International Civil Aviation Organization

THE SIXTH MEETING OF MODE S DOWNLINKED AIRCRAFT PARAMETERS WORKING GROUP (MODE S DAPs WG/6)

Bangkok, Thailand, 28 – 30 March 2023

Agenda Item 10: Future work programme and any other business

A GNSS INTERFERENCE DETECTION ALGORITHM BASED ON MACHINE LEARNING

(Presented by China)

SUMMARY

This paper presents a novel approach for identifying ADS-B signal interference based on a multi-dimensional evaluation system constructed through feature engineering. Using multi-variate logistic regression, the approach determines whether each message is interfered with, and then utilizes the central limit theorem to construct a regional interference index. This work builds upon our previous research in the area and offers a promising approach to detecting and mitigating ADS-B signal interference.

1. INTRODUCTION

1.1 Automatic Dependent Surveillance-Broadcast (ADS-B) is a surveillance technology that relies heavily on Global Navigation Satellite Systems (GNSS) for accurate position and velocity information. However, when GNSS is jammed, the integrity and accuracy of ADS-B data can be seriously affected, which poses a threat to aviation safety. Therefore, analyzing and processing ADS-B data can be a feasible method to detect GNSS interference.

1.2 Our research on the relationship between changes in ADS-B data and GNSS interference began in 2020. Based on our previous studies(**MODE S DAPs WG/5-IP/10-*"A GPS INTERFERENCE IDENTIFICATION METHOD BASED ON ADS-B DATA"***), we concluded that when the rate of velocity and position changes, NIC, NACp, all change simultaneously, it can be considered as ADS-B abnormal due to GNSS interference. If only one or two of these features change, the abnormal data may not be caused by GNSS interference.

1.3 In this working paper, we further analyze the relationships between track angle(TA), flight level(FL), ADS-B equipment version, and GNSS interference in ADS-B data, and establish a more suitable data model through a larger amount of data analysis to improve the experimental results.

1.4 Instead of using the traditional method of calculating indicators, we adopt a data-driven approach, selecting specific feature values and training a machine learning model using historical data. Specifically, we use a multivariate logistic regression model and use the trained machine learning model to detect abnormal data.

1.5 Finally, based on the results of logistic regression, we establish an indicator that indicates whether an area is affected by interference according to the central limit theorem.

2. DISCUSSION

2.1 For the sake of completeness, we provide a brief retrospective on the previous working paper. In our previous study, we conducted experiments using both simulated laboratory scenarios and real-world event data provided by a data center. We concluded that when the rate of change in velocity, rate of change in position, NIC, and NACp occur simultaneously, it can be considered an ADS-B abnormality caused by GNSS interference, take Flight CSN3444 as an example. We then established a comprehensive evaluation index using different mapping functions and ultimately rated the index based on historical experience. While this method represents an improvement over previous single-indicator methods and to some extent eliminates abnormal data caused by other reasons, it is still subject to a certain degree of error due to the manual formulation of the index rating based on historical data. Thus, in this paper, we employ machine learning methods to eliminate human error to the greatest extent possible.

【Flight CSN3444】：

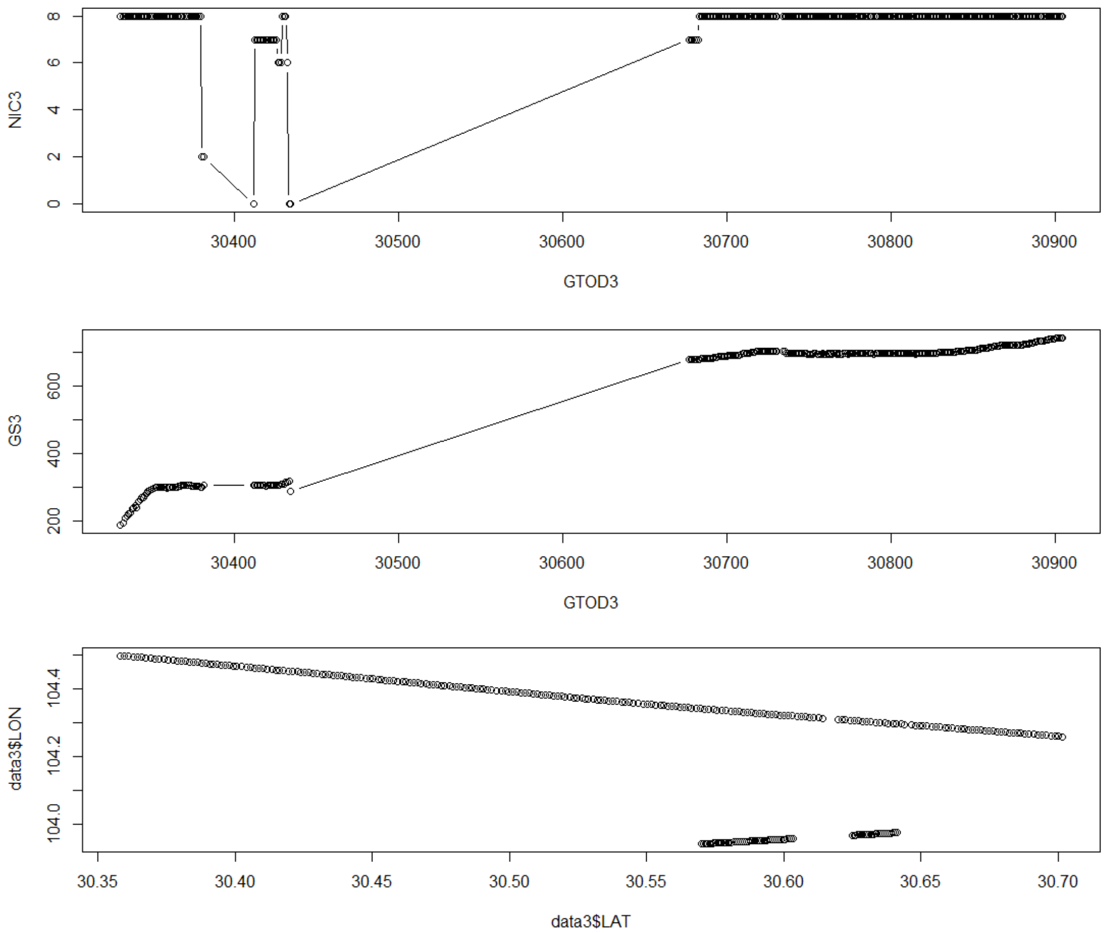

Figure 1: The related index change graph

2.2 Feature engineering

Based on previous work, the present study involved an in-depth analysis of additional ADS-B data items before training the machine learning model. Specifically, we analyzed different versions of ADS-B device data, focusing on five main aspects, including JUMP (large-scale position deviation), BACK (position bounce), Z20 (20 or more swings), FL (barometric altitude jump), and GH (geometric altitude jump). The analysis results are presented in the figure below, and the following conclusions were drawn:

a) The proportion of abnormal aircraft in the DO-260A version is high in all aspects;

b) In terms of large-scale deviation, bounce, Z swings of 20 or more, and barometric altitude jump, the proportion of abnormal aircraft in the DO-260 version is higher than that in the DO-260B version;

c) In terms of geometric altitude jump, the proportion of abnormal aircraft in the DO-260B version is higher than that in the DO-260 version.

Table 1: the comparison of abnormal data occupation between three ADS-B equipment version

Version	JUMP	BACK	Z20	FL	GH
DO-260	44.50%	46.22%	36.07%	22.02%	50.63%
DO-260A	53.11%	43.50%	39.55%	23.73%	72.88%
DO-260B	35.16%	40.10%	31.08%	10.90%	73.26%

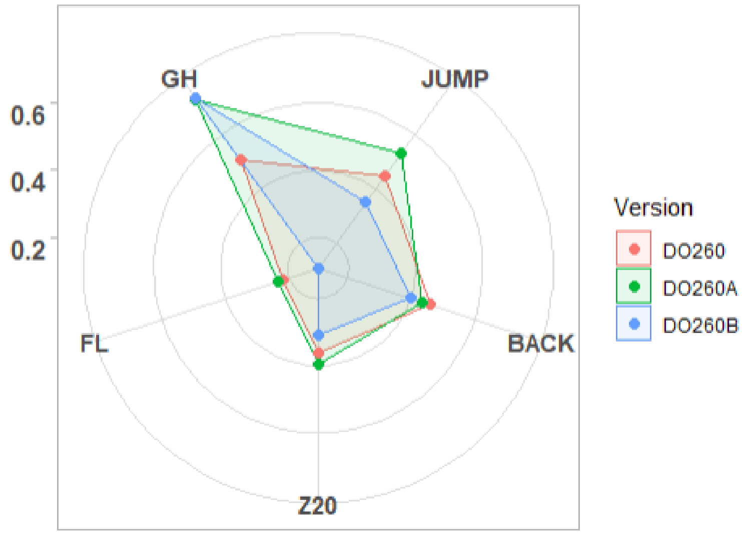

Figure 2: The comparison of various abnormal ratios of three ADS-B equipment version

Also, in the field of civil aviation, the majority of interference sources are of the suppression type, which is characterized by their low cost and easy accessibility. Typically, these interference sources are placed on the ground, and their range of influence is generally limited. Therefore, data in the route area is not usually susceptible to interference. However, it is crucial to fully consider altitude as an intrinsic value to ensure the safety of aircraft operations. The altitude factor plays a vital role in determining the accuracy and reliability of aircraft data. Hence, it is imperative to incorporate altitude as a critical parameter in the evaluation of interference.

In accordance with the fundamental principle of satellite positioning, the heading angle difference of an aircraft will cause the angle between the aircraft's antenna and the satellite to vary depending on which satellite it is working with. Aircraft that work with fewer satellites and have a smaller angle between their antenna and the satellite are more susceptible to interference. Thus, it is important to consider the heading angle as a factor when assessing the susceptibility of an aircraft to interference.

After comprehensive experiments and analyses, we have selected message update interval, velocity change rate, position change rate, NIC, NACp, SIL, FL, TA, and ADS-B version as the feature values for identifying interference in subsequent research.

2.3 Logistic regression

Logistic regression is a method that involves establishing a cost function for regression or classification problems, and iteratively solving optimal model parameters through an optimization method, followed by testing to verify the quality of the solution model. Logistic regression is commonly used in classification problems, where it can be used to classify binary data and solve multi-classification problems using multinomial logistic regression. Hypothesis representation in logistic regression is based on multivariable linear regression and involves obtaining a linear combination of multiple variables. For the two-classification problem, the hypothesis function is represented by the nonlinear Sigmoid function, which standardizes the value of any range into the [0,1] interval. The calculation of the hypothesis function in logistic regression is shown in the following formula.:

$$g(z) = \frac{1}{1+e^{-z}}$$

$$h_\theta(x) = g\left(\theta^T X\right)$$

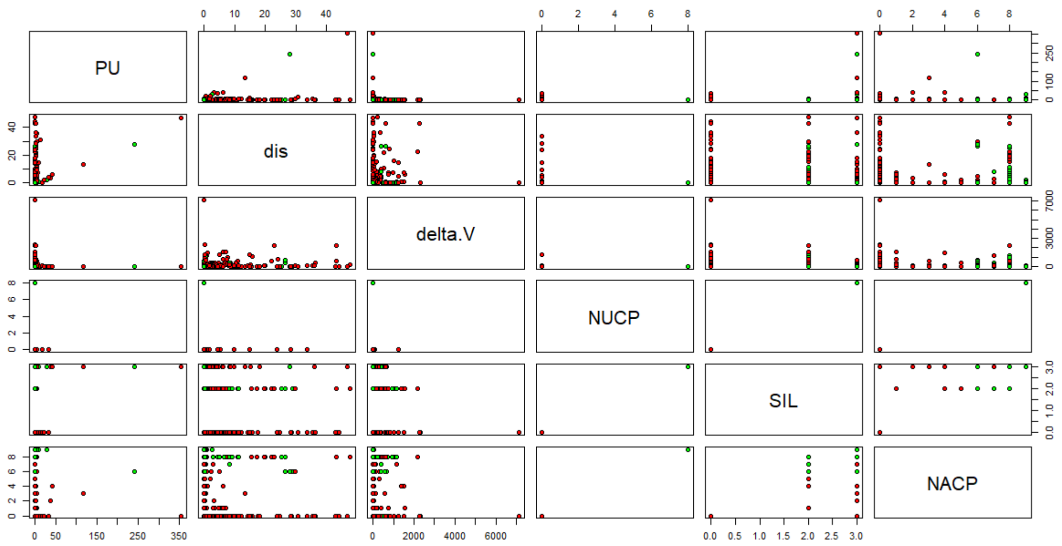

Figure 3: The diagram of the correlation between multiple characteristics

After identifying the necessary features and selecting a specific model, we trained the model parameters using a large amount of actual data. To eliminate abnormal data that may not be caused by GNSS interference, we selected relevant flight data within areas and times confirmed by data centers to have experienced interference. These data were labeled based on established rules. We divided the collected data into training, testing, and validation sets. Our experiments showed that the multivariate logistic regression method achieved a classification accuracy rate of over 95%. Overall, our data-driven approach using multiple indicators has shown promising potential for detecting GNSS interference and enhancing aviation safety.

2.5 Statistical index

Through the use of the trained multivariate logistic regression model described above, we were able to determine whether each ADS-B trajectory point was subject to interference. The results of our analysis are shown in the accompanying figure, which indicates the number of points in a given trajectory or region that are suspected to have experienced interference. However, an important issue to consider is that flight densities differ across regions, so it is not reasonable to rely solely on the number or proportion of points that appear to be affected to make determinations of interference.

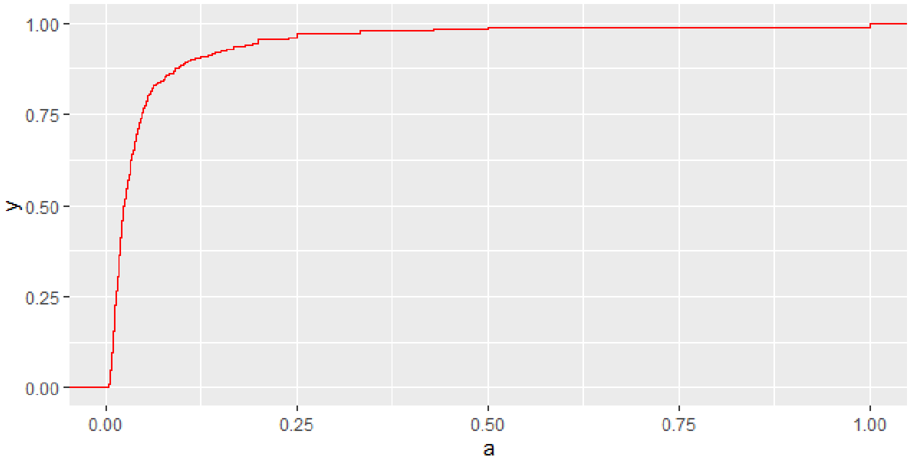

Figure 4: The PDF of statistics A

To address this issue, we employed statistical measures to establish indicators. Specifically, we calculated the cumulative distribution function of the number and proportion of points affected within a defined range, with the example of a range of 0.5 degrees longitude by 0.5 degrees latitude. We then used this information to determine the distribution values corresponding to the 90th, 95th, 96th, 98th, and 99th percentiles of the a-value probability density function (PDF). Next, we identified the corresponding A_require values for each percentile value. Finally, we calculated the statistical measures using the A_require values for analysis.

$$A = \frac{\sqrt{N_{UI}} \times (A_{means} - A_{req})}{\sqrt{N_{UI} \times A_{req} \times (1 - A_{req}) + \frac{A_{req}^2}{2}}}$$

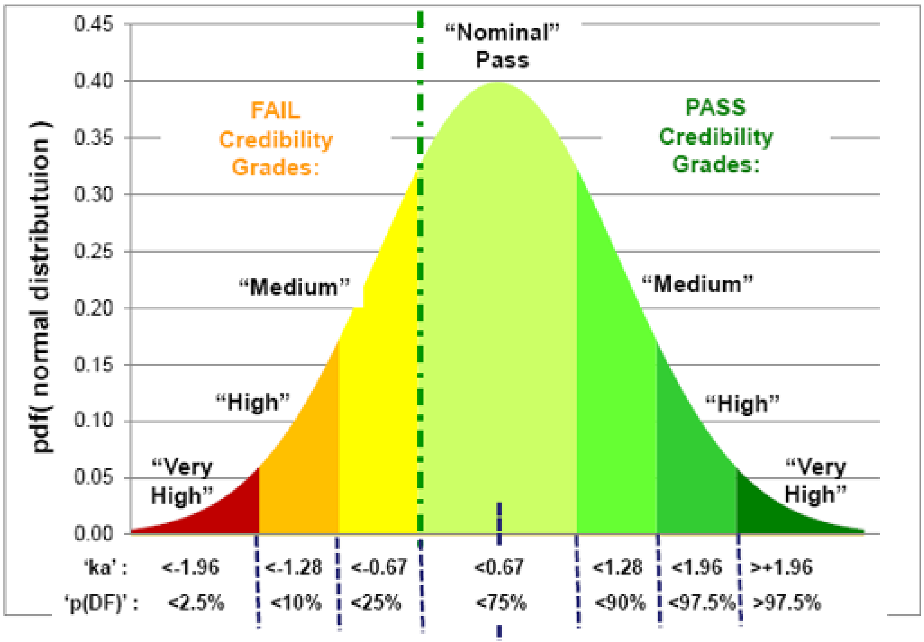

Figure 5: The diagram of the normal distribution of statistics A

Based on the central limit theorem, we assume that the distribution of the indicators follows a normal distribution when the data volume is sufficiently large. The degree of interference represented by the indicator calculation results can be divided into categories as shown in the figure below. It is important to note that the classification of interference levels should be interpreted in the context of the specific application and use case, and may require further validation and calibration. Nonetheless, this approach provides a useful framework for quantifying and assessing ADS-B interference systematically and objectively.

2.6 Conclusion

The Automatic Dependent Surveillance-Broadcast (ADS-B) technology relies on Global Navigation Satellite Systems (GNSS) for accurate position and velocity information. However, GNSS jamming can severely affect the integrity and accuracy of ADS-B data, posing a threat to aviation safety. In this working paper, we have analyzed and processed ADS-B data to detect GNSS interference, using a data-driven approach, including a multivariate logistic regression model trained on historical data to detect abnormal data. We have established an indicator that indicates whether an area is affected by interference, based on the central limit theorem, with an accuracy rate of up to 0.95. Furthermore, we analyzed additional ADS-B data items and incorporated altitude and heading angle as critical parameters to improve the accuracy and reliability of the evaluation of interference. The results of our experiments and analyses demonstrate the effectiveness and potential of using machine learning methods to improve the detection of GNSS interference and enhance aviation safety.

3. ACTION BY THE MEETING

a) note the information contained in this paper; and

b) discuss any relevant matter as appropriate.

– – – – – – – – – – – – –

Appendix

NACp	95% Horizontal and vertical precision boundaries (EPU和VEPU)
0	EPU ≥ 18.52 km (10 NM)
1	EPU < 18.52 km (10 NM)
2	EPU < 7.408 km (4 NM)
3	EPU < 3.704 km (2 NM)
4	EPU < 1852 m (1NM)
5	EPU < 926 m (0.5 NM)
6	EPU < 555.6 m (0.3 NM)
7	EPU < 185.2 m (0.1 NM)
8	EPU < 92.6 m (0.05 NM)
9	EPU < 30 m and VEPU < 45 m
10	EPU < 10 m and VEPU < 15 m
11	EPU < 3 m and VEPU < 4 m

NIC	Horizontal and vertical inclusion boundaries
0	RC ≥ 37.04 km (20 NM)
1	RC ＜ 37.04 km (20 NM)
2	RC ＜ 14.816 km (8 NM)
3	RC ＜ 7.408 km (4 NM)
4	RC ＜ 3.704 km (2 NM)
5	RC ＜ 1852 m (1 NM)
6	RC ＜ 1111.2 m (0.6 NM)
7	RC ＜ 370.4 m (0.2 NM)
8	RC ＜ 185.2 m (0.1 NM)
9	RC ＜ 75 m and VPL ＜ 112 m
10	RC ＜ 25 m and VPL ＜ 37.5 m
11	RC ＜ 7.5 m and VPL ＜ 11 m