819(in Chinese).
- [10] Yang Xiyun, Zhang Yanfeng, et al. Prediction of combination probability interval of wind power based on naive bayes[J]. High Voltage Engineering, 2020, 46(03): 1099-1108(in Chinese).
- [11] He Yaoyao, Liu Rui, et al. Short-term power load probability density forecasting method based on real time price and support vector quantile regression[J]. Proceedings of the CSEE, 2017, 37(03): 768-776(in Chinese).
- [12] Zhao Yuan, Shen Zhijian, et al. Reliability assessment of bulk power systems utilizing sequential simulation and nonparametric kernel density estimation[J]. Automation of Electric Power Systems, 2008(06): 14-19(in Chinese).
- [13] Zhang Kun, Wang Cuirong, et al. Adaptive threshold background modeling algorithm based on Chebyshev inequality[J]. Computer Science, 2013, 40(04): 287-291+297(in Chinese).