

Low-Altitude Status Monitoring and Fault Diagnosis Algorithm for Aircraft Equipment Based on Big Data Mining and Analysis

Wenjing Yin^{1,a}, Sichao Lv^{1,b*} and Xiaofei Li^{2,c}

¹*School of Transportation Engineering, Weifang Engineering Vocational College, Qingzhou 262500, Shandong, China*

²*School of Marxism, Weifang Engineering Vocational College, Qingzhou 262500, Shandong, China*

The monitoring and fault diagnosis of aircraft equipment at low-altitudes refers to the monitoring of the relevant parameters of various operating states of aircraft equipment, effectively determining its operational status based on existing data, and the diagnosis of any faults, so that appropriate technical remedial measures can be taken. This study examines the analysis and application of aircraft equipment low-altitude state monitoring and fault diagnosis methods based on big data mining (DM), and establishes relevant systems. The parameters of aircraft equipment were analyzed by applying association rules, whereby the rules governing the operation of equipment were found. The verification analysis and experimental data of the system showed that in the test set, when the number of aircraft equipment fault samples was 200, the correct number of fault samples for the traditional system and the system proposed in this paper was 171 and 191 respectively. The fault diagnosis time of the two systems was 1.127s and 0.562s respectively. This indicates that the monitoring and fault diagnosis based on big DM analysis of aircraft equipment in the low-altitude condition has certain feasibility and merits further promotion and application.

Keywords: aircraft equipment, big data mining analysis, fault diagnosis algorithm, low altitude condition monitoring

1. INTRODUCTION

The goal of fault diagnosis is to examine the fault closely once it has been detected; this comprises its location, cause, and repair. First of all, it is necessary to solve the problem and formulate corresponding measures to prevent potential economic loss and human casualties. When using big DM technology for fault diagnosis, it is not necessary to establish a complete system model. Only reliable quantitative analysis and quantitative data are needed to diagnose this type of fault. In order to ensure safe and reliable aircraft operations, it is important to conduct low-altitude status monitoring,

diagnosis and prediction of aviation equipment operation in terms of health assessment, life tracking, decision-making in regard to maintenance, and material reserves during the low-altitude flight phase

With the continuous development of society, the research of aircraft equipment has gradually increased. The purpose of Lvov's research was to compare various implementation schemes of optical fiber strain sensors to monitor the stress and strain status of monitored objects. Due to the increased requirements for the safe operation of complex technical objects (such as bridges, structures, aircraft, automobiles, etc.), it is necessary to monitor the technical status of various equipment [1]. Jianzhong introduced the use of existing sensor parameters to extract human health index

*Corresponding Author. Email: ^a yinwenjing@wfec.edu.cn,
^b wgaelcsc@163.com, ^c wfgc1988@163.com

for health monitoring of the air conditioning system in traditional commercial aircraft [2]. Pagnoni found that the prediction in the next decade indicated that the equipment in the engine cabin might exceed its temperature threshold, thus endangering the aircraft safety. In order to meet the compatibility, safety and environmental restrictions, titanium water loop heat pipes were designed, manufactured and tested for the thermal management of aircraft engine equipment [3]. Although these studies promoted the research of aircraft equipment to a certain extent, they were not conducted in the context of actual situations.

At the same time, big DM analysis has gradually attracted widespread attention from the academic community. Sharypov discussed the life cycle monitoring of aircraft components as a tool to identify unreal aviation equipment items, which led to the improvement of big DM analysis [4]. Back's research involved the structural safety analysis of thrust-measuring test equipment on small engines, which promoted the application of big DM analysis to aircraft equipment [5]. Astakhov described the various brakes used on rocket orbital facilities when testing aerospace and rocket technology equipment. This determined the technical capabilities of conventional types and braking modes, including their advantages and disadvantages and their application scope. The acceptable brake types in the high-speed track test were analyzed and studied, indicating the advantages of big DM analysis [6]. Although these research methods were innovative, the amount of experimental data was not adequate enough to prove the reliability of the methods.

This study examined the methods used for the monitoring and fault diagnosis of equipment of aircraft flying at a low-altitude, application of big DM analysis, and the degree of correlation between aircraft equipment parameters and their operational mode. By using big DM analysis technology, the association rules between various parameters of aircraft equipment and the operation rules of various devices were found, and the aircraft equipment status monitoring and fault diagnosis algorithm system was established. Then, the proposed system was tested. The experimental analysis was carried out on the sample of aircraft equipment to determine delay rate, stability, the number of faults detected, the time required to do so.

2. AIRCRAFT EQUIPMENT LOW-ALTITUDE STATUS MONITORING AND FAULT DIAGNOSIS ALGORITHM

2.1 Big DM Evaluation

Big data analysis involves a comprehensive analysis of massive amounts of data [7–8]. Currently, the growth rate of various types of information is unprecedented. It is conservatively estimated that in the next few years, the growth rate will reach an annual rate of 50%. In big data technology, the biggest difficulty is how to choose DM [9]. Whether it is by means of outlier analysis, segmentation, clustering, etc., the data can be mined to yield a great amount

of value. Generally, these algorithms need to process a large amount of data quickly in the shortest time to obtain the most useful information from big data. In order to obtain the maximum value from DM, it is necessary to select the appropriate algorithm [10].

The aircraft equipment low-altitude condition monitoring and fault diagnosis algorithm system studied in this paper was to use big DM analysis technology to mine out the operation rules of various equipment and the relationship between various parameters from a large amount of historical data, and combined its historical development rules to monitor the status of existing equipment and conduct fault diagnosis.

DM technology is used to find useful information in a large amount of data, although this is a complex process [11–12]. When searching, it would be affected by many external environments, such as the lack of some important data, how to reduce the number of data and how to find hidden information behind the data.

Initially, DM is not independent, but connected with the data warehouse. After preprocessing, it collects and stores all data so as to form a data warehouse [13]. However, with the continuous development of DM technology, its concept has gradually been introduced into the concept of data from a simple data warehouse [14]. Therefore, the deep learning that it makes possible enables management personnel to make data-based decisions, which is also an important area of artificial intelligence. Later, with the development of computer technology, more coding tools emerged, thus bringing infinite vitality and new technology to DM. However, the data warehouse plays an important role in the development of DM. Without a data warehouse, there would be no DM today, the concept has played a huge role in promoting the development of DM [15–16].

The specific process of DM for the monitoring of aircraft equipment at low altitudes is shown in Figure 1.

It can be seen from Figure 1 that the DM process includes the following steps:

Selection of required data: A large amount of data may contain a great deal of irrelevant information; hence, DM mining is used to extract only the information of interest. Some of these data are missing and need to be inserted manually, and some are obvious errors that must be deleted. The manual application of DM technology collects and organizes useful information.

Data preprocessing: The data are extracted from different information such as unit location data, system configuration data, content structure data, etc. Because their units and dimensions are different, they must be dimensionless so that the data is unified.

Data conversion: In terms of data preprocessing, there are some data that cannot be used by humans. This requires the building of an analysis model for the conversion of data before conducting DM.

DM: mining and analysis of information in the model: For this process, the most important thing is to choose an algorithm. The quality of an algorithm determines its success or failure, so this step is very important. This is because, except for the selection of algorithm, other tasks can be completed by the computer without the need for manual work. In this step, the user is required to make

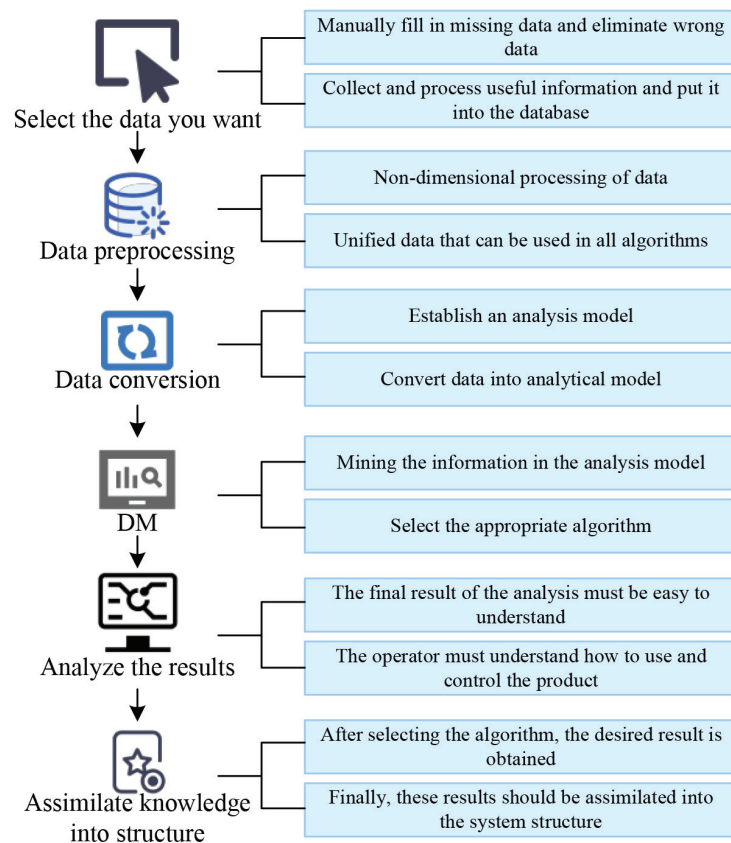


Figure 1 DM process specifically for low altitude aircraft.

assumptions, select appropriate algorithms, and build an appropriate environment.

The results are analyzed: The results may be only one or several numbers or letters. Therefore, the results should be analyzed and the conclusions must clearly presented. DM is a technology whose users need to understand its operation and control methods.

Knowledge should be assimilated into the structure: by selecting the appropriate algorithm, the required results are obtained and assimilated into the architecture.

The goal of DM is to let the data speak for themselves. This is done by mining knowledge from the database, and extracting valuable, characteristic and potential application value from a large amount of data, so as to formulate clear rules that establish the connection between data [17–18]. For fault diagnosis, due to the development of current data collection technology, an increasing amount of information can be collected. This means that the traditional fault diagnosis methods can no longer meet the requirements of fault-related decision making. Hence, there is a need for a new method that can effectively and efficiently extract fault rules. The combination of DM technology and practical application has expanded the former's advantages, subsequently leading to the emergence of innovative hardware and software. The application of DM in practice has also achieved fruitful results [19–20].

The low-altitude fault diagnosis process based on DM is shown in Figure 2.

According to Figure 2, the first step is to information about the subject of diagnosis. Then the collected data is processed

prior to conducting DM and generating the association rule. The operation rule mode should be calculated, and finally the governance measures should be obtained.

2.2 Calculation of Correlation Between Airborne Equipment Parameters at Low Altitude

There are many methods of DM. For each method, the DM angle has its own characteristics, as shown in Figure 3.

It can be seen from Figure 3 that regression analysis is a method of forecasting and correlating data series, which involves using mathematical functions to clarify the correlation between data, thus determining the data attributes.

A neural network has many characteristics such as nonlinear processing and distributed storage, and therefore offers unique advantages in dealing with DM problems.

Association rule mining is used to deduce the possibility of another occurrence based on the presence of one item of data. It can be divided into two steps: From a large number of original numbers, all high-frequency groups are found; then relevant rules are established for those groups with higher frequency..

Clustering is used to classify different data. The difference between similar data is small, but the difference between different data is large.

Classification methods can be used for application classification and trend prediction. By finding the common features

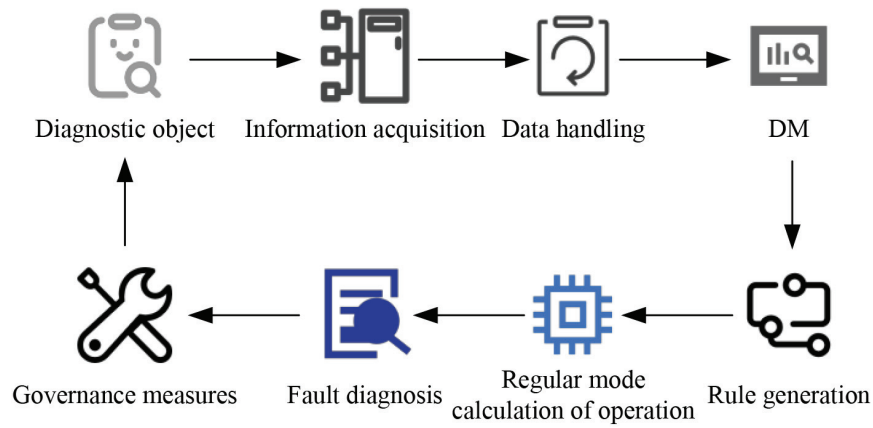


Figure 2 Low-altitude fault diagnosis process based on DM.

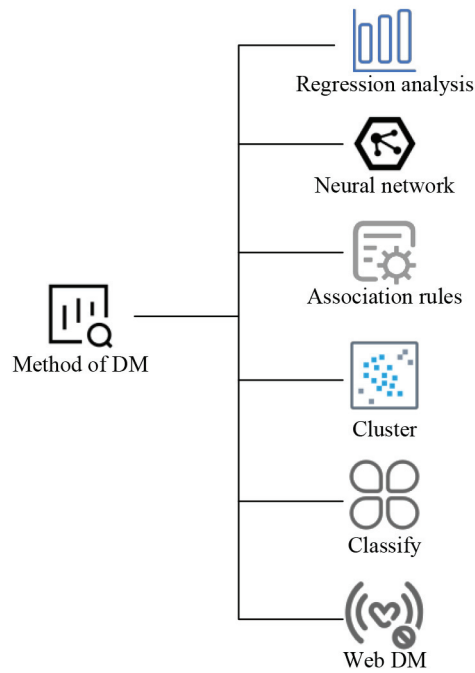


Figure 3 DM method.

of a large amount of data and classifying them, according to different classification modes, each feature in a large amount of data can be classified.

Web DM is a comprehensive technology, including Page Rank algorithm.

A device in an aircraft has many parameters that can include wing, aileron, fuselage, power unit, landing gear, aircraft length, etc. There are complex relationships between these parameters. In practical applications, the combination degree algorithm can effectively reduce the dimensionality of the model. The aviation equipment low-altitude condition monitoring system based on big DM technology is use to eliminate those factors that have little impact on the equipment operation status according to the interrelationship of various parameters, thereby laying the foundation for the selection of an equipment model. In this paper, the grey correlation degree algorithm was used to analyze the operation parameters of important aircraft equipment.

It is assumed that two sequences, namely x and y , are to be analyzed according to their degree of correlation.

The x and y sequences are dimensionless. Because the meaning of x and y sequences may be different, the dimension is not necessarily the same. Therefore, it is difficult to compare them and draw accurate conclusions. Hence, before correlation analysis, it must be preprocessed, making it dimensionless.

The correlation coefficient $\vartheta(x_o)$ of the two sequences of x and y is obtained with Formula (1) where v represents the resolution coefficient. Generally, the value range is 0 to 1, which is usually 0.5.

$$\vartheta(x_o) = \frac{\Delta(\min) + v\Delta(\max)}{\Delta_{po}(l) + v\Delta(\max)} \quad (1)$$

Among them, $\vartheta(x_o)$ represents sequence correlation system; $\Delta(\min)$ represents the second level minimum difference; $\Delta(\max)$ represents the maximum difference between two levels; $\Delta_{po}(l)$ represents the absolute difference between each point on the curve of each comparison sequence x_o and each point on the curve of the reference sequence.

The correlation degree r_o should be calculated.

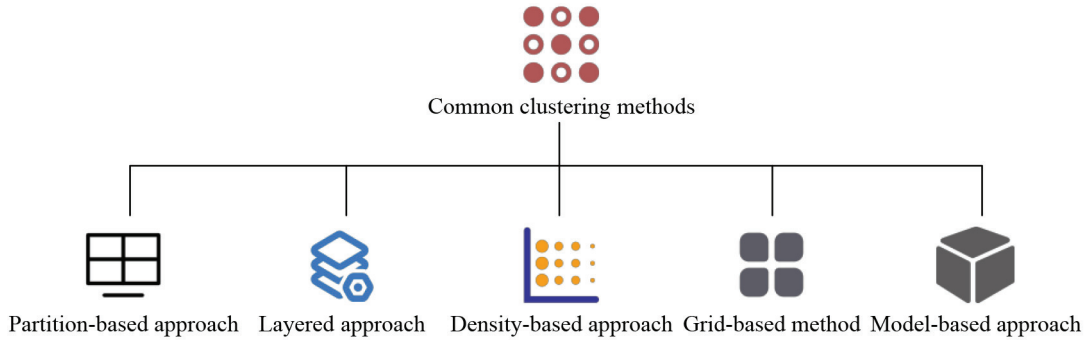


Figure 4 Common clustering methods.

This parameter is calculated with Formula (2), which represents the correlation of each sequence in a certain time.

$$r_o = \frac{1}{M} \sum_{l=1}^M \vartheta_o(l) \quad (2)$$

In the formula, r_o represents the grey correlation degree between sequence x and y . The closer the value of r_o is to 1, the better the correlation between the sequence columns x and y .

2.3 Calculation of Low-Altitude Operation of Aircraft Equipment

The essence of the aircraft equipment low-altitude condition monitoring and fault diagnosis algorithm system based on big DM technology is to discover the relationship between various parameters and the operation rules of various devices. This is done by using big DM to analyse a large amount of historical data so as to establish an equipment-monitoring model. By the monitoring the equipment, its operational status can be determined in real time, which will enable the early diagnosis of the equipment fault. On this basis, the equipment low-altitude condition monitoring and fault diagnosis algorithm system is studied. Firstly, the degree of correlation of each parameter must be determined; secondly, the operation rules of the equipment must be known.

The field of aircraft equipment comprises mostly time series data, which is both crucial and specific. By combining big DM technology with equipment status monitoring technology, the working principle of a device can be found and its operational status can be monitored. In this process, the data collected are continuous data, which can understand the working rules of the device and further explore its regularity. There are three main research areas at present. By means of functional data analysis, DM is studied in depth via clustering technology. These data generally have a high degree of similarity, although this depends on the type of data. Hence, the similarity between certain data types can be very low. There is no specific classification standard for cluster analysis, which is generally classified according to the characteristics of the data itself. Its classification methods are diverse so the potential relationship between data can be determined.

In functional data processing, the development speed of the clustering algorithm is relatively slow. The operation mode involves clustering function types by expanding the function base. In mathematics, it has been proved that on a standard orthogonal basis function, the Euclidean distance of the initial function is consistent with the Euclidean distance of the coefficient vector of the basic expansion. This shows that there are many clustering algorithms in time series data. On this basis, the functional data clustering analysis technology is introduced into the field of equipment monitoring, and the clustering results are used to monitor the equipment, effectively reducing the risk of equipment failure.

Common clustering methods according to different clustering rules are shown in Figure 4:

In all clustering problems, distance is a very important indicator. 111 $x(t), c(t)$ given. Common indicators for measuring distance include the following:

Difference supremum $D_{xc}^{(1)}$:

$$D_{xc}^{(1)} = \sup\{|x(t) - c(t)|\} \quad (3)$$

Consistent difference $D_{xc}^{(2)}$:

$$D_{xc}^{(2)} = \int_a^T |x(t) - c(t)| dt \quad (4)$$

European distance $D_{xc}^{(3)}$:

$$D_{xc}^{(3)} = \int_a^T [x(t) - c(t)]^2 dt \quad (5)$$

The characteristic of Euclidean distance is very good, so it is also a common measure function. However, this method can only describe the difference of curves, and cannot indicate the various types of curves. Therefore, in this paper, a similarity index and similarity index were used based on rank correlation in morphology. However, these two algorithms require a lot of integration operations, and always have certain limitations in regard to calculation speed, which hampers the clustering process. However, although European distance has some shortcomings, it is the best. In order to overcome the above defects, the function can be normalized in advance, or the specific derivatives of the function can be clustered. In general, in this paper, Euclidean distance is still considered as an important measure of functional clustering similarity.

Table 1 Association rule table and association rule learning table.

Name	Type	Whether it can be empty
Association rule name	Int	No
Fact	Varchar50	Yes
Hypothesis	Varchar50	Yes
Fact factor	Float	Yes
Determining factor	Float	Yes
Rule priority	Int	No
Status monitoring	Varchar50	No
Fault diagnosis	Varchar50	Yes
Remarks	Varchar50	Yes

Table 2 User information table.

Name	Type	Whether it can be empty
ID	int	No
Name	Varchar50	No
Age	int	Yes
Gender	Varchar50	Yes
Position	Varchar50	Yes
Remarks	Varchar50	Yes

3. TEST AND EVALUATION OF AIRCRAFT EQUIPMENT LOW-ALTITUDE STATUS MONITORING AND FAULT DIAGNOSIS ALGORITHM SYSTEM

3.1 Design and Implementation of Low-Altitude State Monitoring and Fault Diagnosis Algorithm System

In the field of aviation equipment, there are many expert systems for fault diagnosis. In practical applications, due to the limitation of rule search speed and analysis ability, there is almost no practical algorithm system for the diagnosis of faults in aircraft equipment. In this paper, the grey correlation method was used to analyze the correlation between the parameters, and the working mode of the equipment was obtained by applying the similarity coefficient clustering method, which improved the working efficiency and accuracy of the diagnosis system. Based on this, an aircraft equipment low-altitude condition monitoring and fault diagnosis algorithm system utilizing big DM technology was proposed.

The combination of DM technology and aircraft equipment low-altitude condition monitoring and diagnosis system can greatly improve the operational efficiency of the system. The users of the system are usually experts. The system involves input inquiry, system reasoning, and output of reasoning results. The users (experts) are responsible for knowledge collation, knowledge entry, rule generation, and parameter setting.

The low-altitude condition monitoring and fault diagnosis algorithm system comprises a user database, fact database and knowledge base. The knowledge base consists of a

factor table, an output expression table, a hypothesis logic expression table, a hypothesis expression explanation table, a rule table and inference rule expression table, and an output logic expression table. The factor table in the knowledge quantity database stores the factor name and value. The output expression table stores the name, identification and solution of the output expression. The hypothesis expression explanation table stores the identifier and name of the hypothesis expression, and the keyword logical expression table stores the logical identifier and the keyword. The composition and parameters of access rules for rule tables and inference rule tables. Fact database and user database include user data table, fact data table and association rule learning table. Data tables are used to store different factors. The association rule learning table is used for storing new rules. If no existing rule can be found, a new rule is generated.

Both the association rule table and the rule learning table have the same content as shown in Table 1.

Table 1 shows the association rule table and the association rule learning table comprising the association rule name, fact, hypothesis, fact factor, determination factor, rule priority, status monitoring, fault diagnosis and remarks.

The user information table was used to store user-related information, as shown in Table 2.

Table 2 contains information related to Identity Document (ID), name, age, gender, position and remarks.

In the proposed system, the grey correlation algorithm was used to calculate the correlation between the parameters, and the similarity coefficient clustering method was applied to obtain the working mode of the equipment, so as to apply it to the knowledge representation and rule extraction. The system can achieve real-time monitoring and fault diagnosis of aircraft equipment, and can obtain rule extraction and knowledge expression, thus enhancing the applicability of the system.

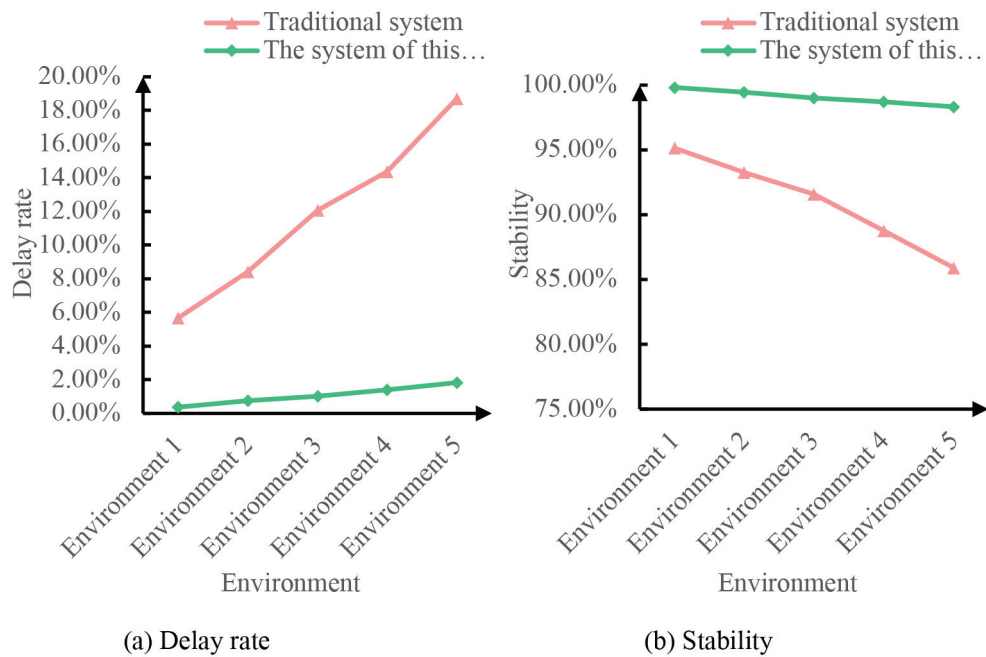


Figure 5 Real-time and stability analysis.

3.2 Verification and Analysis of Low-Altitude State Monitoring and Fault Diagnosis Algorithm System

The following experimental results demonstrate the effectiveness of the aircraft equipment status monitoring and fault diagnosis algorithm system proposed in this paper in comparison with the performance of the traditional algorithm applied to monitor and diagnose faults in the equipment of aircraft flying at low altitudes. Both systems were tested and analyzed under five different influence factors.

Figure 5 shows the delay rate and stability of the two systems.

Figure 5 (a) shows that in environment 1, the delay rate of the traditional system reaches 5.66%, and the delay rate of the proposed system is 0.37%; in environment 2, the delay rates of the two systems for aircraft equipment status monitoring are 8.42% and 0.74% respectively; in environment 3, the delay rates of the two systems are 12.08% and 1.02% respectively; in environment 4, the delay rates are 14.35% and 1.39% respectively; in environment 5, the delay rates are 18.71% and 1.83% respectively. It is evident that the delay rates of the proposed system designed for aircraft equipment status monitoring were consistently less than those of the traditional system.

It can be seen from Figure 5 (b) that under environment 1, the stability of the traditional system is 95.14%, and the stability of the proposed system is 99.82%; in environment 2, the stability of the two systems is 93.27% and 99.46% respectively; in environment 3, the stability of the two systems is 91.59% and 99.02% respectively; under environment 4, the stability of the two systems is 88.75% and 98.71% respectively; under environment 5, the stability of the two systems is 85.91% and 98.33% respectively. The experimental results indicate that the stability of the proposed

system for aircraft equipment status monitoring is greater than that of the traditional system.

In this study, 600 samples of aircraft equipment faults were selected for the experiment. The number of fault samples in the training set and the test set was 400 and 200 respectively.

Figure 6 shows the number of samples used for fault diagnosis of the two systems applied to the training set and the test set.

Figure 6 (a) shows that in the training set, the correct number of 50 aircraft equipment fault samples diagnosed by the traditional system and the system in this paper is 45 and 49 respectively; when the number of aircraft equipment fault samples is 150, the correct number of fault samples for diagnosis of the two systems is 129 and 144 respectively; when the number of aircraft equipment fault samples is 250, the correct number of the two systems' diagnosis fault samples is 209 and 239 respectively; when the number of aircraft equipment fault samples is 350, the correct number of fault samples for diagnosis of the two systems is 286 and 335 respectively.

According to Figure 6 (b), in the test set, the correct number of 40 aircraft equipment fault samples diagnosed by the traditional system and the system in this paper is 37 and 40 respectively; when the number of aircraft equipment fault samples is 80, the correct number of fault samples for the two systems is 71 and 78 respectively; when the number of aircraft equipment fault samples is 120, the correct number of fault samples for the two systems is 106 and 116 respectively; when the number of aircraft equipment fault samples is 200, the correct number of fault samples for the two systems is 171 and 191 respectively.

As shown in Figure 6, when the number of samples is the same, the number of aircraft equipment fault samples diagnosed by the proposed system is greater than the number obtained using traditional system diagnosis.

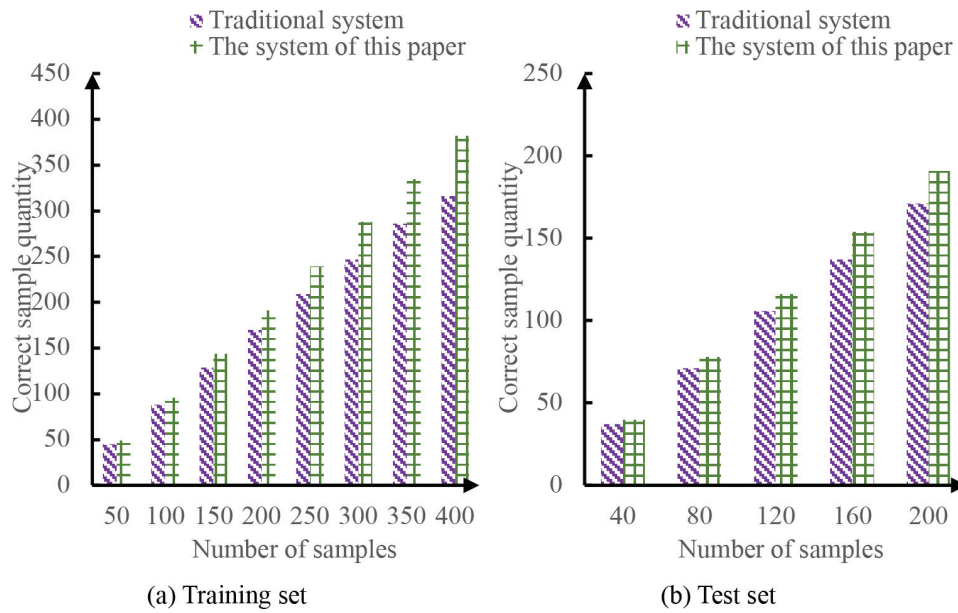


Figure 6 Number of samples for correct fault diagnosis.

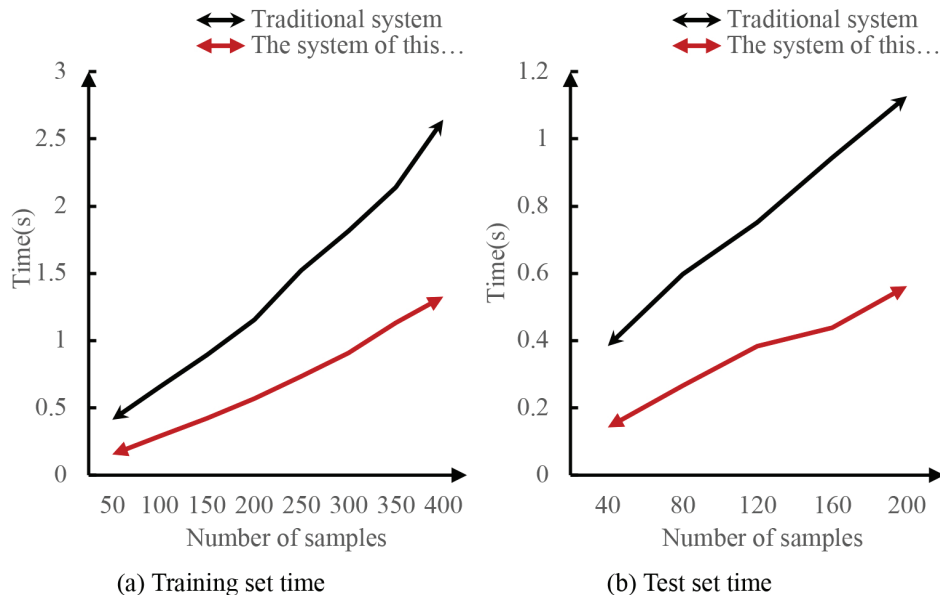


Figure 7 Fault diagnosis time (s).

Figure 7 shows the fault diagnosis time-related data for the two systems in the training set and the test set.

As shown in Figure 7 (a), when the number of aircraft equipment fault samples is 100 on the training set, the fault diagnosis time of the two systems is 0.657s and 0.292s respectively; when the number of aircraft equipment fault samples is 200, the fault diagnosis time of the two systems is 1.154s and 0.570s respectively; when the number of aircraft equipment fault samples is 300, the fault diagnosis time of the two systems is 1.817s and 0.906s respectively; when the aircraft equipment fault sample data is 400, the fault diagnosis time of the two systems is 2.642s and 1.327s respectively.

It can be seen from Figure 7 (b) that in the test set, when the aircraft equipment fault sample data is 40, the fault diagnosis time of the traditional system and the system in this paper is 0.384s and 0.142s respectively; when the aircraft equipment

fault sample data is 80, the fault diagnosis time of the two systems is 0.598s and 0.267s respectively; when the aircraft equipment fault sample data is 120, the fault diagnosis time of the two systems is 0.752s and 0.384s respectively; when the aircraft equipment fault sample data is 200, the fault diagnosis time of the two systems is 1.127s and 0.562s respectively.

The data in Figure 7 shows that when the number of aircraft equipment fault samples are the same, the time required to diagnose the number of samples using this system is shorter than that of the traditional system.

4. CONCLUSIONS

With the rapid development of the aviation industry, the data of aircraft equipment is also constantly being quantified.

In particular, the development of big data and analysis technology have improved the status monitoring and fault diagnosis of aircraft equipment during low-altitude flight. The aim is to have a more complete big data method. China's major airlines have established data management and analysis systems for passenger service, air control, flight status, etc. These have been developed by themselves or in collaboration with information technology companies. Also, the development of big DM technology has greatly advanced many aspects of the aviation industry. At present, big DM technology has been widely used in the low-altitude condition monitoring and fault diagnosis of aviation equipment. However, with the increasing number of various engine data, big DM technology could be widely used for the monitoring of low-altitude conditions and the fault diagnosis of aircraft equipment.

REFERENCES

1. Lvov Nikolai, Khabarov Stanislav, Todorov Aleksander, Barabanov Aleksander. "Versions of fiber-optic sensors for monitoring the technical condition of aircraft structures". *Civil Engineering Journal* 4.12 (2018): 2895-2902.
2. Jianzhong S. U. N., Li Chaoyi, Liu Cui, Gong Ziwei, Wang Ronghui. "A data-driven health indicator extraction method for aircraft air conditioning system health monitoring". *Chinese Journal of Aeronautics* 32.2 (2019): 409-416.
3. Pagnoni Filippo, Ayel Vincent, Bertin Yves, Coulloux Jerome, Zebian Maxime. "Loop heat pipe for thermal management of aircraft engine equipment". *Journal of Thermophysics and Heat Transfer* 35.2 (2021): 323-334.
4. Sharypov A., Brusnikin V., Koval S., Glukhov G., Gubanov O. "Aircraft components life cycle monitoring as a tool for identifying inauthentic aviation equipment items". *International Journal of Mechanical Engineering and Technology (IJMET)* 9.7 (2018): 612-620.
5. Back Kyeongmi, and Hyunbum Park. "A study on structural design and analysis of small engine test equipment for use in aircraft". *Journal of Aerospace System Engineering* 12.1 (2018): 42-46.
6. Astakhov S. A., V. I. Biruykov, and A. V. Kataev. "Effectiveness evaluation of various methods of the retainable equipment braking at the limited length while high-speed track tests of aircraft and rocket engineering products". *VestnikMoskovskogoaviatsionnogo institute* 29.2 (2022): 20-34.
7. Arena Fabio, and Giovanni Pau. "An overview of big data analysis". *Bulletin of Electrical Engineering and Informatics* 9.4 (2020): 1646-1653.
8. Haoxiang Wang, and S. Smys. "Big data analysis and perturbation using data mining algorithm". *Journal of Soft Computing Paradigm (JSCP)* 3.01 (2021): 19-28.
9. Atluri Gowtham, Anuj Karpatne, and Vipin Kumar. "Spatio-temporal data mining: A survey of problems and methods". *ACM Computing Surveys (CSUR)* 51.4 (2018): 1-41.
10. Ali Haseeb, Mohd Salleh Mohd Najib B., RohmatSaedudin Rd., Hussain Kashif. "Imbalance class problems in data mining: A review". *Indonesian Journal of Electrical Engineering and Computer Science* 14.3 (2019): 1560-1571.
11. Edastama, Primasatria, Amitkumar Dudhat, and Giandari Maulani. "Use of Data Warehouse and Data Mining for Academic Data: A Case Study at a National University". *International Journal of Cyber and IT Service Management* 1.2 (2021): 206-215.
12. Zou Hailei. "Clustering algorithm and its application in data mining". *Wireless Personal Communications* 110.1 (2020): 21-30.
13. Sour Alireza, and Rahil Hosseini. "A state-of-the-art survey of malware detection approaches using data mining techniques". *Human-centric Computing and Information Sciences* 8.1 (2018): 1-22.
14. Ageed Zainab Salih, Zeebaree Subhi R. M., Sadeeq Mohammed Mohammed, Kak Shakir Fattah, Rashid Zryan Najat, Salih Azar Abid, et al. "A survey of data mining implementation in smart city applications". *Qubahan Academic Journal* 1.2 (2021): 91-99.
15. Liu Yi, Elsworth Benjamin, Erola Pau, Haberland Valerii, Hemani Gibran, Lyon Matt, et al. "EpiGraphDB: a database and data mining platform for health data science". *Bioinformatics* 37.9 (2021): 1304-1311.
16. Goswami Saptarsi, Chakraborty Sanjay, Ghosh Sanhita, Chakrabarti Amlan, Chakraborty Basabi. "A review on application of data mining techniques to combat natural disasters". *Ain Shams Engineering Journal* 9.3 (2018): 365-378.
17. Irawan Yuda. "Implementation Of Data Mining For Determining Majors Using K-Means Algorithm In Students Of SMA Negeri 1 PangkalanKerinci". *Journal of Applied Engineering and Technological Science (JAETS)* 1.1 (2019): 17-29.
18. Sumiran Keerthi. "An overview of data mining techniques and their application in industrial engineering". *Asian Journal of Applied Science and Technology* 2.2 (2018): 947-953.
19. Zeng, X., &Lv, Z. (2026). Task reconstruction and contrastive enhanced meta-learning for cross-machine bearing fault diagnosis. *Cyber-Physical Systems*, 12(??), 51–69.
20. Poornima S., and M. Pushpalatha. "A survey of predictive analytics using big data with data mining". *International journal of bioinformatics research and applications* 14.3 (2018): 269-282.



Wenjing Yin, born in 1995, is from Weifang, Shandong. She is a lecturer at Weifang Engineering Vocational College and holds a master's degree from Shandong University of Science and Technology, with a bachelor's degree from Shandong Management College. Her primary research focus is safety engineering. E-mail: yinwenjing@wfec.edu.cn



Sichao Lv was born in Tengzhou, Shandong, China, in 1989. He is a lecturer at Weifang Engineering Vocational College. He received the bachelor's degree from Weifang University of Science and Technology, his master's degree from Shenyang Aerospace University. His research interests include aircraft digital manufacturing and aircraft maintenance. E-mail: wgaecslc@163.com



Xiaofei Li, native of Weifang, Shandong Province, born in 1987. Lecturer and postgraduate at Weifang Engineering Vocational College; holds a bachelor's degree from Weifang University and a master's degree from Liaocheng University. His primary research focus is artificial intelligence and educational informatization.
Email: wfgc1988@163.com