Network Equipment Fault Prediction Based on CUSUM Algorithm

Liu Peishun[1, a], Lin LiFang[2,b] and Zhao Changjian[3,c], Jia Peiwei[3,c]

[1] Ocean University of china, Qingdao, Shandong ,China

[2]Qingdao Huanghai University, Qingdao, Shandong, China

[3] Shanxi Heli Creative Technology Limited, Taiyuan, Shanxi, China

[a]liups@ouc.edu.cn, [b]413142048@qq.com, [c]Zhaochj@hlcx.cn

Keywords: Fault Prediction, CUSUM algorithm, Network Equipment, Anomaly detection.

Abstract. In this paper, the CUSUM algorithm is used to detect and predict the state of network equipment. By collecting information on network equipment operating characteristics, features of the device obtained sample set, the application design is complete training sample set and obtained parameters of algorithm, build fault prediction model based on CUSUM algorithm to realize the equipment fault predictive analysis. The experimental results show that the detection algorithm can detect the abnormal state of the system in advance, and provide the decision-making basis for the equipment fault prediction.

Introduction

With the development of Internet technology, network is becoming more and more important in social life. The network equipment such as routers, switches and other equipment are responsible for the transmission of network data. Network equipment failure can cause the entire network paralysis disease. Therefore, to ensure the normal operation of the network equipment, after a failure to quickly and accurately locate the problem and troubleshooting, is the core task of network maintenance and management personnel work[1].

Fault prediction technology is to analyze the characteristics of the equipment fault, according to the historical information to predict the future, it can calculate the future operation of the health status of the equipment. In recent years, the technology of fault prediction has been greatly developed. According to the method of fault feature, the fault prediction technology is divided into 3 categories: the expert system, the data driver and the model based method. The expert system is suitable for the equipment running data is very few, and the model is missing, The data driven approach is applied to the case of a sufficient device to run the data, and The model-based method is suitable for the model easier to obtain system, the inherent failure mechanism more clearly, but the method depends on the accuracy of the model[1,2].

Based on the operating characteristics of network devices for statistical analysis, this paper carries out intelligent forecasting of network equipment failure used CUSUM algorithm. The experiment proves that the target can reach the goal of early warning, and provide important basis for predicting failure.

Network equipment fault detection algorithm

Fault prediction of network equipment is based on the analysis of the operating parameters of the equipment. Run time of the device, chip temperature, CPU usage, and interface can reflect the health status of the device, select them as the characteristic variable predictive model. Common sense can make the following 5 points based on reasonable assumptions [2]:

1) The longer the operation time of the equipment, the greater the probability of failure.
2) The higher the temperature of the chip, the greater the probability of failure.
3) The higher the usage of CPU, the greater the probability of failure.
4) The higher the memory utilization rate, the higher the probability of failure.

5) The higher the packet loss rate, the higher the bit error rate, the higher the probability of failure.

In statistical quality control, the CUSUM (or cumulative sum control chart) is a sequential analysis technique developed by E. S. Page of the University of Cambridge. It is typically used for monitoring change detection. In industrial monitoring, CUSUM algorithm is usually used to detect the abnormal changes of product attributes in the process of production, which has the advantages of small computation, rapid detection and simple implementation [3]. Exception monitoring of production process and network equipment fault detection is similar, so it can be used CUSUM algorithm for network equipment fault detection.

Let x_i be the i th observation on the process and distributed as normal with mean $\delta = (\sum_{i=1}^{n} x_i)/n$ and variance $\sigma = k_\sigma * \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \delta)^2}$. We also assume that the δ as the target value for the quality characteristic x . The Tabular Cusum [3-5]

$$Z_0 = 0, Z_n = \max\{0, Z_{n-1} + x_n - \delta - K\} \quad n = 1, 2, \ldots \quad (1)$$

K is usually called the reference value and often chosen halfway between the target value μ_0 and out-of-control value of the mean μ_1 that want to detect quickly. Normally $\mu_0 = \delta = (\sum_{i=1}^{n} x_i)/n$, thus if $\mu_1 = \mu_0 + y*\sigma$, then [3-5]

$$K = \frac{y}{2} \sigma = \frac{|\mu_0 - \mu_1|}{2}$$

If Z_n exceed the decision interval H, the process is considered to be out of control. The reasonable estimate of H is H = 5*σ .

Using the CUSUM algorithm to detect the abnormal detection is easy to produce cumulative effects problem: the system has returned to normal state from the abnormal state, but the Z_n is higher than the threshold h, so that false positives. In order to solve this problem, this paper provides that Z_n no longer maintained at h when Z_n>H, when the system returns to normal state, the Z_n is less than H, so the cumulative effects is eliminated.

Application of algorithm in network equipment fault prediction

Using CUSUM to detect network equipment failure, first the characteristic sequence of network equipment such as the board temperature, CPU usage, memory usage, packet loss rate, and error rate are obtained. Then, the characteristic parameters of the device are trained by the training algorithm. Finally the Characteristic sequence of the network equipment are collected and calculated by the training parameters. The CUSUM value of the current network equipment is judged to be normal or abnormal, and if the abnormal is issued an alert, the future network equipment may occur. If the false alarm system will adjust the feature parameters, eliminate false positives.

According to this idea, this paper designs a model training algorithm and a fault detection algorithm. The training algorithm flow is shown in Figure 1. The training steps are as follows:

Step 1: Get the characteristic variables of the normal operation of network equipment, the acquisition time is at least 2 days, the time interval is 5-30 minutes, the time interval is longer, the time interval is shorter, and the characteristic is calculated. We mainly collect chip temperature, CPU usage, memory usage, interface packet loss rate, bit error rate as characteristic sequence. Let the sequence is x_n.

Step 2: find the expected value of the characteristic sequence x_n, denoted:

$$\delta = (\sum_{i=1}^{n} x_i)/n . \quad (2)$$

Step 3: the standard deviation of the sequence is calculated by using the normal distribution

approximation method.

$$\sigma = k_\sigma * \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \delta)^2} \qquad (3)$$

k_σ is a factor associated with n the number of samples, when n is greater than 100 is approximately equal to 1.

Step 4: Traversal y = 1,2,3 ..., calculate k =y * σ, If the deviation of the sequence x_n is below K, then stop.

Step 5: record k=y*σ, as the characteristic variables of the network equipment's reference value.

Step 6: calculate the alarm threshold h, the use of the characteristics of the normal operation of the upper limit H, such as the temperature value of the normal work of the upper limit of 70℃.

h=y*(H-σ+k), y is the tolerance, the value following rules

$$y = \begin{cases} 1 & \texttt{Immediate alarm} \\ T/t & \texttt{Allow time T working and t is data acquisition time interval} \end{cases} \qquad (4)$$

Step 7: The normal characteristic sequence x_n is detected, the Zn value of x_n is calculated :

if x_n+Z_{n-1}-δ-k <0 *then* $Z_n=0$, else $Z_n=t_n+Z_{n-1}$-δ-k .

If $Z_n > h$, the h need to be increased, which makes the $Z_n < h$. The value h is the threshold of the characteristic sequence.
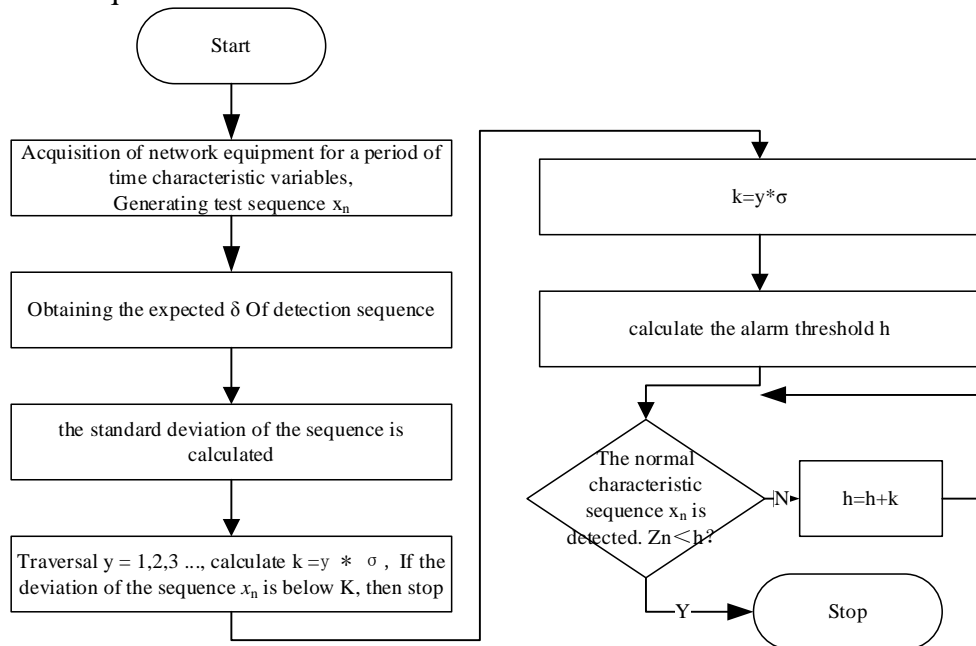

Fig1. Training algorithm flowchart

Through the training algorithm can find out the reference values and threshold values of the specific characteristics sequence, these two values in the detection algorithm has important applications, and they are not fixed, they needs to be adjusted according to the state of the system.

The fault detection algorithm is shown in Figure 2, the main steps are as follows:

Step 1: Get characteristic variable network equipment runtime, the main acquisition chip temperature, CPU usage, memory usage, packet loss rate interface, error rate and other parameters.

Step 2: The collected characteristic sequence x_n is detected, calculating CUSUM value Zn :

If $x_n + Z_{n-1}$ -δ-k <0，then $Z_n=0$ ，else $Z_n=x_n+Z_{n-1}$-δ-k。$Z_0=0$。

Step 3: If $Z_n> h$, then alarm and denote D $= Z_n$, $Z_n = h$, when $Z_n> h$, Z_n no increase is maintained

at h, this can eliminate the cumulative effects.

Step 4: If a false positive occurs, an appropriate increase in h, denote $h=h+k$, so that $D < h$.
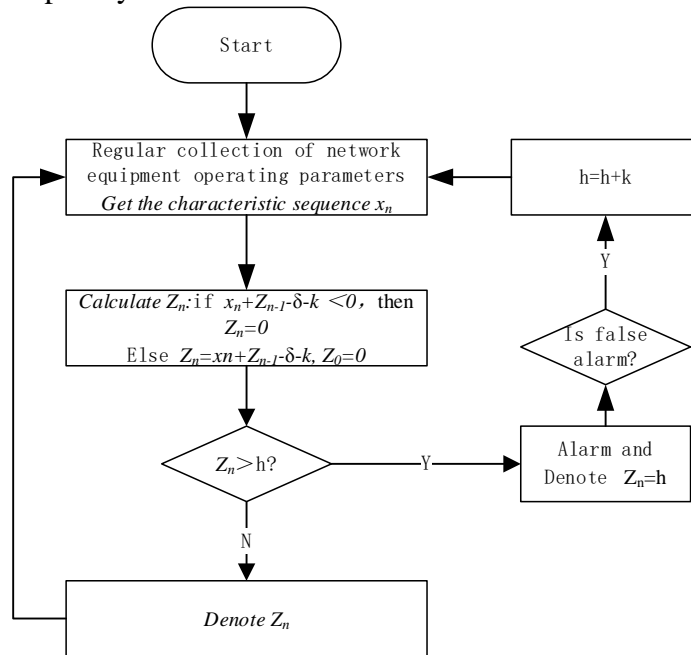
Step 5: Record Zn and report system status.

Fig2. Detection algorithm flowchart

System test

In this paper, the temperature of the board is collected, and the sequence deviation distribution is shown in Figure 3. As shown in Figure 3, when the system is stable, disturbance deviation of detection sequences are located within 2σ, therefore, k selected 2σ.
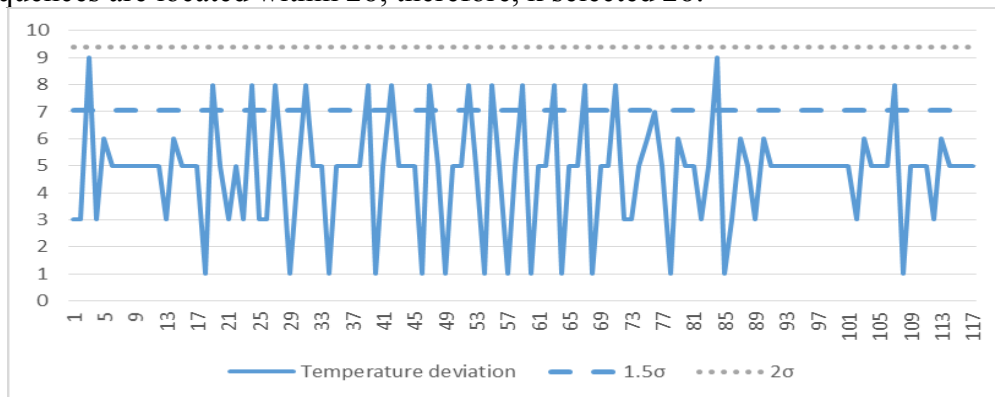

Fig. 3 Standard deviation distribution of temperature detection

Threshold h selection need to adjust according to the actual situation, the router board allows the high temperature threshold is 70℃, the average temperature is 35 ℃, k=9.4, tolerance is 3, then h=3*(70-35+9.4) =133.2, , h is adjusted through the test of the training sequence, in this paper the h is set to 150.

In summary, the parameters CUSUM anomaly detection algorithm uses the temperature detection selection as follows: h = 150, k = 2σ.

CPU usage anomaly detection parameter selection is similar to the temperature, as follows: h = 300, k = 4σ.

Based on the above parameters, this paper uses the denial of service tool to simulate attacks on routers, the experiment mainly collected the CPU usage and temperature. As shown in Figure 4, it can be seen that the use of CPU and the temperature are rising with the attack, the CUSUM value of CPU

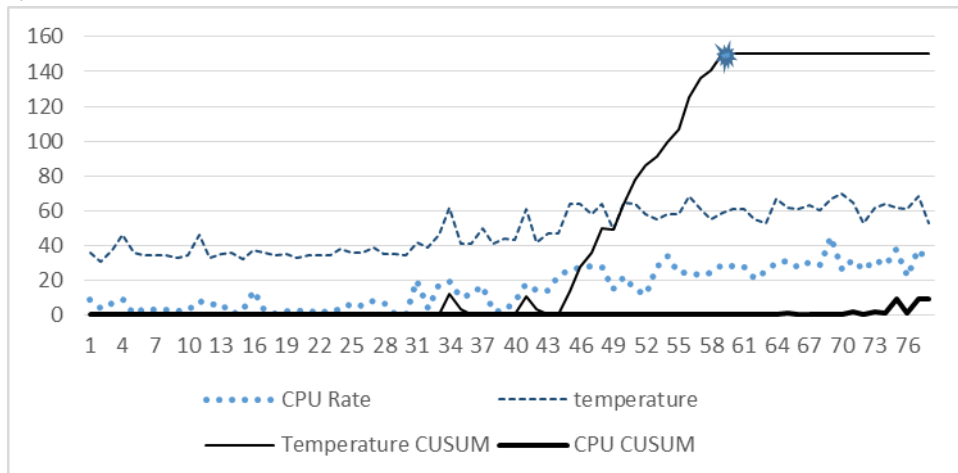is not to reach the alarm level, but the CUSUM value of the temperature reached the alarm level, it issued an alarm.

Fig. 4 The CUSUM algorithm used to detect abnormal of router

Conclusions

In this paper, the CUSUM algorithm is used to detect the health status of the network equipment. The experimental results show that the proposed algorithm can detect the abnormal system state in a short time, and provide early warning. In the follow-up study, we will analyze the detection performance of the algorithm, realize the adaptation of parameters, and try to apply it to other fault detection.

References

[1] Shengchao Su, Wei Zhang. Fault Prediction for Nonlinear System Using Sliding ARMA Combined with Online LS-SVR. Mathematical Problems in Engineering. Vol. 2014 p.1-9.

[2]Hou,Xiao-kai, Li Shiqian. Neural network fault prediction system in network equipment. Journal of shandong University of Technology,Nov 2014,p.29-33

[3]Z. X. SUN, Y. W. TANG, Y. CHENG. Router anomaly traffic detection based on modified CUSUM algorithms[J], Journal of Software, Vol.16(12)(2005), p. 2117-2123

[4] Han Dong, Tsung Fugee, Hu Xijian. A Multi-Chart Approach for Mean Shift Detection. Chinese

Journal of Applied Probability and Statistics.Vol. 24(3) (2008) ,p. 297-311.

[5] YU Ming, CHEN Wei-dong, ZHOU Xi-yuan. Adaptive Nonparametric CUSUM Control Chart. Computer Science. Vol.35(7) (2008) ,p 25-28.

[6]Kerdprasop, K., Kerdprasop, N.: Feature Selection and Boosting Techniques to Improve Fault Detection Accuracy in the Semiconductor Manufacturing Process. In: Int. Multi.Conf. Engineers and Computer Scientists 2011, pp. 398–403.

[7] Hong, S.J., Lim, W.Y., Cheong: Fault Detection and Classification in Plasma Etch Equipment for Semiconductor Manufacturing e-Diagnostics. IEEE Trans.Semi. Manufac. 25(2012),p. 83–93 .