

## Research on the Method of Aeroengine Fault Diagnosis based on Immune Genetic Algorithm

YanJun Li <sup>a</sup>, Jian Zhang <sup>b</sup>, Lina Zhang <sup>c</sup>, Zhengqiang Cheng <sup>d</sup>

College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

<sup>a</sup>email: lyj@nuaa.edu.cn, <sup>b</sup>email: 1540079310@qq.com,

<sup>c</sup>email: 747931881@qq.com, <sup>d</sup>email: 269102349@qq.com

**Keywords:** Aeroengine; Fault Diagnosis; Immune Genetic Algorithm; Gas Path Parameter

**Abstract.** In this paper, an immune genetic algorithm based on the performance parameters of aeroengine and the adaptive, learning and memory characteristics of artificial immune system (AIS) was put forward to diagnose the fault of aeroengine. Firstly, we generate the initial detector according to the fault samples, and optimize the matching relationship to improve the representation and coverage. Then train the detector with the genetic algorithm to learn and remember the fault information. Finally, we identify the faults of aeroengine by the mature detector. The results show that the method can accurately identify the engine fault by using the parameters of the aeroengine gas path provided by Honeywell company.

### Introduction

The great development of the aviation industry results in more attention being given to safety and more attention, and the security of the aviation engine has become a trending research topic. Aviation engine works in high temperatures, at high speed and other harsh working environment, and its structure being rather complex can sometimes be prone to failure. The diagnosis of engine failure can help to improve the reliability of the engine, ensure flight safety, reduce the next time out, and can reduce the ground test resource consumption, increase cost-effective use of spare parts, realize the maintenance of the visual conditions, and improve the operational efficiency.

Aeroengine fault can be diagnosed through the engine gas path parameter, the main methods are neural network [1][2], expert system [3][4], support vector machine [5][6], etc. However, these methods all have shortcomings, for example neural networks have the problems of the selection of network structure and weight algorithm, the hidden layer has no strict theoretical basis and is easy to fall into local optima. The expert system has the problems of incomplete knowledge and is difficulty in obtaining its structure coupled with poor adaptability, etc. The support vector machine has the issues regarding kernel function selection and the solution of quadratic programming function, etc.

Immune genetic algorithm (IGA) is an optimization algorithm that simulates natural selection and evolutionary process, which can track the change of object in a timely fashion. In addition to the advantages of traditional genetic algorithm, IGA can also effectively inhibit the "premature", improve the search speed and the quality of optimization [7]. So this paper uses immune genetic algorithm to accomplish the fault diagnosis of aircraft engine. Firstly, define the relevant parameters of the immune genetic algorithm, then generate initial detector randomly to calculate the matching degree between the sample and the fault and suppress or promote the detector according to the calculation results. Next, use genetic algorithm crossover, mutation detector, and update detector according to a specific strategy. Finally, we make use of fault diagnosis if it is mature, otherwise go to the matching degree calculating, and so forth, until the detector matures.

### Method of Aeroengine Fault Diagnosis based on Immune Genetic Algorithm

#### Sample vector of performance parameter

The corresponding performance parameters will be abnormal when the aero engine failure

occurs, so the fault can be diagnosed by analyzing the performance parameters of the engine. In this paper, the  $N$  key performance parameters of the engine condition are selected as the basis of fault diagnosis and a set of state samples is defined as an  $N$  dimension vector. It is not feasible to directly use the original data to analyze as the unit level of each parameter is different. Therefore, first of all, the data needs to be normalized.

### Method of affinity calculation

Supposed  $J = [j_1, j_2, \dots, j_n]$  and  $K = [k_1, k_2, \dots, k_n]$  are all normalized vectors, in  $N$ -dimensional state space and the affinity between  $J$  and  $K$  is

$$D = \frac{1}{1 + d_E} \quad (1)$$

In the formula, the  $d_E$  is the Euclidean distance between  $J$  and  $K$ , that is

$$d_E = \sqrt{\sum_{i=1}^n (j_i - k_i)^2}, i = 1, 2, \dots, n \quad (2)$$

This method of calculating the affinity makes the affinity in the  $[0, 1]$  range, and the smaller the distance, the greater the affinity, the more intuitive it is.

### Generate initial detectors

Generate a certain number of  $N$  dimensional normalized initial detectors according to the performance parameter samples under the fault condition and define as vector  $x_i^0$ . Initial detectors should meet two requirements: Firstly, the detectors cannot match with the normal sample, so the affinity of the two should be less than the threshold  $r$ ; The second is that the detector distribution should cover the entire range of fault sample, and the more uniform it is the more clearly can it be represented, so the initial detector and the matching threshold between the generated detector should be less than  $q$ . For those detectors that cannot meet the two conditions, they need to be deleted and re-generated.

### Selection, crossover, mutation and update of detectors

The variation and update of the detectors are aimed so as to enable them to identify the fault. The initial detectors are generated according to the normal samples and cannot recognize the fault samples. So the detector should learn the characteristics of the fault samples and form memory of the fault characteristics. Produce new  $N$  individuals by operation of selection, crossover and mutation to genetic population, and the algorithms are as follows:

A. Selection. In this paper, based on the ratio selection method, select these detectors whose affinity with the antigen is greater than the immune selection threshold  $T_i$  for crossover and mutation operation.

B. Crossover. This paper uses the floating-point number encoding, so the following algorithm is used to cross:

$$\begin{aligned} x_i^{t+1} &= \alpha x_j^t + (1 - \alpha) x_i^t \\ x_j^{t+1} &= \alpha x_i^t + (1 - \alpha) x_j^t \end{aligned} \quad (3)$$

In the formula,  $x_i^{t+1}$  and  $x_j^{t+1}$  are the crossover individual,  $x_i^t$  and  $x_j^t$  are a randomly selected,  $\alpha$  is a constant crossover value, the range is  $(0, 1]$ .

C. Mutation. The algorithm is as follows:

$$x_i^{t+1} = \begin{cases} x_i^t + k \cdot (x_{\max} - x_i^t) \cdot r, \text{rand}() \% 2 = 0 \\ x_i^t - k \cdot (x_i^t - x_{\min}) \cdot r, \text{rand}() \% 2 = 1 \end{cases} \quad (4)$$

In the formula,  $x_i^{t+1}$  is the individual after mutation,  $x_i^t$  is the individual before mutation,  $k$  is a constant value of mutation, in the range of  $(0, 1]$ ,  $x_{\max}$  is the upper limit of the individual,  $x_{\min}$  is the lower limit of the individual,  $r$  is randomly generated random number.

### Mature judgment

If the affinity of a certain type of antibody and antigen is greater than diagnostic threshold, we

think the antibody can effectively recognize antigen.

If all antigens can be recognized by antibodies in the memory, that means the termination conditions are met, the antibodies are mature and the algorithm ends.

## Example of Aeroengine Diagnosis based on Performance Parameters

### Sample data preprocessing

In this paper, we verify the accuracy of the immune system for the fault detection of aeroengine with samples of a certain type of aeroengine provided by Honeywell. The unit of performance parameters is different, so we normalize the original data in order to facilitate observation. The algorithm is as follows:

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

In the formula,  $x'_i$  is the sample data after normalized,  $x_i$  is the initial sample data,  $x_{\max}$ 、 $x_{\min}$  are respectively the maximum and minimum value of the feature vector  $\{x_i | i = 1, 2, \dots, n\}$ .

Select 10 group of samples in normal conditions provided by Honeywell, it is shown in table 1 after being normalized. Select 10 group of samples in fault condition, it is shown in table 2 after being normalized. Two sets of data are respectively used as the normal and fault samples.

Table 1 Normal Samples

No.	EGT	N <sub>2</sub>	No.	EGT	N <sub>2</sub>
1	0.550	0.720	6	0.490	0.690
2	0.525	0.701	7	0.51	0.696
3	0.535	0.724	8	0.54	0.697
4	0.505	0.709	9	0.490	0.703
5	0.490	0.697	10	0.530	0.701

Table 2 Fault Samples

No.	EGT	N <sub>2</sub>	No.	EGT	N <sub>2</sub>
1	0.155	0.251	6	0.730	0.720
2	0.770	0.726	7	0.750	0.724
3	0.270	0.653	8	0.275	0.649
4	0.740	0.721	9	0.29	0.652
5	0.155	0.335	10	0.285	0.651

### Selection of parameter

From the above analysis, we know that several parameters need to be defined in advance, which play a key role in the diagnosis speed and quality of the algorithm. Several important parameters are analyzed as follows:

1) Number of initial detector  $n$ . The initial detector should take the operation speed and coverage rate into account, and its value is related to the dimension  $N$  of the input sample and the matching threshold  $q$  of the initial detector. The samples analyzed in this paper are two-dimensional vectors, so  $N=2$ ,  $q=0.95$ ,  $n=400$ .

2) Suppression threshold  $\xi$ . From the analysis above, we can see that the suppression threshold is aimed to reduce the redundancy of detectors in memory, the greater the value, the higher the similarity between detectors which is then acceptable. However if the value is too small, more detectors will be removed from memory, which will easily lead to undetected fault. Therefore, in order to ensure the accuracy of the algorithm, we choose a large suppression threshold as  $\xi=0.98$  by sample spreadsheet and experience.

3) Diagnostic threshold  $\lambda$ . Diagnostic threshold is a sample basis for fault diagnosis, Getting too large in magnitude  $\lambda$  can easily cause the omission, fault samples as normal samples; and when too low, become prone to error detection considering normal samples as fault samples. According to this example, it is found that the best value is about 0.92, so it may take  $\lambda=0.92$ .

### Diagnosis examples

According to the above analysis, the relevant parameters involved in the algorithm are as follows: number of initial detector  $N=400$ ; matching threshold of initial detector  $q=0.80$ ,  $r=0.85$ ; the number of detector with high affinity  $m=20$ ; mutation rate  $\alpha=0.5$ ; memory warehouse selecting threshold  $\delta=0.90$ ; suppression threshold  $\xi=0.98$ ; diagnostic threshold  $\lambda=0.92$ . Figure 1 shows the initial detector space distribution generated by the proposed algorithm. Figure 2 shows the mature detector distribution of memory warehouse.

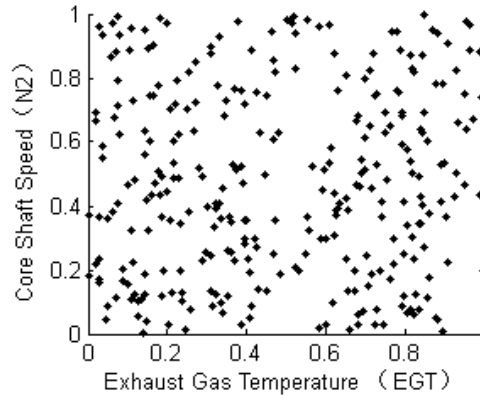


Figure 1 Initial detectors

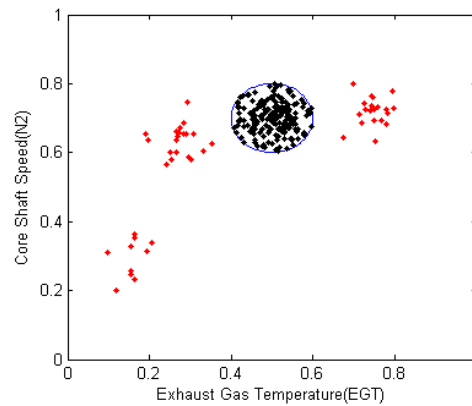


Figure 2. Mature detectors

Choose 10 examples of the same engine as tested samples and calculate the affinity with the mature detector in the memory warehouse respectively. We consider the sample as a fault sample if the affinity with a detector is greater than the diagnostic threshold. 10 sets of data and its analysis results are shown in table 3, all the diagnosis results are correct, the diagnosis results are shown in Figure 3.

Table 3 Test Results

No.	EGT	N <sub>2</sub>	Actual state	Diagnosis results
1	0.185	0.285	Fault	Fault
2	0.495	0.682	Normal	Normal
3	0.515	0.701	Normal	Normal
4	0.295	0.654	Fault	Fault
5	0.52	0.697	Normal	Normal
6	0.485	0.711	Normal	Normal
7	0.755	0.763	Fault	Fault
8	0.49	0.694	Normal	Normal
9	0.535	0.707	Normal	Normal
10	0.265	0.613	Fault	Fault

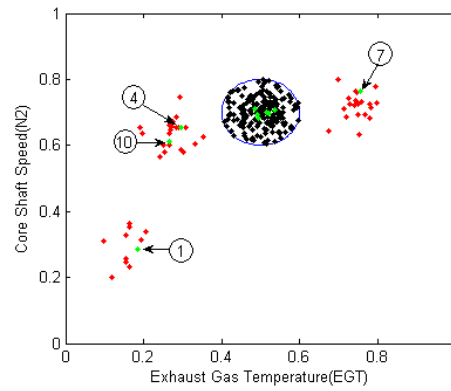


Figure 3 Test Result

## Conclusion

Performance parameter of engine can be an objective response to the state of the engine, the analysis plays an important role towards engine fault diagnosis. The method of engine fault diagnosis proposed in this paper is based on performance parameters and immune genetic algorithm. The method uses the functions of Immune Theory such as memory, learning and identify to dig hidden information of engine state in performance parameters and effectively identify the type of engine failure with the advantages of fast discrimination and strong adaptability.

## Acknowledgement

In this paper, the research was sponsored by Aeronautical Science Foundation of China (Project No. 201510301007131) and Nanjing University of Aeronautics and Astronautics School Open Fund (Project No. kfjj20150701).

## References

- [1] Tian Chen , Jianguo Sun, Weihua Yang, et al. Self-organizing neural network based fault diagnosis for aero engine gas path [J]. Aeronautical journal. 2003 (01) :46-48.
- [2] Chen T, Sun J G, Yang W H, et al. Self-organizing neural network based fault diagnosis for aero engine gas path [J]. Aeronautical journal. 2003(01):46-48.
- [3] Guo Chen, Hongfu Zuo. Expert systems of engine wear fault diagnosis based on knowledge rule [J]. Journal of Aerospace Power, 2004, 19(1):23-29.
- [4] Kehong Lv, Yue Li. Study on Fault Diagnosis Expert system for engine based on ferrography information[J]. Transactions of CSICE, 2003, 21(6):453-457.
- [5] Caoying Sun, Lu Liu, Chuanwu Liu, etc. Aero-engine fault diagnosis based on Boosting-SVM [J]. Journal of Aerospace Power, 2010, 25(11): 2584-2588.
- [6] Zhousuo Zhang, Lingjun Li, Zhengjia He. Research on Diagnosis Method of Machinery Fault Based on Support Vector Machine[J]. JOURNAL OF XI'AN JIAOTONG UNIVERSITY, 2002, 36(12): 1303-1306.
- [7] Ruiming Wang, Yujin Zeng , Jingping Jiang. Application of immune genetic algorithm based fuzzy neural network in AC servo system [J]. Journal of Zhejiang University, 2005, 39 (8):1156-1160.