

## *Retraction*

# **Retracted: Analysis of Abnormal Flight and Controllers Data Based on DBSCAN Method**

### **Security and Communication Networks**

Received 2 December 2022; Accepted 2 December 2022; Published 28 December 2022

Copyright © 2022 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Security and Communication Networks* has retracted the article titled “Analysis of Abnormal Flight and Controllers Data Based on DBSCAN Method” [1] due to concerns that the peer review process has been compromised.

Following an investigation conducted by the Hindawi Research Integrity team [2], significant concerns were identified with the peer reviewers assigned to this article; the investigation has concluded that the peer review process was compromised. We therefore can no longer trust the peer review process, and the article is being retracted with the agreement of the Chief Editor.

The authors do not agree to the retraction.

### **References**

- [1] C. Zeng, R. Wang, and Q. Zuo, “Analysis of Abnormal Flight and Controllers Data Based on DBSCAN Method,” *Security and Communication Networks*, vol. 2022, Article ID 7474270, 8 pages, 2022.
- [2] L. Ferguson, “Advancing Research Integrity Collaboratively and with Vigour,” 2022, <https://www.hindawi.com/post/advancing-research-integrity-collaboratively-and-vigour/>.

## Research Article

# Analysis of Abnormal Flight and Controllers Data Based on DBSCAN Method

Chen Zeng <sup>1</sup>, Rundong Wang,<sup>2</sup> and Qinghai Zuo<sup>2</sup>

<sup>1</sup>Civil Aviation Flight University of China Xinjin Flight College, Sichuan 611431, Xinjin, China

<sup>2</sup>Civil Aviation Flight University of China, Guanghan 618307, Sichuan, China

Correspondence should be addressed to Chen Zeng; zhanghengheng@cafuc.edu.cn

Received 28 March 2022; Revised 23 April 2022; Accepted 5 May 2022; Published 27 May 2022

Academic Editor: Muhammad Arif

Copyright © 2022 Chen Zeng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to quickly and timely analyze the airborne and controllers data of domestic ARJ21 (Advanced Regional Jet for 21st Century) aircraft and find out the existing flight problems and potential safety hazards, this paper proposes a DBSCAN (density-based spatial clustering of applications with noise) clustering analysis method for aircraft airborne and controllers data outlier detection to evaluate and monitor the flight status. The flight QAR (quick access recorder) data stored in the aircraft rapid data recorder are input to the DBSCAN clustering analysis algorithm to detect the flight parameters that are different from the normal range. Compared with the traditional airborne data analysis method, this method can realize the real-time analysis and prediction of data, improve the efficiency of data analysis, and find the potential security risks according to the analysis results and deal with them in time. In this paper, 1,102 ARJ21 aircraft operation data are used to test. The results show that the DBSCAN clustering anomaly data detection method based on density is fast and accurate in detecting the continuous parameters recorded in the flight process, and the display results are easy to analyze, which can predict the potential safety problems in time. The outliers detected by this method can provide support for the controller to detect the outliers and related flight risks in daily flights.

## 1. Introduction

Since the last century, the continuous improvement of air transport safety has been summarized from the lessons of each tragic air crash. However, this post summary method cannot be solved from the source of the accident, and the cost is huge. At present, air transportation has gradually changed from passive processing to active avoidance. By analyzing the airborne data of normal transportation flights, we can analyze abnormal data, identify risks, find out accident symptoms before accidents, and take measures to solve potential safety hazards. Many airlines predict and identify risks by regularly analyzing airborne flight QAR (quick access recorder) data and controller data [1]. Therefore, it is particularly important to process and analyze data through accurate and fast data analysis methods [2].

Usually, the aircraft will have abnormal flight and even cause accidents due to bad weather, air control, mechanical failure, pilot control technology, and other reasons. The

weather is often the primary factor leading to abnormal flights. But “weather reasons” actually include many situations, such as wind shear, thunderstorms, and so on. Aircraft airborne QAR data is the most important and direct evidence to judge the cause of flight accidents. The flight information recording system includes all states of the aircraft during flight, including speed, altitude, heading, deflection, roll, and other data, that are an important basis for the in-depth restoration of the cause of the accident. Among them, there are 16 kinds of important required data, such as aircraft acceleration, attitude, thrust, fuel volume, the position of the control surface, and so on. Usually, after a flight, the relevant personnel will replay the records to reproduce the errors or faults that have been found. Maintenance personnel can easily find the location of the fault by using it. Pilots can use it to check the flight performance and operational deficiencies of the aircraft and improve flight technology. In case of an aircraft crash, this recording system will become the most direct basis for accident analysis.

Due to a large number of flight data parameters, the method of visualizing QAR data is usually used to analyze the abnormal values. The distance between the midpoint of the feature space and the average value of the distribution is used to judge whether the value in the flight data is an abnormal value [3, 4]. With the maturity of data mining technology, using data mining technology to extract outliers from flight data has gradually been widely used in data analysis [5].

The density-based cluster data analysis method identifies and analyzes the potential safety problems existing in flight operation through new data-driven methods and airborne flight data so that relevant professionals can deal with potential safety hazards and avoid accident symptoms in time [6]. Compared with the traditional flight operation data processing methods, the density-based clustering data analysis method has the following advantages [7]:

- (1) It can analyze the causes and predict potential faults and safety problems through the outliers of data.
- (2) The airborne raw data of different models can be directly input and analyzed without initial configuration or adjustment, which improves the simplicity and rapidity of data processing.
- (3) Through the parallel operation, multiple types of parameters can be analyzed at the same time to identify multiple abnormal parameters in one flight data faster. If abnormal data is found, it will be output and displayed immediately. Then it is handed over to the expert team in relevant fields to manually identify the causes of problems and analyze potential risks [8, 9].

The state representation of each part of the aircraft during flight is recorded through QAR data, and then frequent itemset mining, association rule extraction, and similar sequence matching are carried out on the QAR data through data mining technology. Combined with existing fault data analysis, it can be effectively observed different characterization results of different flight states summarize different fault types according to different state characteristics, and thus, analysis results can provide reasonable and powerful technical support for the improvement of flight safety quality and operational efficiency important basis for maintenance, etc. It can effectively determine the existing hidden faults or the possibility of faults of some aircraft components and can remind airlines to carry out maintenance and repairs on the aircraft in time, keep abreast of the current safety situation, and formulate targeted plans to minimize accidents as much as possible. Hidden dangers can even be avoided, and the quality of flight safety can be further improved.

## 2. Experimental Details

The most representative density-based clustering analysis methods are the DBSCAN algorithm, options algorithm, and DENCLUE algorithm [10]. Compared with the traditional  $K$ -means algorithm, the biggest difference of

DBSCAN (density-based spatial clustering of applications with noise) is that it does not need to input the number of categories  $K$  and can find clusters of arbitrary shapes, rather than  $K$ -means, which is generally only used for convex sample set clustering. At the same time, it can also find outliers while clustering, which is similar to the birch algorithm [11]. DBSCAN algorithm is a typical density clustering algorithm. It is generally assumed that the category can be determined by the tightness of sample distribution. The samples of the same category are closely connected, and a cluster category is obtained by dividing the closely connected samples into one category [12]. By dividing all closely connected samples of each group into different categories, the final results of all clustering categories are obtained [13]. This paper uses the method of DBSCAN cluster analysis to analyze the outliers in airborne data. DBSCAN belongs to an unsupervised learning algorithm. The connotation of an unsupervised algorithm is to observe the unlabeled data set, automatically find the hidden structure and hierarchy, and find the hidden law in the unlabeled data. The application of the clustering model in data analysis: it not only can be used as a separate process to find the internal law of data but also can be used as the pre-exploration of other analysis tasks such as classification. DBSCAN is a density-based spatial clustering algorithm considering noise. Simply put, given a set of points, DBSCAN gathers the points close to each other (Euclidean distance), and it also marks the points in the low-density area as outliers.

- (1) Compared with the  $k$ -means method, DBSCAN does not need to know the number of cluster classes to be formed in advance.
- (2) Compared with the  $k$ -means method, DBSCAN can find clusters of arbitrary shapes.
- (3) At the same time, DBSCAN can identify noise points. It has good robustness to outliers and can even detect outliers.
- (4) DBSCAN is not sensitive to the order of samples in the database, that is, the input order of pattern has little effect on the results. However, for the boundary samples between cluster classes, it may swing according to which cluster class is detected first.
- (5) DBSCAN is designed to be used with the database to speed up the query of the region. For example, use the  $R^*$  tree.

According to Figure 1, the above definition can be easily understood. In Figure 1,  $\text{MinPts} = 5$  and red points are core objects because there are at least five samples in  $\text{MinPts} = 5$  neighborhood. The black sample is a noncore object. The samples with a direct density of all core objects are in the hypersphere centered on the red core object. If they are not in the hypersphere, they cannot have direct density. The core objects connected by green arrows in the figure form a sample sequence with a density of up to. In the  $\epsilon$ -neighborhood of these density reachable sample sequences, all samples are density connected to each other [14].

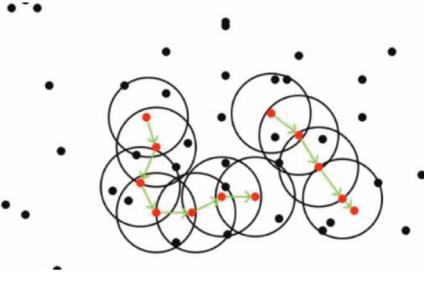


FIGURE 1: Example of density-based cluster analysis method.

### 2.1. DBSCAN Clustering Algorithm

- (1) Take the data point  $x_i$  of each frame of each parameter as the center of the circle and take  $r$  as the radius to form a data circle, which is called the adjacent area of the data point  $x_i$ .
- (2) Count the data points contained in the data circle. If the number of data points in a data circle exceeds the density threshold  $\text{MinPt}$ , the center of the data circle is recorded as the core data point, also known as the core object. If the number of points in the  $r$  neighborhood of a data point is less than the density threshold  $\text{minpts}$ , but the data point falls in the neighborhood of the core point, the data point is called a boundary point. Points that are neither core points nor boundary points are called noise points.
- (3) All data points in the  $r$  neighborhood of the core point  $x_i$  are the direct density of  $x_i$ . If  $x_j$  is directly from  $x_i$  density,  $x_k$  is directly from  $x_j$  density, and  $x_n$  is directly from  $x_k$  density, then  $x_n$  can be reached from  $x_i$  density. This property shows that the density reachability can be deduced from the transitivity of density direct.
- (4) If, for  $x_k$ , both  $x_i$  and  $x_j$  can be reached by  $x_k$  density, then  $x_i$  and  $x_j$  density are connected. Connecting the points connected by density together forms our cluster.

In short, if the total number of data points in the  $r$  neighborhood of a data point is less than the density threshold  $\text{MinPt}$ , then the point is a low-density data point. If it is greater than the density threshold  $\text{MinPt}$ , the point is a high-density data point. If a high-density data point is in the neighborhood of another high-density point, the two high-density points are directly connected, which is the core point. If a low-density point is in the neighborhood of a high-density point, connect the low-density point to the nearest high-density point, which is the boundary point. Low-density points that are not in the  $r$  neighborhood of any high-density point are abnormal points.

**2.2. DBSCAN Clustering QAR Data Outlier Detection Method Based on Density.** DBSCAN clustering QAR data analysis method is a common data mining technology, which is used to identify abnormalities in the detected data [15]. The first step is to convert QAR data into high-dimensional vectors to capture the multi-variate and temporal characteristics of

each flight. In the second step, the dimension of the above vectors is reduced to solve the problems related to data sparsity and multi-collinearity. The third step is cluster analysis of the above dimensionality reduction vectors. The adjacent vector group is a cluster, and the independent vector is abnormal data, as shown in Figure 2.

The 8-minute landing phase and the 3-minute takeoff phase are also called “dangerous 11 minutes” by the civil aviation industry, especially the landing phase [16]. Since 1980, 552 accidents have occurred in these 2 stages of global transport flights. Among these samples, 392 accidents occurred during landing, accounting for 71%, and 160 accidents occurred during takeoff, accounting for 29%. In this paper, the QAR data of domestic flights are used to analyze the 2 flight stages of takeoff and final approach [17].

**2.3. QAR Data Processing Process Based on DBSCAN Clustering Algorithm.** The QAR data processing process based on the DBSCAN clustering algorithm is shown in Figure 3.

The first is a data conversion. Since flight data is recorded in chronological order, it is necessary to convert flight data from time series to multi-dimensional vectors for storage.

Secondly, the multi-dimensional vector data in the first step is reduced to extract key data and remove data redundancy.

Finally, the low-dimensional data are analyzed according to the clustering principle to identify the core data points and discrete data points in the multi-dimensional vector space.

#### Step 1. Data conversion

Before data conversion, format and pre-process the original QAR flight data. Normalize different continuous flight parameters. The median value of the discrete parameter or the median value of the continuous parameter is transformed together with the Euclidean vector.

Then, the QAR data vector in Euclidean space is represented by multi-dimensional vector space, and the abnormal events are marked according to the time series. Starting from the marking event, each parameter type in QAR data is sampled at a time frequency of 1 Hz, making the data of different flights comparable. All sampling values are arranged in time series to form a flight vector, as shown in the following equation:

$$\mathbf{v} = [x_1^1, x_2^1, \dots, x_n^1, \dots, x_j^i, \dots, x_n^m], \quad (1)$$

where data  $x_j^i$  is the value collected by the  $i$ -th airborne flight parameter at time  $j$ . The dimension of the data vector collected in each period is  $m \times n$ , where  $m$  represents the sampling times of flight data during this period and  $n$  represents the number of flight parameters contained in each sampling of flight data. The Euclidean distance is used to evaluate the difference between data vectors in each period of time: if the Euclidean distance is farther, the similarity between vector data is smaller [18].

The calibration range of takeoff and final approach phase is shown in Figure 4. In the takeoff phase, the range is 180 s after the time point when the engine thrust reaches the

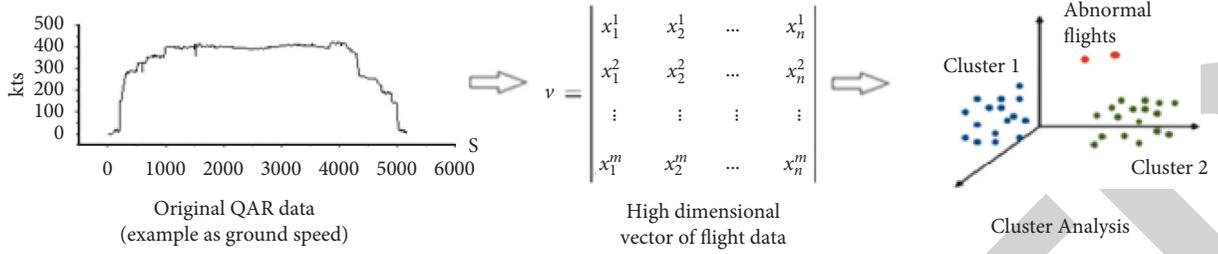


FIGURE 2: QAR data processing steps based on the DBSCAN clustering algorithm.

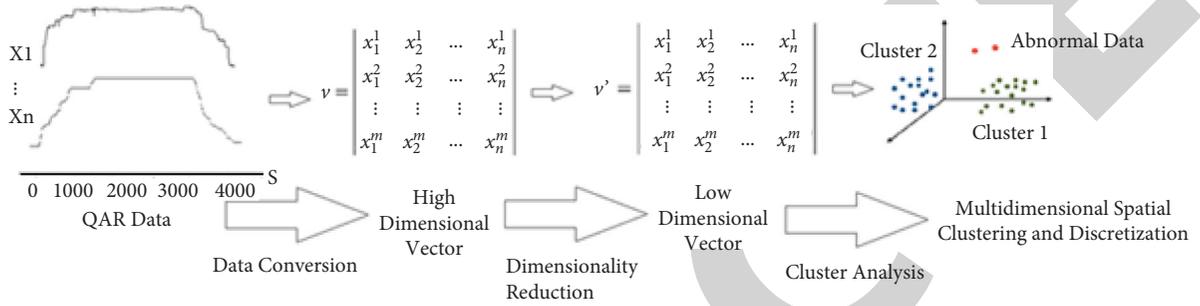


FIGURE 3: The QAR data processing process.

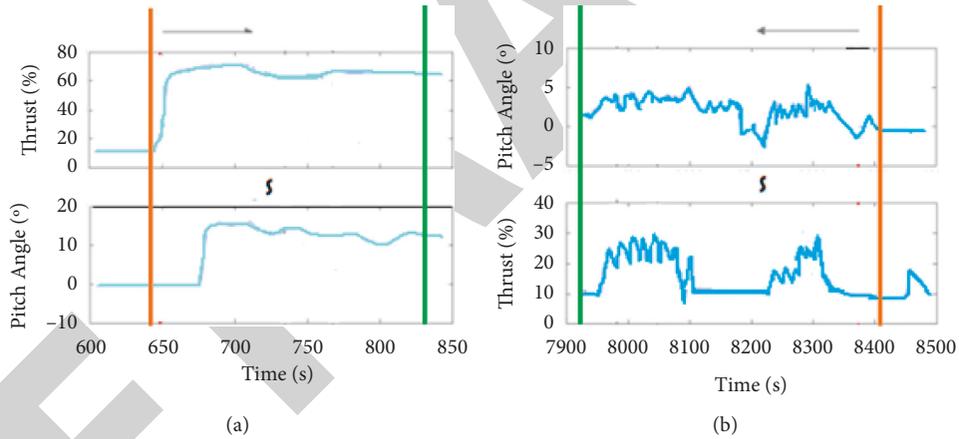


FIGURE 4: Sampling time range in takeoff and approach phase: (a) takeoff phase range and (b) final approach phase range.

takeoff power. For the final approach stage, the range is 480 s ahead of the time point when the wheel touches the ground. For all parameters, the flight process is sampled at a fixed 1 second time interval.

### Step 2. Dimensionality reduction

The data vector generated after data format conversion is a vector space with tens of millions of dimensions. We use principal component analysis (PCA) to merge data sparsely distributed in different dimensions. In this way, the data are transformed into a new orthogonal coordinate system to reduce the dimension of the data vector [19, 20]. In the new system, the coordinates are sorted by the amount of embedded information, and most of the information at the top of each parameter is retained, so the dimension is reduced. In this paper, the data within 80% variance of the first component in the captured data is retained [21] as follows:

$$\sqrt{\frac{\sum_{i=1}^K \lambda_i^2}{\sum_{i=1}^N \lambda_i^2}} > 80\%, \quad (2)$$

where  $\lambda$  is the variance of principal component analysis and the total number of principal components, that is, it is equal to the original dimension. Calculation results is the retained principal component score. In this paper, the vector dimension in takeoff phase is reduced from 14,400 to 75, and that in the approach phase is reduced from 41,760 to 89.

### Step 3. DBSCAN cluster analysis based on density

Through the DBSCAN algorithm for clustering analysis of data vectors [22], we can get the number of clusters in the data vector and identify and deal with the cluster vectors with outliers. DBSCAN algorithm forms MinPt clustering

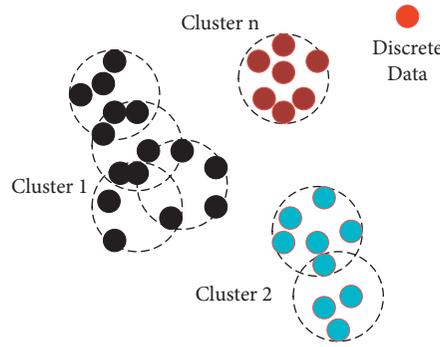


FIGURE 5: DBSCAN cluster analysis based on density.

vector through density restriction. If there is the minimum number of data points required to form a clustering vector in a circle with radius  $r$ , the circle forms a clustering vector. Each clustering vector increases the dimension of the clustering vector by finding the data points in its adjacent region that meet the same density standard. As long as the density standard is met, new clusters will be formed. The algorithm marks the data points that do not belong to any clustering as abnormal data points. As long as the density standard is met, new clusters will be formed. The algorithm marks the points that do not belong to any clustering as outliers as shown in Figure 5.

The algorithm only needs two parameters  $\epsilon$  and MinPt as inputs. MinPt is the minimum number of flights in the data set; the size of  $\epsilon$  value determines the proportion of identifiable flight anomaly data [23]. Through sensitivity analysis, we can determine the size of  $\epsilon$  and MinPt, as shown in Figure 6. Firstly, we determine  $\epsilon$  as a certain value, bring the value of MinPt into the DBSCAN algorithm from small to large, and get the most appropriate value of MinPt after running for many times. After the MinPt value is determined, set the  $\epsilon$  value according to the severity of the abnormal data to be identified [24].

### 3. Results and Discussion

The data set contains 1,102 ARJ21 flight operation data. There are 87 flight parameters in the data set, including but not limited to the engine thrust, aircraft altitude, speed, acceleration, attitude, control surface position, ambient pressure, temperature, and so on. In the final approach phase and takeoff phase, three groups of detection thresholds of 1%, 3%, and 5% are used to detect abnormal flights.

**3.1. Data Pre-Processing and Algorithm Setting.** Firstly, we set the sampling frequency of flight QAR data during takeoff and approach according to the model of airborne equipment. In the takeoff phase, the observation value is obtained every 1 second from the time the pilot applies the takeoff thrust to 180 seconds. In the approach phase, the flight QAR data in the approach phase is collected at the frequency of 1 Hz by pushing back 480 s from the wheel touchdown time. This step retains the data within

80% of the variance of the first component in the captured data, with the dimension from 14400 to 75.

The input parameters  $\epsilon$  and MinPt of the clustering algorithm are set based on sensitivity analysis. The results show that the selection value between 3 and 15 has no significant impact on the results, but when  $\epsilon$  increases, fewer flights are determined as abnormal values, as shown in Figure 6. Therefore, the MinPt is set to a value of 5, and the value of  $\epsilon$  is set to find the first 1, 3, and 5% abnormal values.

**3.2. Abnormal Results in Approach Phase.** In the approach phase, 3 abnormal flights are detected at  $\epsilon = 1\%$  that appears 8 times at  $\epsilon = 3\%$  and  $\epsilon = 5\%$  that appears 15 times. In Figure 7, two abnormal flight data are drawn with reference to the normal flight. The line represents the parameter value of the abnormal flight; the strip represents the range of the flight parameter value of the normal flight; the dark blue strip represents the distribution range of 25%~75% of the normal flight parameter value of all flights; and the light blue strip contains the distribution range of 5%~95% of the normal flight parameter value of the flight. The dark blue area covers 50% of the data, while the light blue area covers 90% of the data.

Figure 7(a) shows the flight situation of low altitude slow thrust abnormal approach. The vertical profile of the aircraft remains below the normal glide path until it returns to the normal glide path 3.7 km before the runway entry point. Before 5.56 km from the runway threshold, the ground speed is lower than the normal flight condition. The flap is set to the landing configuration at least 11.12 km before landing, which is earlier than normal. In order to join the normal glide path 5.56 km to 3.7 km before landing, the pilot used a higher engine thrust and pitch attitude than normal.

Figure 7(b) shows an approach with an abnormal flap angle setting. From 11.1 km before the runway threshold to the runway threshold position, the flap setting is maintained at  $25^\circ$ , while the  $30^\circ$  flap angle is used in the normal flight-landing configuration. In the final stage of the approach, the engine thrust is reduced, and the main indexes of approach performance, altitude, airspeed, and pitch are within 80% of the normal range.

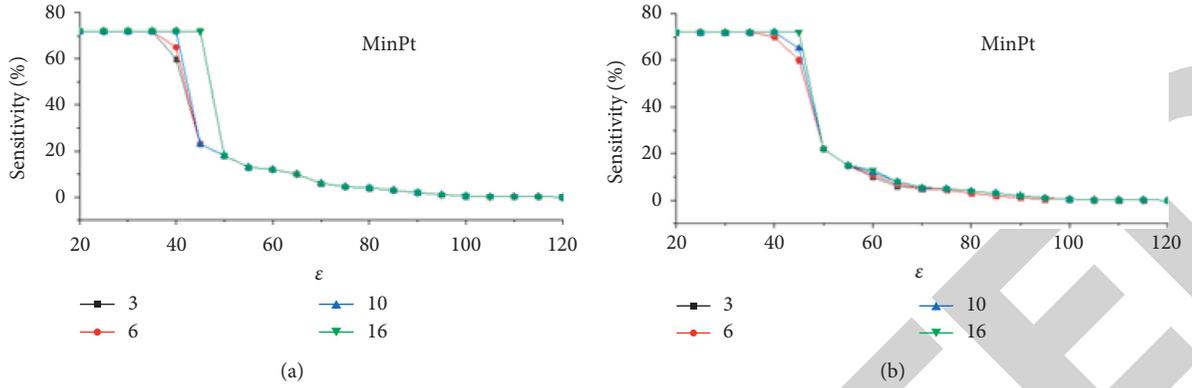


FIGURE 6: MinPt and  $\epsilon$  sensitivity to test results: (a) approach stage and (b) takeoff phase.

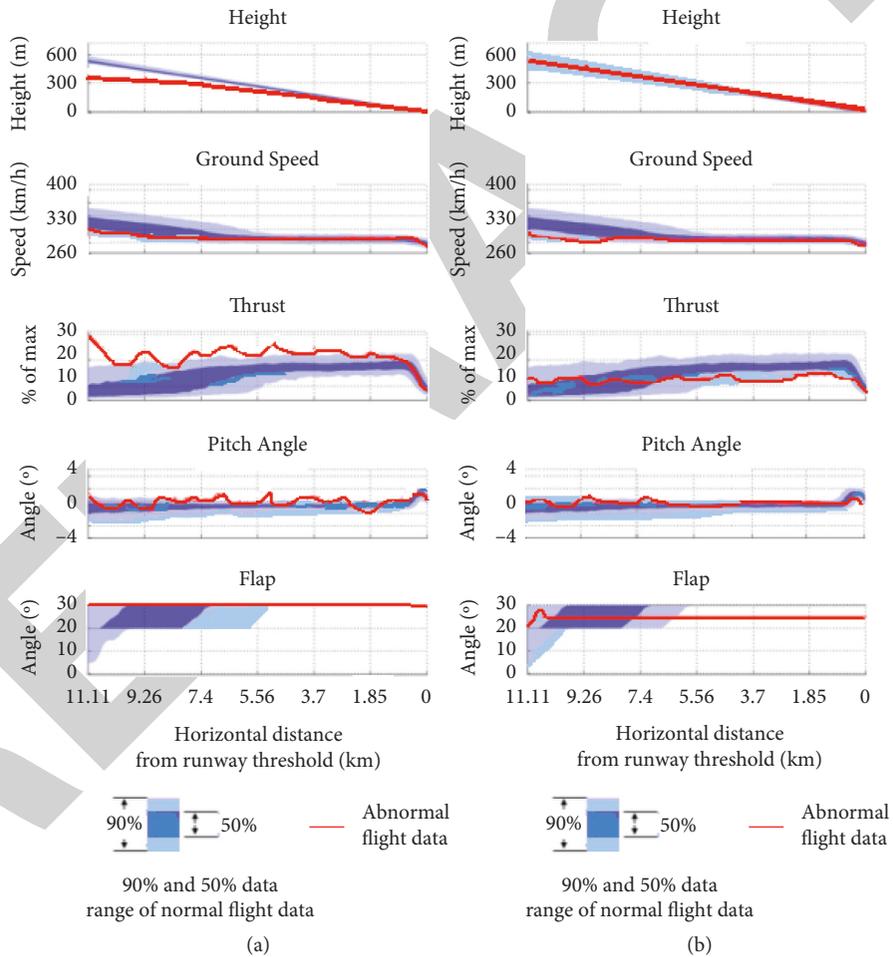


FIGURE 7: Two abnormal flights in the approach phase: (a) flight data of low altitude abnormal thrust and low-speed approach and (b) flap angle setting abnormal data.

3.3. *Abnormal Results during Takeoff.* In the final approach phase, four abnormal flights are detected at  $\epsilon = 1\%$ ,  $\epsilon = 3\%$  that appears 9 times, and  $\epsilon = 5\%$  that appears 22 times. Taking the normal flight as a reference, two abnormal flights are illustrated. Figure 8(a) shows an abnormal flight with

full-load thrust reduction takeoff. The aircraft accelerates slowly; it takes longer to reach the required takeoff speed; the taxiing distance becomes longer; and the ground clearance time is delayed. The pitch angle of the roller lift off the ground is 15 degrees, and the pitch angle reaches 15 degrees

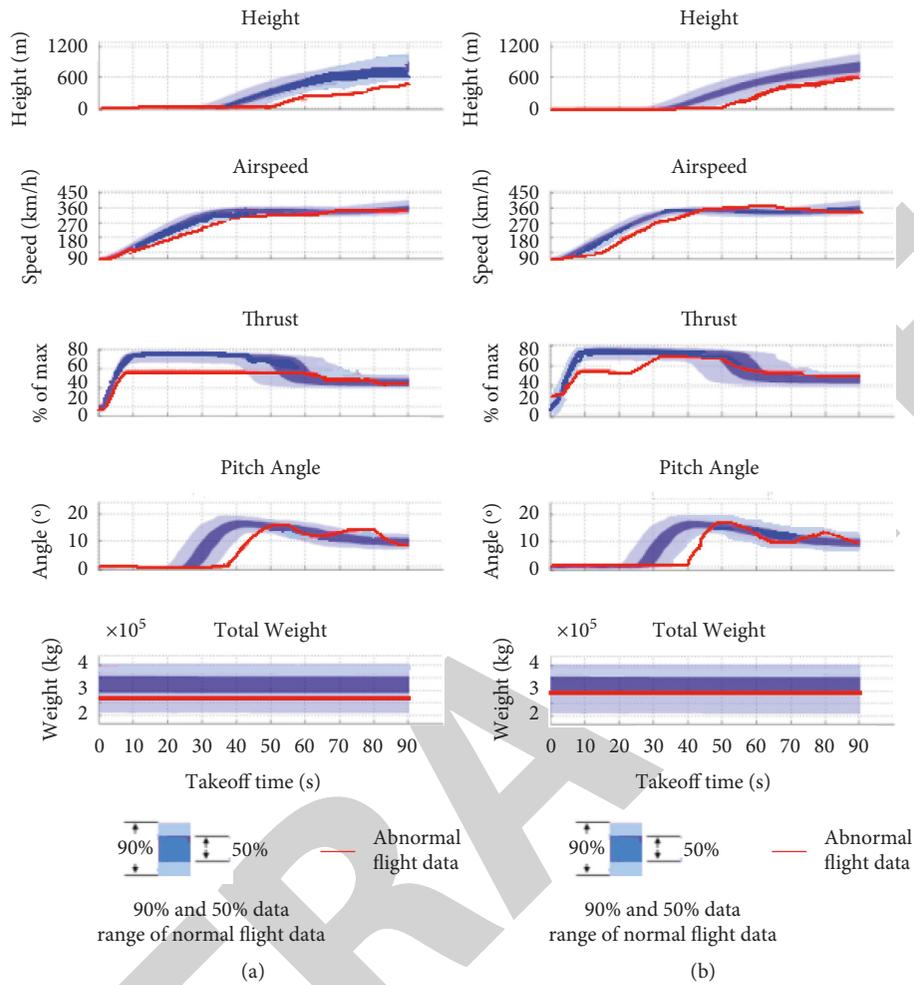


FIGURE 8: Two abnormal flights during takeoff: (a) reduce thrust takeoff and (b) variable thrust takeoff.

again 80 seconds after takeoff. Figure 8(b) shows the case of variable thrust takeoff. In the first 20 seconds, the reduced thrust takeoff is used. After the lifting wheel leaves the ground, the engine thrust power setting returns to the normal level, but the climb rate and acceleration are still lower than the normal flight.

#### 4. Conclusion

This paper proposes a new DBSCAN cluster analysis method to analyze these flight data and extract relevant information according to the requirements. The abnormal data detection method can use cluster analysis to automatically detect abnormal flight data from conventional airline flights, compare and analyze it with normal flight data, and draw relevant conclusions. Then, the abnormal information obtained shall be submitted to experts in relevant fields to determine the existing problems and potential risks. The data analysis results show that the DBSCAN airborne data clustering analysis method can identify a small amount of abnormal data from a large number of flight data and even find potential high-risk factors. Cluster data analysis is more sensitive to detecting the anomalies of continuous parameters. However, the

DBSCAN clustering analysis method has poor clustering quality if the density of the sample set is uneven and the cluster spacing is very different. If the sample set is large, the clustering convergence time is longer. Parameter adjustment is slightly more complex than traditional clustering algorithms such as  $k$ -means. It mainly needs to jointly adjust the parameters of distance threshold  $\epsilon$  and neighborhood sample number threshold  $MinPt$ . The different parameter combinations have a great impact on the final clustering effect. The future work direction is to develop reinforcement learning ability. Abnormal flights with the same symptoms can be automatically classified without repeated human evaluation, so as to improve the efficiency and accuracy of data analysis.

#### Data Availability

The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This article was published with the support of the following funds: National Key Research and Development Plan (2021YFF0603904), air traffic management information standard technology research and application verification, and Sichuan Science and Technology Plan, key research and development project (2021YFS0319) and key technology research of general aviation in mountain fire rescue.

## References

- [1] Y. Aref, K. Ö. S. E. Cemal, Y. Asef, and S. Amir, "Fuzzy-DBSCAN algorithm for morphological and overlapping datasets," *Journal of Systems Engineering and Electronics*, vol. 31, no. 6, pp. 1245–1253, 2020.
- [2] M. Li and X. Bi, "A method of two-stage clustering learning based on improved DBSCAN and density peak algorithm," *Computer Communications*, vol. 167, pp. 25–31, 2021.
- [3] C. Wang and W. Li, "A robust segmentation framework for closely packed buildings from airborne LiDAR point clouds," *International Journal of Remote Sensing*, vol. 41, no. 14, pp. 128–133, 2020.
- [4] A. Starczewski, P. Goetzen, and J. Meng, "A new method for automatic determining of the DBSCAN parameters," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 10, no. 3, pp. 256–261, 2020.
- [5] S. K. Ku and H. Baik, "Analysis of runway occupancy time using ADS-B message about landing airplane," *The Journal of Advanced Navigation Technology*, vol. 20, no. 3, pp. 412–417, 2016.
- [6] S. M. Stigler, "Francis galton's account of the invention of correlation," *Statistical Science*, vol. 11, no. 3, pp. 321–324, 2019.
- [7] M. Schwabacher and N. B. Oza, "Unsupervised anomaly detection for liquid-frphmonitoring," *Journal of Aerospace Computing, Information, and Communication*, vol. 6, no. 7, pp. 464–482, 2009.
- [8] S. A. A. Hofmeyr, "Intrusion detection using sequences of system calls," *Journal of Computer Security*, vol. 2, pp. 151–180, 2018.
- [9] C. Möller-Levet, "Fuzzy clustering of short time-series and unevenly distributed sampling points," *Advances in Intelligent Data Analysis V*, vol. 5, pp. 330–340, 2013.
- [10] Ge Qian, S. Hou, and W. Zhao, "GPS positioning algorithm based on Kalman filter and improved DBSCAN clustering combination," *Global positioning system*, vol. 46, no. 1, pp. 28–35, 2021.
- [11] W. Pan, K. Liu, R. Wang, and Q. Zuo, "Research on the architecture of Civil Aviation ATC big data processing platform," *Computer applications and software*, vol. 37, no. 06, pp. 48–52, 2020.
- [12] Z. Yuan and H. Zhuang, "Hierarchical clustering algorithm for variable scale density data," *Computer Applications*, vol. 40, no. S2, pp. 54–59, 2020.
- [13] V. E. Atasoy and C. Cetek, "Enhanced cruise range prediction for narrow-body turbofan commercial aircraft based on QAR data," *Aeronautical Journal*, vol. 12, pp. 1286–1294, 2020.
- [14] W. Li, Y. Li, and Y. Bian, "An improved method of code package hierarchical reconstruction based on DBSCAN algorithm," *Computer engineering and science*, vol. 43, no. 4, pp. 689–696, 2021.
- [15] W. Pan, J. Ren, R. Wang, Li Chen, and H. Liu, "Evaluation and analysis of control quality based on key indicators of control operation," *Aviation computing technology*, vol. 50, no. 2, pp. 1–6, 2020.
- [16] Y. Wang, H. Liu, P. Song, and Z. Hu, "Abnormal operation data cleaning method of wind turbine based on multi-stage progressive identification," *Renewable Energy*, vol. 38, no. 11, pp. 1470–1476, 2020.
- [17] C. Zi Tian, W. Yang, D. Yang, Y. Wang, and S. Sun, "Analysis and Research on student behavior based on comprehensive data of colleges and Universities under the background of K-means and DBSCAN clustering algorithm," *Science and technology innovation*, no. 32, pp. 86–88, 2020.
- [18] J. Jin and T. Wu, "Research on statistical model of runway occupation time of landing aircraft," *Electronic components and information technology*, vol. 3, no. 10, pp. 98–101, 2019.
- [19] W. Pan, K. Liu, X. Zhu, X. Wang, and S. Wang, "Research on the architecture of civil aviation airport emergency rescue big data processing platform," *Computer and digital engineering*, vol. 49, no. 2, pp. 280–283, 2021.
- [20] D. Zhang, "Evaluation model of runway occupancy interval of paired approach aircraft," *Henan science and technology*, vol. 16, no. 11, pp. 111–112, 2016.
- [21] K. Hua, "Application and Prospect of flight quality big data analysis based on QAR," *Network security technology and application*, vol. 27, no. 3, pp. 116–117, 2021.
- [22] H. Shi, "Implementation and development of fault monitoring by airlines using ACMS and QAR data," *Aviation maintenance and engineering*, vol. 2021, no. 2, 20 pages, 2021.
- [23] H. Shi, J. Yang, L. Lin, Z. Chen, and Y. Wu, "Analysis of air turbulence based on QAR data and manual records of Shenzhen Airlines from 2016 to 2018," *Journal of civil aviation*, vol. 5, no. 1, pp. 30–36, 2021.
- [24] Z. Duan, H. Du, and P. Zhang, "Feature extraction of QAR data based on qar2vec model," *Chinese Journal of safety science*, vol. 31, no. 1, pp. 145–152, 2021.