

ADS-B Anomaly Data Detection Model Based on Deep Learning and Difference of Gaussian Approach

WANG Ershen^{1*}, SONG Yuanshang¹, XU Song¹, GUO Jing², HONG Chen³,
QU Pingping¹, PANG Tao¹, ZHANG Jiantong⁴

1. School of Electronic and Information Engineering, Shenyang Aerospace University, Shenyang 110136, P.R. China;

2. China Academy of Civil Aviation Science and Technology, Beijing 100028, P.R. China;

3. College of Robotics, Beijing Union University, Beijing 100101, P.R. China;

4. China Transportation Telecommunications and Information Center, Beijing 100011, P.R. China

(Received 28 May 2020; revised 30 June 2020; accepted 25 July 2020)

Abstract: Due to the influence of terrain structure, meteorological conditions and various factors, there are anomalous data in automatic dependent surveillance-broadcast (ADS-B) message. The ADS-B equipment can be used for positioning of general aviation aircraft. Aim to acquire the accurate position information of aircraft and detect anomaly data, the ADS-B anomaly data detection model based on deep learning and difference of Gaussian (DoG) approach is proposed. First, according to the characteristic of ADS-B data, the ADS-B position data are transformed into the coordinate system. And the origin of the coordinate system is set up as the take-off point. Then, based on the kinematic principle, the ADS-B anomaly data can be removed. Moreover, the details of the ADS-B position data can be got by the DoG approach. Finally, the long short-term memory (LSTM) neural network is used to optimize the recurrent neural network (RNN) with severe gradient reduction for processing ADS-B data. The position data of ADS-B are reconstructed by the sequence to sequence (seq2seq) model which is composed of LSTM neural network, and the reconstruction error is used to detect the anomalous data. Based on the real flight data of general aviation aircraft, the simulation results show that the anomaly data can be detected effectively by the proposed method of reconstructing ADS-B data with the seq2seq model, and its running time is reduced. Compared with the RNN, the accuracy of anomaly detection is increased by 2.7%. The performance of the proposed model is better than that of the traditional anomaly detection models.

Key words: general aviation aircraft; automatic dependent surveillance-broadcast (ADS-B); anomaly data detection; deep learning; difference of Gaussian (DoG); long short-term memory (LSTM)

CLC number: TN279

Document code: A

Article ID: 1005-1120(2020)04-0550-12

0 Introduction

The automatic dependent surveillance-broadcast (ADS-B) data which reflect the flight state and operation state of the aircraft are a series of flight parameters. The state of the aircraft and the operational behavior of pilots were recorded by the parameters of the aircraft^[1]. For the next generation air traffic surveillance, ADS-B is becoming the primary method to obtain more accurate data with wide cov-

erage, which establishes the foundation for automatic and intelligent air traffic management system^[2]. Many countries in the Asian-Pacific region have started to test and evaluate whether ADS-B can provide radar-like services, and American and Australian have planned and deployed ADS-B stations in their countries^[3]. Track refers to the flight path of an aircraft. There are inevitable errors or interference in the data transmission process, resulting in inaccurate or incorrect information. For guarantee-

*Corresponding author, E-mail address: wanges_2016@126.com.

How to cite this article: WANG Ershen, SONG Yuanshang, XU Song, et al. ADS-B anomaly data detection model based on deep learning and difference of Gaussian approach[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2020, 37(4): 550-561.

<http://dx.doi.org/10.16356/j.1005-1120.2020.04.006>

ing the airlines security, it is the primary requirement to ensure flight safety^[4]. In order to ensure the sustainable development and flight safety of civil aviation, it is necessary to predict and correct the anomaly data by means of methods.

Anomaly detection refers to finding data patterns in the data that contradict with the expected behaviors. In time series, there is often a context dependent relationship (time or space) among adjacent data points. Point anomalies due to the breakdown of dependencies can also be called context exceptions^[5]. According to the anomalies, a novel method of identifying instantaneous anomalies for retrospective safety analysis using energy-based metrics is proposed^[6]. The anomalies were detected by the constraints which were based on constraint from the operation database. But this model is sensitive to the operation data instead of position data. A novel partition-and-detect framework is proposed for the trajectory outlier detection method, which partitions a trajectory into a set of line segments. The outlying line segments were detected for trajectory outliers based on constraints from trajectory database^[7].

In recent years, a number of anomaly detection methods based on deep learning have been proposed in various fields. Deep learning methods can automatically learn task-related features through end-to-end training, and obtain high-level abstract representation of raw data through multilayer nonlinear transformation.

In order to detect the anomalies, the intelligence methods were used for detection, like neural networks^[8] and sequence to sequence (seq2seq)^[9]. A deep learning model based on seq2seq was proposed^[10]. In this model, the seq2seq model was used to reconstruct the ADS-B time series based on the data characteristics, and the reconstruction error can detect the anomalous ADS-B messages. Extending the feature space of time series enables the model to better capture the improver's effect of anomaly detection^[11]. The concept of time series anomaly was proposed by Keogh et al.^[12], which indicates that the anomaly of time series is independent of the size of the sliding window, and the extension of the characteristic of low-dimensional data to high-dimen-

sional data is not necessarily valid. So, in order to get the detail information of samples, the difference of Gaussian (DoG) approach was proposed^[13]. And deep learning is used to train the classifier to distinguish abnormal trajectories. But they all neglected to study the direction of the trajectory point. Other traditional types of detecting anomalies were based on the detecting abnormal compact position report (CPR) protocol, but they were hard to be implemented^[14-16].

The research showed that the method based on deep learning had completely surpassed the traditional method^[17-18]. Above studies reflect that anomaly detection has attracted growing attention from researchers.

The contribution of this paper is to propose an approach for anomalies based on deep learning and DoG approach. And this paper is organized as follows. A new method is presented for converting coordinate system with less computation and constraints for outliers in Section 1. In Section 2, DoG is used for enhancing the detail of ADS-B data. Seq2seq model and ADS-B anomaly data detection are proposed in Section 3. In the other sections of the paper, the problem is formulated, the method is described and the experiment validations are presented.

1 ADS-B Position Data Preprocessing

1.1 ADS-B position data acquisition

In order to get the tracks of the common aero vehicle, the ADS-B designed for common aero vehicle was used. First of all, the positions of the aircraft were positioned by BeiDou Navigation Satellite System (BDS) or Global Positioning System (GPS). Then, the tracks would be consisted by the points. Finally, the data of flights would be displayed on terminal and saved in database (Fig.1). In order to save the accurate position data with less anomalies, the data were processed by this model on the terminal.

Because of the terrain structure, meteorological conditions and various factors, there are anoma-

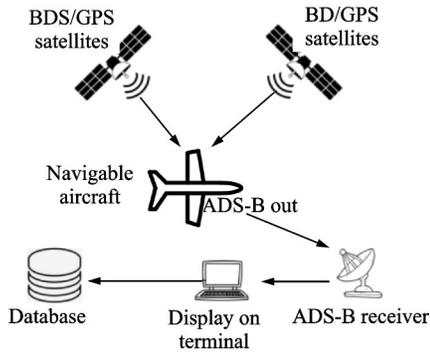


Fig.1 Transmission of ADS-B data

lies in ADS-B message. Tracks selected for processing should not preferably include tracks with the following exceptions.

(1) Incomplete points on track. Because of terrain and other reasons, message might be lost during the transmission processes. The points or a section of track will be incomplete. Some incomplete tracks with a spot of missing points can be completed with the interpolation method, but other incomplete tracks with many missing points must be deleted, which contain massive loss of points. An incomplete track is a whole track with too many missing points. The removal of a whole track does not affect the continuity of other tracks.

(2) Repeated points on track. Because of the error of clock or multi-path, the track data were received twice or more at the same time. In addition, the same points of track might be received at the different time due to the lack of data inspection. So, the less the repeated points, the better the tracks. And the tracks with less repeated points were chosen.

(3) Interference points on track. Different points on incomplete tracks of different flights might be received at the same time due to the similarity. Therefore, it is necessary for us to separate the independent tracks.

1.2 Conversion of ADS-B data into new coordinate system

The received ADS-B position data were written in World Geodetic System-84 (WGS-84). Because of the large computation in WGS-84 and Cartesian Coordinate System (CCS), the position can be converted into a new coordinate system where

the origin was defined by springboard as follows

$$s=2 \times 6387.137 \times t$$

$$t=\arcsin \sqrt{\sin ^2(a / 2)+\cos \left(\operatorname{lat}_1\right) \times \cos \left(\operatorname{lat}_2\right) \times \sin ^2(b / 2)} \quad (1)$$

where s is the plane distance of A and B , A the aircraft position and B the take-off point position; a means the difference-value of the latitude between A and B ; b is the difference-value of the longitude between A and B ; lat_1 is the latitude of A and lat_2 the latitude of B . The radius of the earth is 6387.137 km. When the distance is calculated on the X -axis, $b=0$. The positive direction on X -axis is east. When the distance is calculated on the Y -axis, $a=0$. The positive direction on Y -axis is south. When $a < 0$, the value on X -axis is negative. When $b > 0$, the value on X -axis is negative in Northern Hemisphere. The value on Z -axis is difference-value of altitude between A and B . In this model, the value of B is minuend. And spatial distance of A and B is

$$S=\sqrt{s^2+\Delta h^2} \quad (2)$$

where Δh is the difference-value in height between A and B .

1.3 Constraints from datasheet of general aviation aircraft

The constraints were calculated from the datasheet of the general aviation aircraft named RX1E, shown in Table 1.

Table 1 Datasheet of RX1E

Parameter	RX1E
The maximum overload / g	+4/-2
Cruise velocity / $(\text{km} \cdot \text{h}^{-1})$	120
The maximum flight altitude / m	3 000
Maximum velocity / km	150
Duration / min	about 60

From Table 1, the constraints $C(x)$ were gotten as follows

$$C(x)=\begin{cases} \left|v_i-v_{i-1}\right| < 42 \text{ m/s} \\ \left|v_i-v_{i-1}\right| > 0 \text{ m/s} \\ \left|H_i-H_{i-1}\right| < 42 \text{ m} \\ \left|H_i-H_0\right| < 3000 \text{ m} \\ \left|t_i-t_0\right| < 60 \text{ min} \end{cases} \quad (3)$$

where $v_i (i = 0, 1, 2, \dots, n)$ are the velocities of the aircraft, $H_i (i = 0, 1, 2, \dots, n)$ the altitude of the aircraft, and $t_i (i = 0, 1, 2, \dots, n)$ the time when the data were received.

1.4 Constraints from flight data and kinematics theory

The acceleration produced by gravity was decomposed to three directions including X-axis, Y-axis, Z-axis, shown as

$$G = \begin{bmatrix} G_x \\ G_y \\ G_z \end{bmatrix} = \begin{bmatrix} -mg \sin\theta \\ mg \cos\theta \sin\phi \\ mg \cos\theta \cos\phi \end{bmatrix}_{3 \times 1} \quad (4)$$

where G is a matrix, θ represents the angle between the direction of the flight speed and the positive direction of the X-axis, and ϕ the angle between the direction of the flight speed and the positive direction of Y-axis. According to the Newton’s second law and the instantaneous, accelerations of gravity were calculated as

$$A = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} = \begin{bmatrix} -g \sin\theta \\ g \cos\theta \sin\phi \\ g \cos\theta \cos\phi \end{bmatrix}_{3 \times 1} \quad (5)$$

where A is a matrix, and the maximum acceleration (a_k) of the aircraft has been obtained according to the data sheet of the aircraft. Thus, the theoretical maximum values of accelerations including (a_{xx}, a_{yy}, a_{zz}) were calculated as

$$A' = \begin{bmatrix} a_{xx} \\ a_{yy} \\ a_{zz} \end{bmatrix} = \begin{bmatrix} |-g \sin\theta + a_k \sin\theta| \\ |g \cos\theta \sin\phi + a_k \cos\theta \sin\phi| \\ |g \cos\theta \cos\phi + a_k \cos\theta \cos\phi| \end{bmatrix}_{3 \times 1} \quad (6)$$

where A' is a matrix. According to the velocity of the aircraft at the moment of $i + 1$, the actual accelerations of a_{iX}, a_{iY}, a_{iZ} in three directions are less than the theoretical maximum, that is

$$A'' = \begin{bmatrix} a_{iX} \\ a_{iY} \\ a_{iZ} \end{bmatrix} = \begin{bmatrix} \frac{v_{i+1} \sin\theta - v_i \sin\theta}{\Delta t} \\ \frac{v_{i+1} \cos\theta \sin\phi - v_i \cos\theta \sin\phi}{\Delta t} \\ \frac{v_{i+1} \cos\theta \cos\phi - v_i \cos\theta \cos\phi}{\Delta t} \end{bmatrix}_{3 \times 1} \quad (7)$$

where A'' is a matrix, a_{iX}, a_{iY}, a_{iZ} are the accelerations at time of $i + 1 (i = 0, 1, 2, \dots, n)$, and Δt is the time difference between time of $i + 1$ and i . The

actual accelerations in the X-axis, Y-axis, and Z-axis were calculated according to the flight data. And the constraints were calculated at time of $i + 1$ as

$$\begin{cases} a_{iX} \leq \frac{v_{i+1} \sin\theta - v_i \sin\theta}{\Delta t} \\ a_{iY} \leq \frac{v_{i+1} \cos\theta \sin\phi - v_i \cos\theta \sin\phi}{\Delta t} \\ a_{iZ} \leq \frac{v_{i+1} \cos\theta \cos\phi - v_i \cos\theta \cos\phi}{\Delta t} \end{cases} \quad (8)$$

where $i = 0, 1, 2, \dots, n$. If a_{iX}, a_{iY}, a_{iZ} conflict with the constraints, the velocity is the abnormal data at time of $i + 1$.

After removing the outlier in the flight data by using the constraint condition of acceleration, there may also be a continuous sequence of anomaly in the flight data that cannot be detected. In order to exclude outliers from these data, the aircraft data sheet and kinematics were used for detecting anomaly by calculating the theoretical position of the aircraft

$$\begin{cases} X_{i+1} = X_i + a_{iX} \Delta t \\ Y_{i+1} = Y_i + a_{iY} \Delta t \\ Z_{i+1} = Z_i + a_{iZ} \Delta t \end{cases} \quad i = 0, 1, 2, \dots, n \quad (9)$$

where $X_{i+1}, Y_{i+1}, Z_{i+1}$ are the theoretical positions of the aircraft on the X-axis, Y-axis, and Z-axis, such as

$$\begin{cases} X_{i+1} - x_{i+1} \leq v_{Xi} \Delta t + 1/2 a_{iX} \Delta t^2 \\ Y_{i+1} - y_{i+1} \leq v_{Yi} \Delta t + 1/2 a_{iY} \Delta t^2 \\ Z_{i+1} - z_{i+1} \leq v_{Zi} \Delta t + 1/2 a_{iZ} \Delta t^2 \end{cases} \quad (10)$$

where $i = 1, 2, 3, \dots, n$, x_{i+1} is the actual position of the aircraft at the time of $i + 1$. If the above constraints are not met, the positions of the aircraft on the X-axis, Y-axis and Z-axis are abnormal data.

2 Detail Enhancing of ADS-B Data

2.1 Position series extraction from ADS-B data

There are two dimensions in ADS-B data: time series and position series. When ADS-B data are arranged in two dimensions on a plane, it is different from the data arranged in one dimension on a line. In order to provide more information and more variation, the section will be used for exacting the position series of the ADS-B position data.

The set $T = \{T_1, T_2, \dots, T_n\}$ was composed by consecutive time series which include the points of tracks from 0 to n , and n is the length of the time series. T_i is a vector where the length of sequence is n , shown as

$$T_i = \{t_1, t_2, t_3, \dots, t_n\} \quad (11)$$

There is only one feature for each dimension. The contents for each vector, including position, course angle, height, etc., were broadcasted by ADS-B, shown as

$$\begin{aligned} t_1 &= \{x_1, y_1, z_1, v_1, a_1\} \\ t_2 &= \{x_2, y_2, z_2, v_2, a_2\} \\ &\vdots \\ t_n &= \{x_n, y_n, z_n, v_n, a_n\} \end{aligned} \quad (12)$$

where x_1, y_1, z_1, v_1, a_1 are the positions on X -axis, Y -axis, Z -axis, velocity and angle of aircraft on time of t_1 .

In order to extract the position sequence from T_i , multiple cross sections are extracted from the position sequence, and sample observations are selected on these cross sections to form the sample data. And each section starts at the same time.

$$\begin{aligned} X &= \{x_1, x_2, x_3, \dots, x_n\} \\ Y &= \{y_1, y_2, y_3, \dots, y_n\} \\ Z &= \{z_1, z_2, z_3, \dots, z_n\} \\ V &= \{v_1, v_2, v_3, \dots, v_n\} \\ A &= \{a_1, a_2, a_3, \dots, a_n\} \end{aligned} \quad (13)$$

where X, Y, Z, V, A are the position sections. In order to make the algorithm invariant to data scaling and reduce the influence of data changes on series, the method of establishing data Gaussian space to form data pyramid structure will be used.

2.2 Detail enhancing of ADS-B data by DoG

Data pyramid is a series of data from the same original section which is arranged in pyramid shape and gradually reduces resolution. It is the basis of multi-scale calculation based on the gradient data of the original section, the up-sampled data and the down-sampled data adjacent to the gradient data, shown in Fig.2. The down-sampled data are 1/4 of the original data, and the up-sampled data are 4 times of the original data.

The data details play an irreplaceable role in recognition of anomaly data. Anomalies are usually

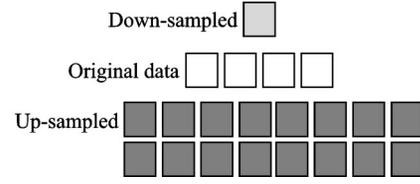


Fig.2 Block diagram of pyramid data

reflected in the details of the data, which are the result of the accumulation of local details. The local details of the section will be calculated by DoG. As an enhancement algorithm, DOG can be used to increase the visibility of edges and other details.

The level of detail for position sections will be reduced according to the function of Gaussian smoothing. Different scales of Gaussian smoothing have different smoothing effects on position sections. Each layer of data pyramid is smoothed data pyramid, forming three-layer Gaussian smoothing position sections.

$$L_i(n, m) = G_i * X_i(n, m) \quad (14)$$

where $G_i = \exp[-(x^2 + y^2)/2\sigma_i]$, $*$ denotes the convolution operation, $X_i(n, m)$ is the section of X , n the time, and m the value of dimension.

As shown in Fig. 3, L_0 is taken as input for Gaussian smoothing. The data are formed four layers with the three-layer smooth data obtained after the data processing of this layer, and the DoG data are obtained by subtracting from the adjacent filtered data.

$$L_i = L_{i-1} - L_i \quad i = 1, 2, 3 \quad (15)$$

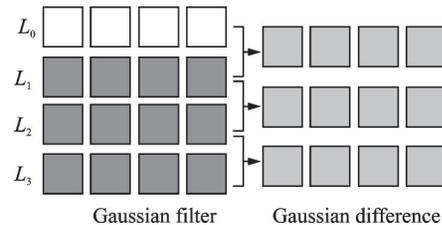


Fig.3 Block diagram of DoG

The pyramid data of each layer produce three-layer DoG data. The data obtained by Gaussian difference are a local detail description of the data at different scales. The detail description of the data can improve the expression ability of the position data section to the details. Pyramid data on each layer are superimposed with three-layer DoG data of this

layer section. To prevent overflow, the average value of data superposition is taken to form the detail superposition data, as shown in Fig.4. The value after superposition is L , shown as

$$L = \frac{1}{4} \left(L_0 + \sum_{i=1}^3 D_i \right) \quad (16)$$

After enhancing the details of the position section data, the details of the information are more abundant, and the extracted local features are more representative.

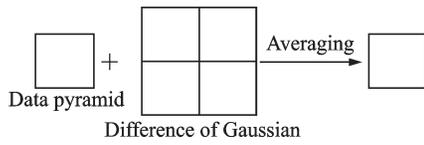


Fig.4 Block diagram of superposition

3 Seq2seq Model and ADS-B Anomaly Data Detection

3.1 Optimized seq2seq model by long short-term memory

The seq2seq model, which was used for conversion of sequence, is a model with the frame of Encoder-Decoder and often used in automatic abstracting, machine translation, deep learning and so forth. Encoder, decoder and state vector were included by seq2seq. State vector is a connection between encode and decoder. Recurrent neural network (RNN) is a unit used by encoder and decoder^[19].

The sequence of input will be learned by encoder and then be encoded into a fixed size state vector, which is then transmitted to the decoder part. Finally, the decoder learns the state vector to carry out the relevant output.

Because of the advantage of short-term memory for RNN, it is often used as the first choice neural network for the training of time sequence data. With the length of training data for RNN increasing, the gradients for training data will fade away. The gradients will disappear and the process of training will be broken while the length of training data is long enough. So, the limit time interval is essential

for training data when RNN is used. The length of the position data which is broadcasted by ADS-B in flights is so long that the feature can not be obtained by RNN. In order to resolve this problem, the improved long short-term memory (LSTM) units were used in seq2seq, which have excellent effects on many problems and now have been widely used, shown in Fig.5. The LSTM model can replace the RNN cells in the hidden layer with LSTM cells, so as to provide the hidden layer with long-term memory^[10]. The long-term dependency could be avoided effectively by the particular design of LSTM. It is easy for LSTM to acquire the feature of the time sequences with long interval.

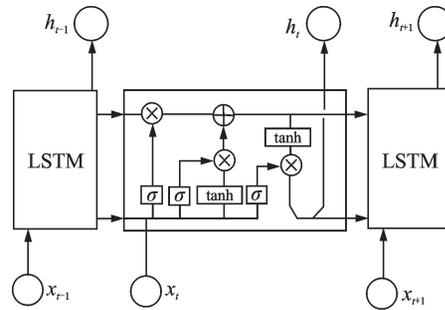


Fig.5 Chart for LSTM structure

In Fig.5, each line represents the transfer of the output of one node to the input of another node of vectors. Circles represent element-level operations of vectors, such as addition or multiplication. The boxes represent the layers of the neural network. A line merge represents a join of vectors, and a line bifurcation represents a copy of a vector. x_i is the input, and h_i the output.

The forgetting gate $f_t = \sigma(\mathbf{W}_s \cdot [h_{t-1}, x_t] + \mathbf{b})$, where \mathbf{W}_s is the weight matrix of the forgetting gate, σ the function of sigmoid, \mathbf{b} the bias term of forgetting gate, h_t the output at time of t , x_t the input at time of t , and i_t and \bar{c}_t are all apart of input of gate, shown as

$$\begin{aligned} \bar{c}_t &= \tanh(\mathbf{W}_c \cdot [h_{t-1}, x_t] + \mathbf{b}_c) \\ i_t &= \sigma(\mathbf{W}_i \cdot [h_{t-1}, x_t] + \mathbf{b}_i) \end{aligned} \quad (17)$$

where \tanh is the hyperbolic function, \mathbf{b}_c the weight matrix of \bar{c}_t , and \mathbf{b}_i the weight matrix of i_t . The unit state c_t at the current time is the product of the input

of the forgetting gate and the state at the previous c_{t-1} time adds the product of the two parts of the input gate (i_t and \bar{c}_t), shown as

$$c_t = f_t \odot c_{t-1} + i_t \odot \bar{c}_t \quad (18)$$

Finally, the sigmoid function is used to calculate information that needs to be output o_t , and then the output is obtained by multiplying the value which was calculated by hyperbolic function for current cell state.

$$\begin{aligned} o_t &= \sigma(\mathbf{W}_o \cdot [h_t - 1, x_t] + \mathbf{b}_o) \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad (19)$$

So, there are three gates: forgetting gate, input gate and output gate. In this model, LSTM was used. Core model consists of three parts for LSTM units of encoder and decoder, as well as the state vector (encoder state), respectively. In the training process, the training data are input to the encoder and encoded to a fixed state vector. Then, the vector is transferred to the decoder, and its sequence is reconstructed by the decoder.

3.2 Anomaly assessment index

The reconstruction error between the realistic time sequence and the predicted time sequence will be detected by using Cosine Similarity, that is

$$\cos(T_i, \tilde{T}_i) = \frac{\sum_{i=1}^n (t_i \times \tilde{t}_i)}{\sqrt{\sum_{i=1}^n T_i^2 \times \sum_{i=1}^n \tilde{T}_i^2}} \quad (20)$$

where n is the dimension, and \tilde{t}_i the reconstruction sequence. So the abnormal-probability p_a can be calculated as

$$p_a = \sum_{i=1}^{i=n} [1 - \cos(t_i, \tilde{t}_i)] \times 100\% \quad (21)$$

3.3 Improved anomaly data detection model

There are five main steps in this improved model (Fig.6).

Step 1 Data (latitude, longitude, height) of ADS-B in WGS-84 were converted to the data in a new coordinate, the origin of which is defined by the springboard.

Step 2 The constraint was created by the datasheet of the flight and the kinematical equation.

And the data which meet the constraints will be selected by the comparison between constraint and data. The constraints and the outliers will be detected.

Step 3 Extract position sections of the ADS-B data. And DoG will be used for detail enhancement of section.

Step 4 Parameters in seq2seq model were initialized. The original time series length is 12, the loss function used in the training is mean square error (MSE), the number of LMST unit is 256 with the batch-size of 128, and the ratio of dropout is 25%. The seq2seq model will be used for predicting the data. And the error between the realistic time sequence and the predicted time sequence will be detected by using this model.

Step 5 The abnormal data will be detected by Cosine similarity. After training, the threshold of the training sets is calculated. For actual flight data, the value of the anomalous score is defined with a threshold of 97.5%.

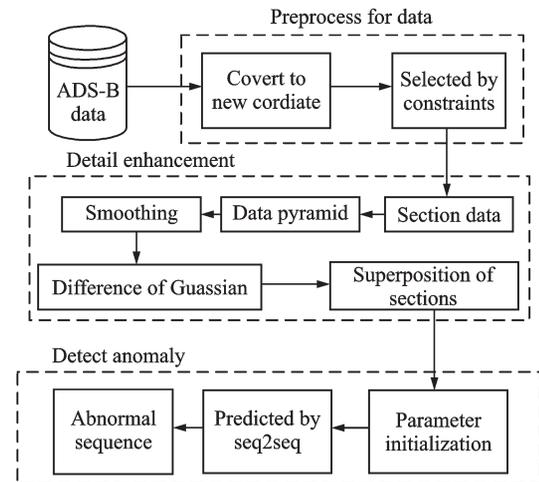


Fig.6 Main steps in this improved model

4 Simulation and Results Analysis

4.1 Strategies for experiment

The data come from the ADS-B which is broadcasted by common aero vehicle named RX1E. Some anomalies like outliers were produced while the data broadcasting. Outliers were detected by constraints.

Some anomalies were deliberately injected

such as network delay, route replacement, random noise and velocity drift, and almost all of them will be detected by seq2seq. Table 2 shows the strategies for the following injecting anomalies.

Table 2 Methods for injecting anomalies

Value	Strategy for producing error
Value on X-axis	Network delay
Value on Y-axis	Route replacement
Altitude	Random noise
Velocity	Velocity drift

(1) Network delay

In the selected ADS-B sequence, for every 10 consecutive messages, only the first message is retained and the last nine messages are deleted. This situation simulates the reality that the receiver can only receive valid message if the time interval is large enough.

(2) Route replacement

Given a certain route message, another correct route message is injected to replace the selected ADS-B sequence.

(3) Random noise

Part of the original ADS-B data are multiplied by a number between 0 and 2 randomly. Seven trajectory points of random noise will be injected per 180 track points.

(4) Velocity drift

The speed characteristic information contained in ADS-B message is gradually changed as a multiple of five sections. Specifically, in the selected ADS-B sequence, the speed characteristic contained in the first vector is increased by five sections, and the second by 10 sections.

In the course of the experiment, after the operations of normalization, the complete data of 10 flights from take-off to landing for seven consecutive days were selected as training samples, and the data of each flight were between 1 500 and 2 000. The data of flight for 20 min and 180 points in track were used for test set. And the test sets were taken from the flight records of Ruixiang RX1E aircraft in Faku General Aviation Airport of Shenyang at 10 o'clock on November 13, 2019. Its track is shown in Fig.7.

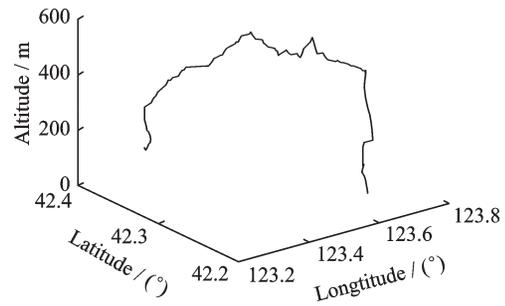


Fig.7 Flight track of Ruixiang RX1E aircraft

4.2 Result analysis for anomalies detection

After the operations for outliers, the model of seq2seq was used for anomalies in altitude. Fig.8 is the detection result used for the anomalies in the data of altitude. The anomaly type is the random noise. At the points 24, 59, 65, 78, 90, 101 and 114, random noise is injected. Due to the use of DoG approach, the details of the data are obvious. The threshold value of abnormal-probability for anomalies is 2.5%, and the time dependence of the data features is more easily to detect. At point 24, an anomaly just above the threshold was detected. Thus, this model is sensitive to anomalies. Another six anomalous points were detected by the seq2seq model.

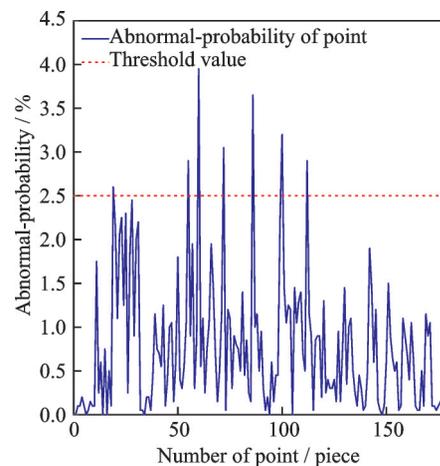


Fig.8 Anomaly processed by seq2seq for altitude

Similarly, Fig.9 is the detection result used for the anomalies in the data on X-axis. The network delay was injected on data of X-axis. At the points 11—16, 42, 58, 62, 94, 104, 11 and 128—145, anomalies of network delay are injected. The anomalies of the network delay are injected, resulting in

the increase of the abnormal-probability of anomalies, thus the anomalies of the abnormal network delay are detected. As shown in Fig.9, the threshold value of abnormal-probability for anomaly is 3%. All injected anomalies were detected by the seq2seq model.

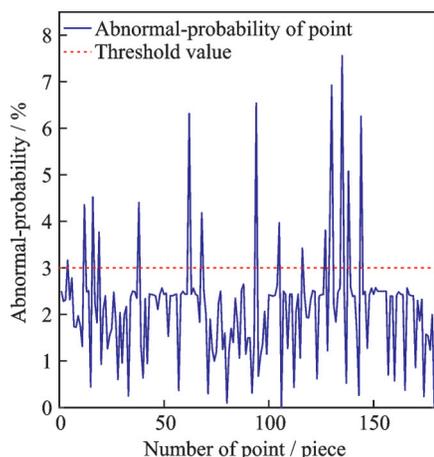


Fig.9 Anomaly processed by seq2seq for X-axis

After the operation for outliers, the model of seq2seq is used for abnormal of Y-axis. The detection result is shown in Fig.10. At the points 9, 11, 14, 31, 59, 119, 165 and 175, another route replaces the trajectory of the sample. As shown in Fig.10, the threshold value of abnormal scores is 3%, and the points above the threshold value are anomalies. The anomalies are detected by the seq2seq model.

Fig.11 is the detection result used for the outliers and anomalies in the data of track. As shown in

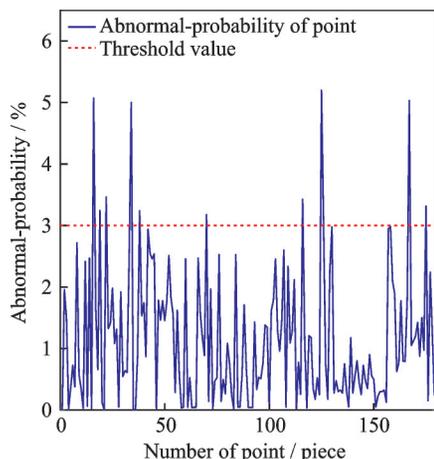


Fig.10 Anomaly processed by seq2seq for Y-axis

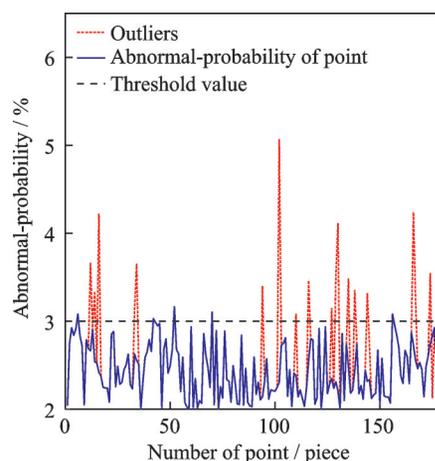


Fig.11 Anomaly processed by seq2seq for track

Fig.11, the dotted line is the outliers detected by the constraint conditions. The outliers are also identified by the seq2seq model. The injected type is velocity drift. According to the threshold value of abnormal-probability, at the points 51, 70 and 159, the anomalies were detected by seq2seq model instead of constraints. Thus, the seq2seq model is more sensitive than constraints.

4.3 Performance analysis of the model for running time and average precision

4.3.1 Analysis of running time

Running time is the time of converting ADS-B data to the corresponding coordinate system. Table 3 shows the running time in different coordinate systems. Since WGS-84 is the coordinate system with the center of the earth as the origin, the data value is very large, which makes the running time longest. The transition from the object coordinate system to the inertial coordinate system requires rotation, and the transition from the inertial coordinate system to the world coordinate system requires translation, thus resulting in a longer running time than the coordinate system in this model. The coordinate system in this model reduces the redundant bits in ADS-B data, and the calculation formula is simple, so the running time is decreased 0.14 s. Reducing the complexity of computation can reduce the running time and make the performance of real-time better, which is beneficial to the real-time processing.

Table 3 Running time in different coordinate systems

Coordinate	Network delay	Route replacement	Random noise	Velocity drift
Inertial coordinate	2.86	2.92	2.83	2.46
Coordinate in this model	2.82	2.89	2.79	2.19
WGS-84	2.97	2.95	2.86	2.48

4.3.2 Analysis of average precision

Precision indicates the number of correctly classified samples to the total number of samples for a given data set, shown as

$$\text{precision} = \frac{TP + TN}{TP + FN + FP + TN} \times 100\% \quad (22)$$

where TP is the true positive, FN the false negative, FP the false positive, and TN the true negative. In this paper, precision refers to the proportion of the correctly detected abnormal ADS-B samples to the actual detected abnormal samples. Table 4 shows the average precision in different neural networks.

Table 4 Average precision in different neural networks

Time series	RNN	LSTM
Network delay	80.5	83.6
Route replacement	82.7	87.3
Random noise	83.3	85.8
Velocity drift	85.6	86.2

As shown in Table 4, due to the gradient disappears severely, accuracy of RNN as the unit of seq2seq model is low, especially for the abnormal pattern of network delay when other conditions remain unchanged. In order to remember the long-term state, LSTM adds one input and one output to the RNN, which is a cell state, thus LSTM is more sensitive to ADS-B data which is a long time sequence. In Table 4, the average precision of LSTM is better than that of RNN, and for network delay, the average precision of LSTM is 3.1% higher than RNN. For the route replacement, random noise and velocity drift, the performance of LSTM is better than that of RNN due to disadvantages of gradient disappearing. Compared with RNN, the average precision has increased 2.7% for LSTM.

Aim to analyze the average precision in differ-

ent models, the Bayes, SVM and 4D-flight track prediction are used for comparison. After converting the data to the coordinate system of this model, the same training and test sets are selected to train and test the SVM model. These models use the known track characteristics to predict the track points at future moments. Table 5 shows the average precision in different models.

Table 5 Average precision in different models

Time series	Seq2seq	Bayes	SVM	4D
Network delay	83.6	74.6	73.5	76.1
Route replacement	87.3	76.3	76.1	78.0
Random noise	85.8	81.5	79.7	82.4
Velocity drift	86.2	78.4	80.8	82.3

(1) SVM

The kernel function of SVM is radial kernel function (RKF). The parameters of the model were optimized by using the grid search method. The kernel width parameter $\sigma = 0.02$ and the regularization parameter $\gamma = 100$ are determined.

(2) Bayes

The 180 track points constitute a set of observation data. Distribution parameters are calculated from training sets. The posterior probability can be calculated by Bayesian formula using prior probability and likelihood function.

(3) 4D-flight track prediction

The maximum likelihood law and Newton-Raphson iterative algorithm are used to identify the corrected airspeed in the model. Finally, the identification results and equiangular track prediction model are used to calculate the aircraft's track. The parameter identification is calculated by using maximum likelihood estimation. After 40 iterations of the calculation, the parameter identification is determined to be 221.6.

Bayes and SVM are prediction models based on classification and have strict requirements on training set samples. Therefore, when the actual flight samples are used as training sets, the accuracy is lower than 4D-flight track prediction. In the process, the model can effectively capture the characteristics of a track, so its accuracy is higher than that

of Bayes and SVM models. The seq2seq model in which neural network is composed of LSTM units has a good abnormal detection effect for ADS-B time series. And it takes into account the target change information in the adjacent time of the current point, so the detection performance is high and the accuracy can be improved by locking the target into a sequence. When the data are processed by network delay, route replacement, random noise, and velocity drift, there is a better performance in this model than in Bayes, SVM, and 4D-flight track prediction models.

5 Conclusions

The ADS-B anomaly data detection model based on deep learning and DoG approach is proposed. In order to reduce the complexity of this model, the kinematics theory is used for removing the outliers. The reconstruction error which is produced by seq2seq is used to detect the anomaly data. Through the simulation and verification, compared with SVM, Bayes, 4D-track models, the average precision is improved, and the running time is decreased.

References

- [1] MATTHEWS B, DAS S, BHADURI K, et al. Discovering anomalous aviation safety events using scalable data mining algorithms[J]. *Journal of Aerospace Information Systems*, 2014, 11(7): 482.
- [2] LI Tengyao, WANG Buhong, SHANG Fute, et al. Dynamic temporal ADS-B data attack detection based on sHDP-HMM[J]. *Computers & Security*, 2020. DOI: 10.1016/j.cose.2020.101789.
- [3] ZHANG Jun, LIU Wei, ZHU Yanbo. Study of ADS-B data evaluation[J]. *Chinese Journal of Aeronautics*, 2011, 24(4): 461-466.
- [4] DING Songbin, GU Qianqian, LIU Jiayu. Flight safety system evaluation and optimal linear prediction[J]. *Transactions of Nanjing University of Aeronautics and Astronautics*, 2019, 36(2): 205-213.
- [5] CHANDOLA V, BANERJEE A, KUMAR V. Anomaly detection: A survey[J]. *ACM Computing Surveys*, 2009, 41(3):15.1-15.58.
- [6] PURANIK T G, MAVRIS D N. Identifying instantaneous anomalies in general aviation operations[C]// *Proceedings of the 17th AIAA Aviation Technology, Integration and Operations Conference*. [S. l.]: AIAA, 2017: 3779-3794.
- [7] LEE J G, HAN J, LI X. Trajectory outlier detection: A partition-and-detect framework[C]//*Proceedings of ICDE 2008*. Washington: Computer Society, 2008: 140-149.
- [8] LI Ruiying, KANG Rui. Failure rate forecasting method based on neural networks[J]. *Acta Aeronautica et Astronautica Sinica*, 2008, 29 (2) : 357-363. (in Chinese)
- [9] GRAVES A, JURGEN S. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. *Neural Networks*, 2005, 18 (5/6): 602-610.
- [10] DING Jianli, ZOU Yunkai, WANG Jing, et al. ADS-B anomaly data detection model based on deep learning[J]. *Acta Aeronautica et Astronautica Sinica*, 2019, 40(12):167-177.(in Chinese)
- [11] LI Kexin, LI Jing, LIU Shuji, et al. GA-iForest: An efficient isolated forest framework based on genetic algorithm for numerical data outlier detection[J]. *Transactions of Nanjing University of Aeronautics and Astronautics*, 2019, 36(6): 1026-1038.
- [12] KEOGH E, LIN J, LEE S H, et al. Finding the most unusual time series subsequence: Algorithms and applications[J]. *Knowledge & Information Systems*, 2007, 11(1): 1-27.
- [13] ZHOU S, LIANG Y, WAN J, et al. Facial expression recognition based on multi-scale CNNs[C]//*Proceedings of Chinese Conference on Biometric Recognition*. [S.l.]: Springer International Publishing, 2016: 503-510.
- [14] YANG Youqiong. Analysis of the strategy of wisdom library service implementation[J]. *Guizhou Library Publication*, 2018 (2) : 67-70. (in Chinese)
- [15] YANG H, HUANG R, WANG X, et al. EBAA: An efficient broadcast authentication scheme for ADS-B communication based on IBS-MR[J]. *Chinese Journal of Aeronautics*, 2014, 27(3): 688-696.
- [16] WESSON K D, HUMPHREYS T E, EVANS B L. Can cryptography secure next generation air traffic surveillance?[J]. *IEEE Security & Privacy*, 2018. DOI: http://radionavlab.ae.utexas.edu/images/stories/files/papers/adsb_for_submission.pdf.
- [17] EDAN H, ASAF S. Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages[J]. *Computers & Security*, 2018(78): 155-173.

- [18] HUAN-KAI P, RADU M, GOMEZ S. Multiscale compositionality: Identifying the compositional structures of social dynamics using deep learning[J]. Plos One, 2015, 10(4): e0118309.
- [19] BAEK J, HABLEEL E, BYON Y J, et al. How to protect ADS-B: Confidentiality framework and efficient realization based on staged identity-based encryption[J]. IEEE Transactions on Intelligent Transportation Systems, 2016. DOI: <http://10.1109/TITS.2016.2586301>.

Acknowledgements The work was supported by the National Key R & D Program of China (No. 2018AAA010080-4), the Talent Project of Revitalization Liaoning (No. XLYC1907022), the Key R & D Projects of Liaoning Province (No. 2020JH2/10100045), the Capacity Building of Civil Aviation Safety (No. TMSA1614), the Natural Science Foundation of Liaoning Province (No. 2019-MS-251), the Scientific Research Project of Liaoning Provincial Department of Education (Nos. L201705, L201716), the High-Level Innovation Talent Project of Shenyang (No. RC190030), and the Second Young and Middle-Aged Tal-

ents Support Program of Shenyang Aerospace University.

Author Prof. WANG Ershen received the Ph.D. degree from Dalian Maritime University in 2009. He is a teacher of College of Electronic and Information Engineering at the Shenyang Aerospace University. His research currently focuses on Global Navigation Satellite System (GNSS) positioning algorithms, BDS/GNSS receiver signal process techniques.

Author contributions Prof. WANG Ershen designed the study, revised and modified the manuscript. Mr. SONG Yuanshang simulated the model. Prof. GUO Jing contributed to data for the analysis of the model. Mr. XU Song contributed to the background of the study. Dr. HONG Chen contributed to the discussion of the study. Prof. QU Pingping contributed to the analysis of the model. Dr. PANG Tao conducted the analysis of the results. Mr. ZHANG Jiantong contributed to the discussion of the study. All authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing interests.

(Production Editor: ZHANG Huangqun)

基于深度学习与高斯差分法的ADS-B异常数据检测模型

王尔申¹, 宋远上¹, 徐嵩¹, 郭婧², 宏晨³, 曲萍萍¹,
庞涛¹, 张建通⁴

(1. 沈阳航空航天大学电子信息工程学院, 沈阳 110136, 中国; 2. 中国民航科学技术研究院, 北京 100028, 中国;
3. 北京联合大学机器人学院, 北京 100101, 中国; 4. 中国交通通信信息中心, 北京 100011, 中国)

摘要:受地形结构、气象条件等多种因素的影响,用于低空通航飞行器定位的广播式自动相关监视(Automatic dependent surveillance-broadcast, ADS-B)设备获取的位置信息存在异常数据。为检测异常数据,提出一种基于深度学习与高斯差分法的ADS-B异常数据检测模型。首先,依据ADS-B位置数据的特点,将ADS-B位置数据转换到以起飞点为原点的坐标系中,利用运动学原理去除ADS-B位置数据中的离群点。然后,利用高斯差分法(Difference of Gaussian, DoG)获取位置数据的细节信息。最后,利用长短期记忆单元(Long short-term memory, LSTM)神经网络优化在ADS-B位置数据中梯度减小严重的循环神经网络(Recurrent neural network, RNN)。通过LSTM神经网络构成的seq2seq(Sequence to sequence)模型对位置数据进行重构,利用重构误差检测异常数据。通过实际数据对模型进行验证和对比分析表明:利用seq2seq模型对ADS-B位置数据重构的方法能有效地检测异常数据,运行时间得到减少,而且相较于RNN神经网络,检测的平均准确率提高了近2.7%,相较于传统的异常检测模型具有更高的准确率。

关键词:通用航空飞行器;广播式自动相关监视;异常数据检测;深度学习;高斯差分;长短期记忆单元