Research on Comparison of Different Algorithms in Diagnosing Faults of Aircraft Engines

Liao Li^{1,}*🕩

1. COMAC Shanghai Aircraft Customer Service Co., Ltd – Shanghai – China. *Correspondence author: qinliao3140358@163.com

ABSTRACT

For the aircraft, the engine is its core component. Once the engine fails, the flight safety will be seriously affected; therefore, it is necessary to diagnose the failure in time. This paper briefly introduced three aircraft engine fault diagnosis algorithms based on support vector machine (SVM), random forest, and particle swarm optimization-back-propagation (PSO-BP) and carried out a simulation experiment on the performance of the three algorithms in MATLAB software. The results showed that the PSO-BP-based diagnosis algorithm had the highest recognition accuracy and the SVM-based diagnosis algorithm had the lowest, both for artificial fault data and real fault data. The PSO-BP-based diagnosis algorithm took the least average recognition time, and the SVM-based diagnosis algorithm took the longest time.

Keywords: Aircraft engine; Fault diagnosis; Random forest; Particle swarm optimization-back-propagation; Air flow; Fuel-air ratio.

INTRODUCTION

Since its invention, the aircraft, a flying device, has become more and more sophisticated, intensive, intelligent and efficient in its overall structure with the progress of technology (White and Karimoddini 2020). It can be said that a large number of advanced modern technologies have been applied in modern aircraft. The application of these technologies allows aircraft to fly faster and higher and to provide better service to users. However, the precise structure and its use environment also make the safety and reliability of aircraft subject to more stringent challenges (Li 2020). The working environment of aircraft, especially passenger aircraft, is generally in the stratosphere at high altitude. In case of failure, especially engine failure, flight safety will definitely be affected. In mild cases, the flight may make an emergency landing, which will affect operational efficiency; in severe cases, the flight may stop or explode, causing loss of life and property (Ali *et al.* 2010). Therefore, accurate fault diagnosis of an aircraft engine is an important means to maintain safe aircraft operation. Zhao *et al.* (2020) proposed an improved extended least square support vector machine (e-LSSVM), and verified the effectiveness of the method in aero-engine fault diagnosis method for aero-engine wear faults and verified the effectiveness of the method by experiments. Zhao *et al.* (2017) proposed a flight state partition-based sensor fault diagnosis method to improve the fault diagnosis accuracy of aero-engine and verified the effectiveness of the method by experiments.

Received: Apr. 24, 2021 | Accepted: Aug. 02, 2021 Peer Review History: Single Blind Peer Review. Section Editor: Dimitrios Pavlou

This is an open access article distributed under the terms of the Creative Commons license

through simulation. This paper briefly introduced three aircraft engine fault diagnosis algorithms based on support vector machine (SVM), random forest, and particle swarm optimization-back-propagation (PSO-BP) and carried out a simulation experiment on the performance of the three algorithms in MATLAB software. The main goal of this paper is to rapidly detect aircraft engine faults with an intelligent algorithm. The basic process is to collect relevant data that can reflect the engine condition and then use intelligent algorithms to identify the collected data. This paper proposes the PSO-BP algorithm to identify the engine data and compares it with SVM and random forest algorithms. The final results verified the effectiveness of the PSO-BP algorithm for diagnosis of engine faults.

BRIEF INTRODUCTION OF AIRCRAFT ENGINE AND FAULT DIAGNOSIS

The original aircraft was relatively simple in structure. Its engine used the expansion thrust generated by the rapid combustion of fossil fuels in a confined space, which drives the piston to produce reciprocating motion, and the reciprocating motion power drove the propeller through transmission devices such as belt and gear to provide lift airflow to realize flight (Furse *et al.* 2021). As the demand for aircraft flight speed increases, turbojet engine appears, and its structure is shown in Fig. 1, including inlet, compressor, combustion chamber, turbine, and nozzle. The turbine is the center of gravity of the whole jet engine. The structure of the turbine mainly includes a rotating shaft, a rotor, and working blades connected to the rotor (Yang *et al.* 2016). The performance of aircraft engines has been greatly improved with the progress of electronic and material technology, but it also makes the structure more complicated and maintenance more difficult.



Figure 1. A schematic diagram of aircraft engine structure.

The traditional engine diagnosis method is regular manual maintenance, and different parts are exchanged and maintained according to the designed working life. This method is simple and regular; however, it is impossible to avoid sudden failures only by regular maintenance, and differences between different engine bodies and flight environments make the occurrence of failures more uncertain (Ma *et al.* 2015). Monitoring and diagnosis need to be realized by collecting data from sensors; the intensive and intelligent aircraft engine makes sensors collect massive data. An appropriate fault diagnosis algorithm is needed in order to diagnose faults quickly.

FAULT DIAGNOSIS ALGORITHM

Support vector machine-based diagnosis algorithm

The diagnosis algorithm that can be applied to aircraft fault detection is essentially a classification algorithm. It takes data in the samples to be detected as the recognition features and divides the samples into corresponding categories, i.e., fault types.

Firstly, the SVM model (Sevil and Dogan 2015) is trained with the training sample set to calculate the "hyperplane" of the sample in the vector space. The optimization problem is as follows (Eqs. 1 and 2):

$$\min\frac{1}{2}\omega^T\omega + \mathcal{C}(\sum\delta_i) \tag{1}$$

Condition: $y_i[(\omega^T x_i) + b] - 1 + \delta_i \ge 0$ $\delta_i \ge 0$ (2)

where x_i is the feature vector, y_i is the category label of sample *i*, *w* is the weight, *b* is the bias term, *C* is the penalty parameter, and δ_i is the non-negative relaxation variable of sample *i* (Loza *et al.* 2015). When solving the optimization problem (Eq. 1), the problem is transformed into a dual universal function by constructing a Lagrange function, and the decision function is obtained after solution (Eq. 3):

$$\begin{cases} f(x) = sgn(\sum_{i=1}^{l} a_i y_i K(x_i, x_j) + b) \\ \sum_{i=1}^{l} a_i y_i = 0 \quad 0 \le a_i \le C \end{cases}$$
(3)

where α_i is the Lagrange coefficient (Kiran *et al.* 2018), *l* is the sample size, and $K(x_i, x_j)$ is a kernel function. When it is applied to aircraft engine fault diagnosis, a decision function is obtained by carrying out fitting training on the engine fault training samples. Finally, the engine fault samples to be detected are classified by the decision function.

Diagnosis algorithm based on random forest

Solution:

The basic flow of the random forest algorithm for aircraft engine fault identification is shown in Fig. 2 (Li *et al.* 2017). Firstly, decision trees are constructed by using the training samples, and the random forest is formed to obtain the classification rules. Then, the engine data to be detected are classified based on the constructed random forest. The specific steps are as follows:





- Firstly, the training samples are input, and the training samples are resampled to form a new sample set *X*. The number of samples in *X* is also *n*, but there are duplicate samples in *X* as the samples are obtained by resampling (Song and Jiang 2016). Resampling is performed many times, and *m* new sample sets are obtained (*m* is determined according to demands). Every sample set *X* is the training sample of a decision tree.
- 2. The decision tree is constructed based on the sample set *X* obtained by resampling. The branching feature of the branch node in the decision tree is randomly selected from the sample feature attributes. One branch node can only choose one attribute, the branch growth is maximized, and the pruning operation is not carried out.
- **3.** Every sample set *X* constructs one decision tree, and all the decision trees form a random forest.
- 4. The samples to be detected are classified by the random forest.

Λ

Diagnosis algorithm based on PSO-BP

Back propagation (BP) neural network can fit the nonlinear law well (Lu *et al.* 2016). For the BP network in the training process, the most important thing is adjusting hidden layer parameters. In the traditional BP network, the parameters are adjusted step by step according to the set learning rate. On the one hand, the adjustment speed is slow; however, when there are many layers in the network, the differences will be weakened layer by layer when they propagate backward, resulting in little change in the parameters, and there will be an illusion that the learning has been completed. In order to reduce the influence of prematurity as much as possible, this study used the PSO algorithm to improve the reverse learning method of BP (Yuan *et al.* 2017). The parameters that need to be adjusted in the BP network in the training process are regarded as a group of coordinates of the particles. More than one particle represents more than one parameter optional scheme. The optimal particle is searched in the solution space to replace the parameter adjustment of the traditional BP network. Compared with the traditional BP algorithm, which uses the learning rate to adjust parameters step by step, the PSO algorithm makes particles fly in the solution space, and the parameters of each layer are adjusted synchronously, which will not weaken layer by layer. Also, the particle flight search will not adjust the fixed step size, but search for the globally optimal particle, which not only improves the learning speed but also avoids "premature" (Li 2020).

Before using the BP algorithm improved by the PSO algorithm for engine fault diagnosis, the basic training process is shown in Fig. 3, and the specific steps are as follows:



Figure 3. The training process of the PSO-BP-based fault diagnosis algorithm.

- 1. Training samples are input. Back propagation and PSO algorithms are initialized. The PSO algorithm mainly initializes population size and particle parameters. In this study, each particle in the population represents a BP parameter, and it will be substituted into the BP algorithm for a calculation to obtain a fitness for subsequent PSO adjustment; therefore, the population size should not be too large.
- 2. The parameter scheme represented by PSO particles are brought into the BP network to make a forwarding calculation on the input training samples, and the formula is shown in Eq. 4:

$$y = f(\sum_{i=1}^{n} \omega x_i - \beta) \tag{4}$$

where *y* is the output, β is the adjustment item of each layer, $f(\cdot)$ is the activation function, *w* is the weight between layers, x_i is the sample input.

- **3.** After the forward calculation of the BP network, the algorithm is judged. If the algorithm meets the termination conditions, the training will be stopped; if it does not, the next step of the PSO adjustment will be carried out.
- **4.** When the algorithm does not reach the termination condition, the population is iterated once using the formula of the PSO algorithm. The iteration formula is shown in Eq. 5:

$$\begin{cases} v_i(t+1) = \overline{\omega} v_i(t) + c_1 r_1 (P_i(t) - x_i(t)) + c_2 r_2 (G_g(t) - x_i(t)) \\ x_i(t+1) = x_i(t) + v_i(t+1) \end{cases}$$
(5)

where $v_i(t+1)$ and $x_i(t+1)$ represent the speed and position of particle *i* after one iteration, $v_i(t)$ and $x_i(t)$ are the speed and position of particle *i* before iteration, \overline{w} is the inertia weight of particles, c_1 and c_2 are learning factors, r_1 and r_2 are random numbers between 0 and 1, $P_i(t)$ is the optimal position that particle *i* that has experienced, and $G_g(t)$ is the best position that the particle swarm has experienced. The parameters represented by the particles after one iteration are substituted into the BP network for forward calculation again.

Steps 2, 3, and 4 are repeated until the algorithm reaches the termination condition. The trained algorithm can be used in aircraft engine fault diagnosis.

SIMULATION EXPERIMENT

Experimental environment

In this study, a comparative simulation experiment was carried on the above three fault diagnosis algorithms using MATLAB software in a laboratory server, configured with Windows 7 system, i7 processor, and 16 G memory (Gao *et al.* 2017).

Experimental data

This study took the fault data of Airbus 340-300 (engine type: CFM56-5C) as the subject of the simulation experiment. The normal and fault detection data of the aircraft were from AVIC Civil Aircraft Maintenance Co., Ltd. These samples contained characteristic data for aircraft engine diagnosis and corresponding normal or fault type tags. There were four types of characteristic data for diagnosis in each sample, which were high-pressure compressor speed (NH), low-pressure compressor speed (NL), exhaust gas temperature (EGT), and fuel flow (FF) (Zhao *et al.* 2017). The reason for choosing these four characteristics is that they could reflect the changes in the engine most directly after the engine failure and were easier to collect data, i.e., the cause of the failure could be inferred backwards based on the four characteristics of the change due to the failure. The types of engine faults in the fault data are shown in Table 1. There were 500 engine normal data samples and 800 engine fault data samples, 100 for each fault.

Fault number	Fault type	Fault number	Fault type		
1	Combustion chamber fault	2	TCC system fault		
3	BETA angle deviation	4	High-pressure turbine fault		
5	Efficiency reduction of the low- pressure compressor	6	Efficiency reduction of the high- pressure compressor		
7	Air bleed valve fault	8	First-grade turbine fault		

 Table 1. Types of engine faults.

The main cause of the combustion chamber fault was that the high temperature generated by the combustion of fuel resulted in the structural stress changes, and moreover the carbon buildup generated by combustion changed the shape of the nozzle fuel outlet, leading to abnormal combustion and aggravating the combustion chamber fault. The Torque Converter Clutch (TCC) system fault referred to the failure of the lightweight, high-speed C compiler used to execute the command file in the aircraft engine system. The BETA angle deviation refers to an abnormal opening and closing of the BETA angle in the engine. The highpressure turbine fault was an abnormal change in the structure of the high-pressure turbine in the engine. Efficiency reduction of the low-pressure compressor and high-pressure compressor meant that the low-pressure and high-pressure compressors of the engine could not reach the normal working efficiency. The air bleed valve fault means that the valve used to control the air outlet in the engine was not flexible enough to open and close on command. The first-grade turbine fault meant that the structure of the first-grade turbine in the engine has changed badly and was not working properly. One sample of each type of fault is shown in Table 2. 6

Number	1	2	3	4	5	6	7	8
NL	0.2	0.8	0.2	0.7	1.2	0.5	-0.7	-1.7
NH	0.2	5.3	0.5	0.4	0.5	-0.1	-0.4	0.7
EGT	17	86	9	5	1	-2	-5	-2
FF	3.5	16.2	2.5	2.5	2.7	2.5	-2.5	-2
	0.9	O.1	O.1	O.1	O.1	O.1	O.1	O.1
	O.1	0.9	O.1	O.1	O.1	O.1	O.1	O.1
	0.1	O.1	0.9	O.1	O.1	O.1	O.1	O.1
Standard autout	0.1	O.1	O.1	0.9	O.1	O.1	O.1	O.1
Standaru output	O.1	O.1	O.1	O.1	0.9	O.1	O.1	O.1
	O.1	O.1	O.1	O.1	O.1	0.9	O.1	O.1
	0.1	O.1	0.1	O.1	O.1	O.1	0.9	O.1
	O.1	O.1	O.1	O.1	O.1	O.1	O.1	0.9

Table 2. Typical fault samples.

Experimental setup

Through the orthogonal test, the relevant parameters of the SVM-based algorithm were finally confirmed as follows: sigmoid kernel function, penalty parameter = 1, and slack variable = 0.5. As to the random forest-based diagnosis algorithm, the number of decision trees was set as 100. The relevant parameters of the PSO-BP-based diagnosis algorithm were: sigmoid activation function (Wong *et al.* 2016), PSO population size = 30, learning factors = 1.5, the maximum number of iterations = 2000, and inertia weight = 0.8.

Experimental project (1): firstly, the fault diagnosis effects of the three algorithms were preliminarily compared using the artificial fault data set. The artificial fault data were obtained by adding interference to the normal data, and its formula is shown in Eq. 6:

$$X_{noise} = X_{normal} + k\sigma(rand() - 0.5)$$
⁽⁶⁾

where X_{noise} is the artificial fault data after adding interference, X_{normal} is normal data, k is data dispersion, σ is the standard deviation of the normal data set, and rand() is the random number in [0,1]. When fault data were generated using Eq. 6, the types of artificial faults were distinguished according to the value of k. In this study, the value of k was set as 0.1, 0.2, and 0.3; then, there were three kinds of artificial faults, 200 for each kind of fault. 60% of the normal samples and the samples of three kinds of faults were taken as the training set, and the remaining 40% were taken as the test set.

Experimental project (2): the three diagnosis algorithms were verified using real fault samples. 60% of the normal data and eight kinds of real fault data (780 samples) were taken as the training set, and the remaining 40% (520 samples) were taken as the test set.

Experimental results

Before testing the real fault data set with the three algorithms, the artificial fault data set was used for preliminary performance verification. The artificial fault data used for the preliminary verification of the algorithm recognition performance formed by adding interference to the normal data, and the interference mode is as described above. According to different dispersion degrees of the interference data, three fault types were constructed. Finally, the algorithm was tested by the normal data and fault data, and the results are shown in Fig. 4. The k = 0.0 meant that there was no interference, i.e., normal data. It was seen from Fig. 4 that the accuracy of the PSO-BP algorithm was the highest, with a comprehensive accuracy of 98.3%, the accuracy of the random forest algorithm was the second, with a comprehensive accuracy of 85.7%, and the accuracy of the SVM algorithm was the lowest, with a comprehensive accuracy of 79.2%, whether for normal engine data or for artificial engine fault data.



Figure 4. The recognition performance of three diagnosis algorithms on an artificial fault data set.

After the preliminary recognition performance test with the artificial fault set, the real fault data were used for performance verification. Artificial fault data formed by adding interference. Although the randomness of interference increased by random numbers, artificial pseudo-random made the randomness limited. Different fault types were distinguished obviously, and the characteristics of the same fault type were similar, i.e., the characteristics of artificial data were too ideal and not in line with reality. Compared with the artificial fault data, the real fault data in different fault types and within the same fault type not only had a larger dispersion degree in data characteristics but also had stronger randomness and greater diagnosis difficulty. It was seen from Fig. 5 that in the face of different kinds of real faults, the PSO-BP-based diagnosis algorithm had the highest recognition accuracy of 97.0%. The accuracy of the random forest algorithm was the second, with a comprehensive recognition accuracy of 79.0%; the accuracy of the SVM algorithm was the lowest, with a comprehensive recognition accuracy of 66.8%.





For aircraft engines, once there is a fault, the flight safety will be greatly shaken. Especially in the flight process, in the face of sudden failure, the sensor needs to make a judgment in time after receiving relevant data. Therefore, the recognition time of the diagnosis algorithm was also a part of measuring algorithm performance. As shown in Fig. 6, it took 0.982 and 1.121 s for the SVM-based diagnosis algorithm to identify artificial faults, 0.521 and 0.612 s for the random forest-based diagnosis algorithm to identify artificial faults, 0.521 and 0.612 s for the random forest-based diagnosis algorithm to identify real faults. As can be seen from Fig. 6, whether it was an artificial fault or a real fault, the PSO-BP-based diagnosis algorithm took the least time, and the SVM-based diagnosis algorithm took the most time; whatever the algorithm was, it always took more time to identify the real fault data.

8



Figure 6. The average time consumption of three algorithms for fault data identification.

CONCLUSION

This paper briefly introduced the SVM, random forest, and PSO-BP-based fault diagnosis algorithms for aircraft engines and carried out a simulation experiment on the performance of the three algorithms in MATLAB software. The experimental data used for simulation included the artificial fault data obtained by adding interference to the normal data and the real fault data obtained from AVIC Civil Aircraft Maintenance Co., Ltd. The final simulation results are as follows: (1) in the recognition of artificial fault data, the comprehensive recognition accuracy of the PSO-BP-based diagnosis algorithm was 98.3%, the comprehensive accuracy of the random forest-based diagnosis algorithm was 85.7%, and the comprehensive accuracy of the SVM-based diagnosis algorithm was 79.2%; (2) in the recognition of real fault data, the comprehensive recognition accuracy of the PSO-BP-based diagnosis algorithm was 97.0%, the comprehensive recognition accuracy of the random forest-based diagnosis algorithm was 79.0%, and the comprehensive recognition accuracy of the SVM-based diagnosis algorithm was 66.8%; (3) in terms of the average time consumption of fault data recognition, the average time of the PSO-BP-based diagnosis algorithm was 66.8%; (3) in terms of the average time consumption of fault data recognition, the average time of the PSO-BP-based diagnosis algorithm was the shortest, and the average time of the SVM-based diagnosis algorithm was the longest; the three algorithms took more time to identify the real fault data.

In this paper, the faults of the engine were determined by the PSO-BP algorithm based on the collected data, and its effectiveness was verified by comparing it with two other intelligent algorithms. The shortcoming of this study is that the type of the engine used in the experiment was single, and the variety of data for determining engine faults was low. The future research direction is to increase the types of data used to determine engine faults and to generalize the intelligent algorithm so that it can diagnose more types of aircraft engines.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable.

FUNDING

Not applicable.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

Ali AH, Markarian G, Tarter A, Kölle, R (2010) Feasibility Demonstration of Diagnostic Decision Tree for Validating Aircraft Navigation System Accuracy. Journal of Aircraft 47(6):2096-2102. https://doi.org/10.2514/1.C000323

Furse C, Kafal M, Razzaghi R, Shin Y-J (2021) Special Issue on Embedded Sensors for Fault Diagnosis in Electrical Wiring Interconnection Systems, Power Grids, Structural Cables, Pipelines, and Electrical Machines. IEEE Sensors Journal 21(2):886-887. https://doi.org/10.1109/JSEN.2020.3035852

Gao Z, Ma C, Song D, Liu Y (2017) Deep quantum inspired neural network with application to aircraft fuel system fault diagnosis. Neurocomputing 238:13-23. https://doi.org/10.1016/j.neucom.2017.01.032

Kiran V, Hemantha K, Gangadharan KV, Salih D, Bendaya M (2018) Engine gearbox fault diagnosis using machine learning approach. J Qual Maint Eng 24(3):345-357. https://doi.org/10.1108/JQME-11-2015-0058

Li J (2020) Research Progress on Typical Fault Monitoring and Diagnosis of Aircraft Engine. IOP Conf Ser: Mater Sci Eng 772:012030.

Li Z, Gao P, Zhao D, Liu J (2017) Fault diagnosis and location of the aero-engine hydraulic pipeline based on Kalman filter. Adv Mech Eng 9(12):1-9. https://doi.org/10.1177/1687814017742811

Loza AF, Cieslak J, Henry D, Dávila J, Zolghadri A (2015). Sensor fault diagnosis using a non-homogeneous high-order sliding mode observer with application to a transport aircraft. IET Control Theory Appl 598-607.

Lu R, Li X, Li Y, Hu W, Teng C (2016) Design of uncontained engine rotor fragments failure during airplane design. Acta Aeronautica et Astronautica Sinica 37(1):351-363. https://doi.org/10.7527/S1000-6893.2015.0341

Ma A, Li Y, Cao Y, Wang Z, An G (2015) Intelligent diagnosis for aircraft engine wear fault based on immune theory. Acta Aeronautica et Astronautica Sinica 36(6):1896-1904. Chinese. https://doi.org/10.7527/S1000-6893.2014.0357

Ma J, Chen L, Liu H (2015) Fault Diagnosis for the Heat Exchanger of the Aircraft Environmental Control System Based on the Strong Tracking Filter. PloS ONE 10(3):e0122829.

Sevil HE, Dogan A (2015). Fault Diagnosis in air data sensors for receiver aircraft in aerial refueling. J Guid Control Dyn 38(10). https://doi.org/10.2514/1.G000527

Song M, Jiang W (2016) Engine fault diagnosis based on sensor data fusion using evidence theory. Adv Mech Eng 8(10):1-9. https://doi.org/10.1177/1687814016673291

White A, Karimoddini A (2020) Event-based diagnosis of flight maneuvers of a fixed-wing aircraft. Reliab Eng Syst Saf 193:106609. https://doi.org/10.1016/j.ress.2019.106609

Wong PK, Zhong J, Yang Z, Vong CM (2016) Sparse Bayesian extreme learning committee machine for engine simultaneous fault diagnosis. Neurocomputing 174(Part A):331-343. https://doi.org/10.1016/j.neucom.2015.02.097

Yang X, Pang S, Shen W, Lin X, Jiang K, Wang Y (2016) Aero Engine Fault Diagnosis Using an Optimized Extreme Learning Machine. Int J Aerosp Eng 2016:7892875. https://doi.org/10.1155/2016/7892875

Yuan Y, Liu X, Ding S, Pan B (2017) Fault detection and location system for diagnosis of multiple faults in aeroengines. IEEE Access 5:17671-17677. https://doi.org/10.1109/ACCESS.2017.2744639

Zhao Y-P, Song F-Q, Pan Y-T, Li B (2017) Retargeting extreme learning machines for classification and their applications to fault diagnosis of aircraft engine. Aerosp Sci Technol 71:603-618. https://doi.org/10.1016/j.ast.2017.10.004

Zhao Y-P, Wang J-J, Li X-Y, Peng G-J, Yang Z (2020) Extended least squares support vector machine with applications to fault diagnosis of aircraft engine. ISA Trans 97:189-201. https://doi.org/10.1016/j.isatra.2019.08.036

Zhao Z, Zhang J, Sun Y, Liu Z (2017) Sensor fault diagnosis of aero-engine based on divided flight status. Rev Sci Instrum 88:115007. https://doi.org/10.1063/1.5000105