

A Dynamic Adaptive Failure Detection Algorithm based on Grey System Theory

Daosheng Mu^{1, a}, Haoming Wang^{1, b} and Lijuan Gao^{1, c}

¹ Equipment Academy, Beijing 101400, China;

^amds19041@sina.com, ^bwanghaoming89@126.com, ^cmaicaipiaola@126.com

Keywords: Grey System, Adaptive, PULL, Failure detection.

Abstract. As an important part of disaster emergency response technology, failure detection technology is the foundation of the whole disaster emergency system. In this paper, an adaptive heartbeat detection mechanism GTFD based on grey system theory is studied, with the general armament department test message system as the background. The reliability and real-time of detection mechanism are improved by prediction of failure probability of occurrence and warning which based on a small amount of message on the system monitoring.

Introduction

The grey system [1-2] is an important theory to solve problem with uncertain factors, such as incomplete parameters, imperfect construction, incomplete factor message, which can obtain the knowledge and understanding of all the message through incomplete and uncertain message. There are problems of a small amount of monitoring message and incomplete parameters in the failure detection determining process of the general armament department test message system, which accord with the "little data uncertainty" application background of the grey system theory. In this paper, in order to solve the problems complex computation, high time cost to acquire accuracy[3], GTFD (Grey Trends Failure Detection) failure detection method is proposed combined with the grey system theory, which can be the supplement of the traditional method. With certain intelligence, the proposed method can make the prediction more practical.

GTFD algorithm

GTFD algorithm use in PULL mode heartbeat detection method[4] which has higher reliability and pertinence, applying the grey system theory to the detection process. Its core idea is to use a differential equation represented by a grey prediction model of a degree, then calculate time grey value of the next received message through all the previous heartbeat detection times, and finally determine whether the node has failed through the actual arrival time.

The implementation process is mainly divided into four stages: heartbeat detection, grey dynamic prediction model establishment, prediction model test, prediction and detection phase, as shown in figure 1. Also the Crash-stop model is described as follows with the two nodes a,b as the master and slave node.

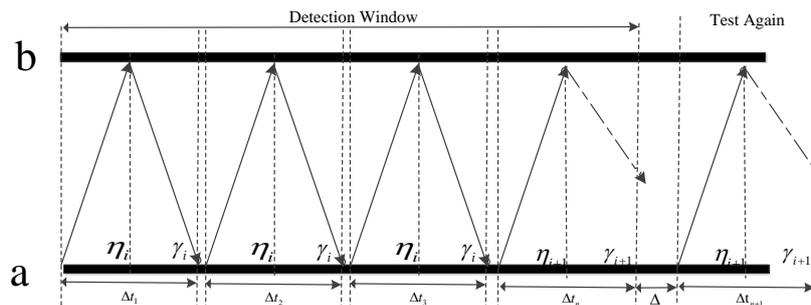


Fig.1 GTFD algorithm

Heartbeat detection phase: Periodically send the PULL heartbeat signal from a to b in timing system, and all the heartbeat signal will be labeled, also the feedback time will be recorded. Select n consecutive heartbeat feedback time lengths as the reference sequence $t^{(0)}$ represented by the array as: $t^{(0)} = (t^{(0)}(1), t^{(0)}(2), t^{(0)}(3), \dots, t^{(0)}(n))$. During the whole testing process, The n feedback message is needed to be recorded constantly and to be processed by AGO, then array $t^{(1)}$ is generated, whose expression is $t^{(0)} = (t^{(0)}(1), t^{(0)}(2), t^{(0)}(3), \dots, t^{(0)}(n))$.

Grey dynamic prediction model establishment: The major process of the grey prediction is the establishment of GM(1.1) prediction model[5-7], namely, to establish a first-order differential equation for the array $t^{(1)}$.

$$\frac{dt^{(1)}(k)}{dk} + at^{(1)}(k) = b \quad (1)$$

By the least square method, using matrix B and vector $T_{(n)}$:

$$[\hat{a}, \hat{b}]^T = [B^T, B]^{-1} B^T T_{(n)} \quad (2)$$

Here, matrix B

$$B = \begin{bmatrix} -\frac{1}{2}(t^{(1)}(1) + t^{(2)}(2)), 1 \\ -\frac{1}{2}(t^{(1)}(1) + t^{(1)}(3)), 1 \\ \vdots \\ -\frac{1}{2}(t^{(1)}(n) + t^{(1)}(n)), 1 \end{bmatrix} \quad (3)$$

$$T_{(n)} = [t^{(0)}(2), t^{(0)}(3), t^{(0)}(4), \dots, t^{(0)}(n)]^T \quad (4)$$

$\hat{t}^{(1)}$ can be obtained by equation(3-9). Assume that $\hat{t}^{(0)}$ is the proper prediction array, then:

$$\hat{t}^{(0)} = (\hat{t}^{(0)}(1), \hat{t}^{(0)}(2), \hat{t}^{(0)}(3), \dots, \hat{t}^{(0)}(n)) \quad (5)$$

Here, $\hat{t}^{(0)} = t^{(0)}(1)$, process $\hat{t}^{(1)}$ by AGO, then

$$\hat{t}^{(0)}(k) = \left(\hat{t}^{(0)}(1) - \frac{\hat{a}}{\hat{b}} \right) \times (1 - e^{\hat{a}}) \times e^{\hat{a}(k-1)} \quad (6)$$

$\hat{t}^{(0)}(n+1)$ is the grey value of arrival time of the next message.

Prediction model test: Before the above algorithm is executed, the grey model GM(1.1) should be tested using the posterior variance method. First we define the residual sequence $\varepsilon^{(0)} = \{\varepsilon^{(0)}(k) = t^{(1)}(k) - \hat{t}^{(1)}(k) \mid K = 1, 2, 3, \dots, n\}$, then calculate variances S_1^2 and S_2^2 of the initial sequence $t^{(0)}$ and the residual sequence $\varepsilon^{(0)}$.

$$S_1^2 = \frac{1}{n} \sum_{k=1}^n (t^{(0)}(k) - \bar{t}^{(0)})^2, \quad S_2^2 = \frac{1}{n} \sum_{k=1}^n (\varepsilon^{(0)}(k) - \bar{\varepsilon}^{(0)})^2 \quad (7)$$

Calculate the minimum error probability $p = P\{|\varepsilon^{(0)}(k) - \bar{\varepsilon}^{(0)}| < 0.6745S_1\}$ and the experimental contrast $C = S_2 / S_1$. $0.95 \leq p$. Precision detection model is divided as shown in Table 1.

Table 1 Detection model prediction

| degree index | excellent | good | medium | poor |
|-----------------|---------------|----------------------|----------------------|------------|
| P | $0.95 \leq p$ | $0.80 \leq p < 0.95$ | $0.70 \leq p < 0.80$ | $p < 0.70$ |
| C | $c \leq 0.35$ | $0.35 < c \leq 0.5$ | $0.5 < c \leq 0.65$ | $0.65 < c$ |

GTFD algorithm analysis

Dynamic and adaptivity. After running for a certain time, the system will be changing because of various internal factors. But simple detection model GM(1.1) can not reflect the changes in the system. In order to avoid the credibility of detection mode to reduce as time goes, we will change the basis of prediction according to the change of the arrival time of the heartbeat detection information.

The prediction result of using GM(1.1) model for long term is affected by the system cycle list size and data changes. If the heartbeat arrival time list is too short, it will make it difficult to establish long-term prediction model. However, if too long, it will make the system unstable, thus reducing the prediction accuracy. Therefore, a fixed window size method is adopted in prediction. Set that heartbeat arrival time T_n is the size of the table window size, setting T_n as the original class table window. Dynamically slide and change the list over time. The process is as follows:

Set that the initial sequence $t^{(0)} = (t^{(0)}(1), t^{(0)}(2), t^{(0)}(3), \dots, t^{(0)}(n))$. Establish model GM(1.1) after AGO. Obtain the n+1 prediction value by (3-9). After getting the n+1th actual heartbeat message, make an improvement of the GM (1.1) model. That is to delete $t^{(0)}(1)$ in the sequence $t^{(0)} = (t^{(0)}(1), t^{(0)}(2), t^{(0)}(3), \dots, t^{(0)}(n))$, and add $t^{(0)}(n+1)$ to make a new sequence $t_{(1)}^{(0)} = (t^{(0)}(2), t^{(0)}(3), t^{(0)}(4), \dots, t^{(0)}(n+1))$, where T_n is the window size. So repeatedly, ensure that the data used is updating in the non-stop, and reflect the latest state of the reaction system. Improve the prediction accuracy in the premise, the algorithm not only has the dynamic data updating, adaptation also has forecast the results change with the change of system.

Integrity performance. During the heartbeat detection procedure, the right process will receive the message b sends in a moment of the future t_i . It has a upper limit value. So in the cases without considering the detection algorithm, the data and parameters which are based on are determined. The predicted values obtained this way have limits. And the next detection is to add a PULL detection while the process a had not receive the process heartbeat messages from process b . This will only increase the detection time, and won't change boundedness of detection time. Therefore, we can conclude that there is a moment t_n , when process a determine that process b fails. Here the moment t_n is limited. The formal specification is: $\exists t_n : \forall t \geq t_n, \forall a \in correct(t), \forall b \in failed, b \in suspect_a(t)$. Thus, the algorithm has strong integrity.

Strong accuracy. The algorithm judges whether the process fails, depending on whether the heartbeat messages arrival time is longer than the time the algorithm predicts. In the next detection process, the time next heartbeat message arrives is obtained by the dynamically changing heartbeat message obtained. The detection time increases along with the detection times. In the end t_i , the return message process sent by b will be received by a during the detection. The formal specification is: $\exists t_i : \forall t \geq t_i, \forall a, b \in correct(t), b \notin suspect_a(t)$. Thus, the algorithm has strong accuracy.

GTFD algorithm test

Prediction accuracy. Same as the algorithm in chapter 3.2, GTFD uses the established server cluster as the simulation test platform. The difference is that GTFD adopts the method of sending PULL detection information from mater server to the slave server. When the platform goes stable, 6 pull continuous feedback delay from the unified time system continuous feedback delay as the raw data, shown as table 2.

Tab. 2 Original data time

| sequence (n) | 1 | 2 | 3 | 4 | 5 | 6 |
|----------------|-------|-------|-------|--------|-------|-------|
| time (s) | 1.050 | 1.036 | 1.020 | 1. 002 | 0.982 | 0.960 |

It can be obtained through formula 2 and 3 that the original sequence $\hat{t}_1^{(0)}(n+1) = 1.05814e^{-0.019n}$. In order to test the model accuracy, we use residual method to calculate error. Sequence 2~6 is the

analog values, while 7~8 is the predicted value. The error is shown in table 3.

Tab.3 Error

| Sequence (n) | Original data $x_1^{(0)}(n)$ | Analog data $\hat{x}_1^{(0)}(n)$ | Residual $\varepsilon_1(n) = x_1^{(0)}(n) - \hat{x}_1^{(0)}(n)$ | Relative error $\Delta_1 = \frac{ \varepsilon_1(n) }{x_1^{(0)}(n)}$ |
|-------------------|---------------------------------|-------------------------------------|--|--|
| 2 | 1.036 | 1.0382 | -0.0022 | 0.212% |
| 3 | 1.020 | 1.0187 | 0.0013 | 0.127% |
| 4 | 1.002 | 0.9995 | 0.0025 | 0.250% |
| 5 | 0.892 | 0.9807 | 0.0013 | 0.132% |
| 6 | 0.960 | 0.9622 | -0.0022 | 0.229% |
| 7 | 0.936 | 0.9441 | -0.0081 | 0.865% |
| 8 | 0.910 | 0.9264 | -0.0164 | 1.802% |

It can be concluded by analyzing the above statistics that the accuracy of GTFD is “Excellent”, but when the predicted number is greater than a certain number, the prediction accuracy will decrease, which proves that GTFD has high accuracy in short-term prediction process.

Let the first prediction gray value be the original data to start the next prediction, and calculate the new error class table, shown as table 4.

Tab.4 Next Prediction Error

| Sequence (n) | Original data $x_1^{(0)}(n)$ | Analog data $\hat{x}_1^{(0)}(n)$ | Residual $\varepsilon_1(n) = x_1^{(0)}(n) - \hat{x}_1^{(0)}(n)$ | Relative error $\Delta_1 = \frac{ \varepsilon_1(n) }{x_1^{(0)}(n)}$ |
|-------------------|---------------------------------|-------------------------------------|--|--|
| 3 | 1.020 | 1.0191 | 0.0009 | 0.088% |
| 4 | 1.002 | 1.0009 | 0.0011 | 0.110% |
| 5 | 0.892 | 0.9812 | 0.0008 | 0.081% |
| 6 | 0.960 | 0.9618 | -0.0018 | 0.186% |
| 7 | 0.936 | 0.9434 | -0.0074 | 0.791% |
| 8 | 0.910 | 0.9248 | -0.0148 | 1.626% |
| 9 | 0.902 | 0.9345 | -0.0168 | 1.826% |

Prediction error obtained by introducing the reprediction method is small, which improves the monitoring accuracy.

Adaptive performance test. In the test, when PULL mode is started, Algorithm in automatic adjust the timeout process is shown in figure 2. The vertical axis represents the timeout the system calculates when the corresponding message arrives, the abscissa represents information number. The horizontal axis represents the information number. Compared with traditional algorithm Hayashibara under the same circumstances, the adaptive performance of GTFD is better.

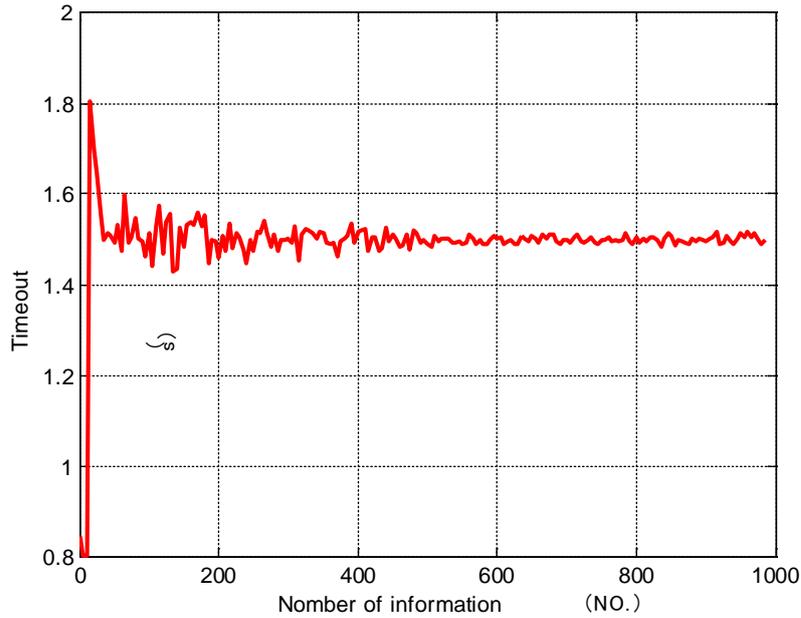


Fig.2 Automatically adjusting the timeout interval simulation

Influence of the window size. With server clusters as a simulation platform, within the 10-hour test period, the window size value is from 10 to 200, whose influence on the average detection time and error rate is tested. Figure 3 shows the average detection time comparison as the window size values vary from 10 to 200. It can be seen that along with the increase of the number of window, the average detection time have some fluctuations, but the overall rule is to increase with the number of windows. Figure 4 shows the influence of window size on the error rate. Overall, the lager the window size, the less the error rate is.

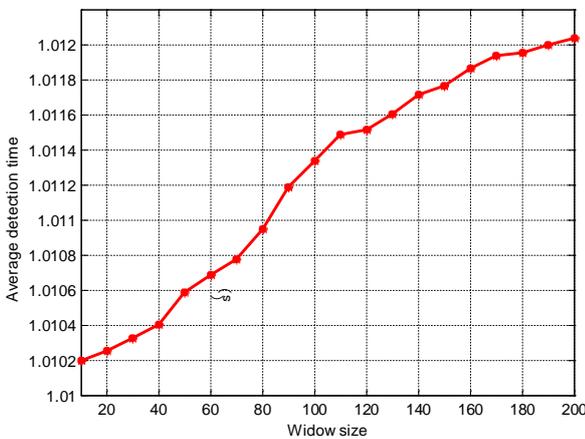


Fig.3 Influence of window size on average detection time

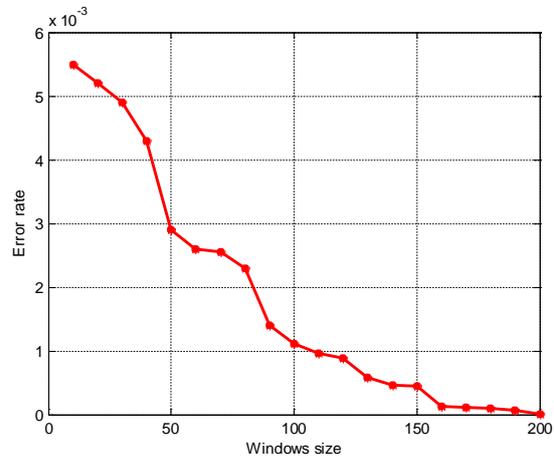


Fig.4 Influence of window size on the error rate

Performance comparison. In order to reflect the algorithm performance better, we compare the GTFD with Chen under the same circumstances. Set the window size as 200, heartbeat cycle 1s. In the Chen algorithm, set the security boundary $\alpha = 0.005$. The test results are shown in table 5.

Tab.5 Comparison of GTFD and Chen

| | GTFD algorithm | Chen algorithm |
|------------------------|----------------|----------------|
| Window size | 200 | 200 |
| Average detection time | 1.010841 | 1.010838 |
| Error rate | 0.002214 | 0.002242 |

The table above shows that: When the window size is 200 and $\alpha = 0.005$ in Chen, in the case when average detection time is approximately the same, the error rate of GTFD is smaller than Chen, thus improving the accuracy of failure detection.

Conclusion

In this paper, we proposed grey dynamic adaptive detection algorithm GTFD by improving traditional grey prediction model GM (1, 1), and proved its dynamic adaptability, strong integrity and strong accuracy through theory analysis and test. Then, through the experiment on algorithm accuracy, adaptive ability, the window size and the detection time, accuracy, the following conclusions were drawn: GTFD algorithm has stronger predictive ability in the case where data is less; It has the adaptive ability to meet failure detection on different business of information system, and the adaptive speed compared to the Hayashibara algorithm is improved in a certain degree. The window size is proportion to the detection time. Finally, under the same network conditions, compared with and the classical Chen algorithm, GTFD is improved in both accuracy and speed in a certain degree.

References

- [1]. Deng J L. Introduction to grey system theory [J]. The Journal of grey system, 1989, 1(1): 1-24.
- [2]. Liu S, Lin Y, Forrest J Y L. Grey systems: theory and applications [M]. Springer, 2010.
- [3]. Yang Z, Li C, Liu Y, et al. An Exponential Smoothing Adaptive Failure Detector in the Dual Model of Heartbeat and Interaction [J]. Journal of Computing Science and Engineering, 2014, 8(1): 17-24.
- [4]. Wei C, Kai S, Sen S, et al. A heuristic chunk scheduling algorithm reducing delay for mesh-pull P2P live streaming[C]. Broadband Network & Multimedia Technology, 2009. IC-BNMT'09. 2nd IEEE International Conference on. IEEE, 2009: 706-710.
- [5]. Wen J, Xiao X. Grey multi-variables GM (1, N/ τ , r) model with delay and its application[C]. Grey Systems and Intelligent Services, 2009. GSIS 2009. IEEE International Conference on. IEEE, 2009: 357-361.
- [6]. Zhao N, Yang J, Li S, et al. A GM (1, 1) Markov Chain-Based Aeroengine Performance Degradation Forecast Approach Using Exhaust Gas Temperature [J]. Mathematical Problems in Engineering, 2014, 2014.
- [7]. Liu S, Lin Y, Forrest J Y L. Grey systems: theory and applications [M]. Springer, 2010.