ADS-B Anomaly Detection Algorithm Based on LSTM-ED and SVDD

Jiao Yi(✉), Lin Lin, Li Nisi, and Wang Jintao

CCAC Academy of Flight Technology and Safety, Civil Aviation Flight University of China,
Guanghan 618300, Sichuan Province, China
jy18980203626@gmail.com

Abstract. With the rapid development of radio broadcasting technology and information technology, the possibility of ADS-B surveillance system suffering from radio frequency interference and network attack is increasing. Firstly, the types of attacks that ADS-B system is vulnerable to are analyzed, and the abnormal data of the corresponding state are simulated. Then, in order to detect ADS-B data attacks accurately, an anomaly detection algorithm based on Long Short-Term Memory Encoder–Decoder (LSTM-ED) and hypersphere classifier is proposed. Using the normal historical data to train the LSTM-ED model so that the reconstructed values can be obtained. Then, the difference values between the actual values and the reconstructed values are put into SVDD (support vector data description) for training, and a hypersphere classifier that can detect abnormal data is obtained. It is judged whether there is an abnormality according to the set threshold. Experiments show that the LSTM-ED-SVDD model can detect abnormal ADS-B data generated by various attacks and has high accuracy.

Keywords: ADS-B · Anomaly detection · LSTM-ED · SVDD classifier

1 Introduction

As one of the important components of air traffic monitoring, ADS-B system has been widely used around the world with the advantages of high monitoring accuracy, low deployment cost, wide coverage, and data sharing [1]. In order to reduce the system cost, the ADS-B system chooses to broadcast data information in plaintext, which leads to security communication problems such as information leakage and information tampering [2]. With the development of radio technology and the wide application of ADS-B system, the security drawbacks left in the early stage of system development are gradually exposed and amplified, becoming the biggest obstacle to the full application of ADS-B.

As of January 2020, the Federal Aviation Administration (FAA) has abandoned about 175 secondary surveillance radars. According to regulations, aircraft in specific types of airspace need to be forcibly equipped with ADS-B avionics to provide accurate information for air traffic management to replace primary and secondary radar systems. At the same time, China is also building a complete civil aviation ADS-B operation

Chinese Society of Aeronautics and Astronautics: CASTYSF 2022, LNEE 972, pp. 245–257, 2023.
https://doi.org/10.1007/978-981-19-7652-0_23

monitoring system and information service system. At present, most countries around the world have been equipped with ADS-B equipment. The comprehensive application of ADS-B system not only improves the performance of aviation surveillance, but also increases the risk of aviation safety.

McCallie et al. [3] analyzed ADS-B security vulnerabilities and their overall impact on transport aviation, classified ADS-B attack types according to the difficulty of implementation, and put forward suggestions for improving ADS-B security. Andrei [4] sent the modification information by a universal software radio peripheral (USRP) unit and completed the message reception by a commercial receiver. Magazu and Schäfer et al. [5, 6] demonstrated the generation and injection of fake ADS-B messages, and show information injection attacks in specific scenarios. The authors [7, 8] analyzed the flow of European air traffic on the 1090 MHz data link, and discussed the challenges and safety problems of ADS-B on the 1090es data link. The results show that in dense airspace, the information loss rate of more than 50% is also a serious security problem. Mohsen et al. [2] used Piccolo autopilot to run simulations on hardware to study the impact of ADS-B message injection attacks. The study showed that such attacks can significantly distract pilots and ground controllers, forcing them to execute emergency handling plans away from ghosts Aircraft, in the case of heavy traffic load, will endanger airspace safety.

All of the above studies have shown that ADS-B data can be easily corrupted by existing cheap hardware and software. An attacker can receive and modify ADS-B messages, delete legitimate ADS-B messages, and inject misleading ADS-B messages. Security flaws in ADS-B systems raise concerns about the direction and deployment of next-generation aviation. New security protocols take a long time and are expensive to roll out. Therefore, the research on ADS-B data anomaly detection algorithm has certain practical significance. Ensure the accuracy and validity of ADS-B data from the data analysis level, so as to assist controllers and flight personnel to make correct decisions and maintain the safety and order of air traffic.

2 Related Works

2.1 ADS-B Security Vulnerabilities

The rapid development of software radio provides potential attackers with a low-cost attack method. The attacked objects in the ADS-B system are mainly ground stations and aircraft. Attacks can be divided into passive and active. Passive attacks are mainly message eavesdropping; active attacks include tampering, injection, deletion and DOS attacks [9].

- Message eavesdropping: The message is not encrypted and broadcast on a fixed frequency. An attacker can regularly monitor the ADS-B message and obtain the location information of a flight without interrupting the system.
- Message tampering: The ADS-B system does not have an inquiry mechanism similar to the secondary radar, and neither the sender nor the receiver can authenticate. Attackers use stronger jamming signals to obscure the original signal, or modify aircraft position information and warning codes by bit flipping. Gradually steer the plane away from its original course without the controller and pilot noticing it.

- Message deletion: The message information is a radio frequency signal, and an attacker can attenuate and damage the message information by superimposing the reverse signal and the error signal.
- Message injection: Attackers use software to simulate and generate fake aircraft information in the same format, and use transceiver equipment to inject fake information into the air traffic control surveillance system and collision avoidance system, eventually creating ghost planes and disrupting the normal operation of air traffic control.
- DOS attacks: A denial of service attack refers to interfering with data link communications in an ADS-B system, preventing participants from receiving or sending messages. Attackers target ground stations or aircraft. For busy airspace or airports, ground radio devices are used to transmit messages to achieve message blocking, which will have a huge impact on the normal operation of air traffic control.

2.2 Anomalous Data Construction

The Opensky network is primarily an air traffic control data provider that collects, processes and stores flight data [10]. The data used for the experiments in this paper comes from the Openky dataset. The ADS-B data is basically normal data. The data that has suffered from abnormal attacks is rare and cannot be obtained. In order to verify the effectiveness of the anomaly detection model, the anomaly data is simulated and generated according to the ADS-B attack types. The test data set in this paper selects the data set of a certain flight for construction processing. For example, select the middle 100 pieces of data from the 1600th to 1900th pieces of a certain flight to simulate the following four types of anomalies, as illustrated in Fig. 1.

- Message injection attack: The route replacement method is used to simulate, and the flight data of another flight with a similar flight status is replaced with the original part of the track.
- Message tampering (location): Modify the original track data from the characteristics of longitude and latitude, and add Gaussian noise with a mean value of 0 and a variance of 0.1 to the flight longitude and latitude within a certain range, so that the attacked trajectory deviates from the actual trajectory.
- Message tampering (altitude): Gradually increase or decrease the flight height in a certain proportion within the flight range, so that the attacked flight trajectory deviates from the original flight trajectory in height.
- DOS attack: simulate message blocking and interfere with normal communication, choose to delete flight data within a certain period of time, and replace it with a value of 0, that is, the receiver cannot receive any messages.

2.3 Long Short-Term Memory Encoder–Decoder

Long Short Time Neural Network (LSTM) is a variant of RNN recurrent neural network [11]. It adds a memory component (forgetting gate, input gate, output gate) on the basis of RNN network to help the network transfer the information learned at T time to $T + n$

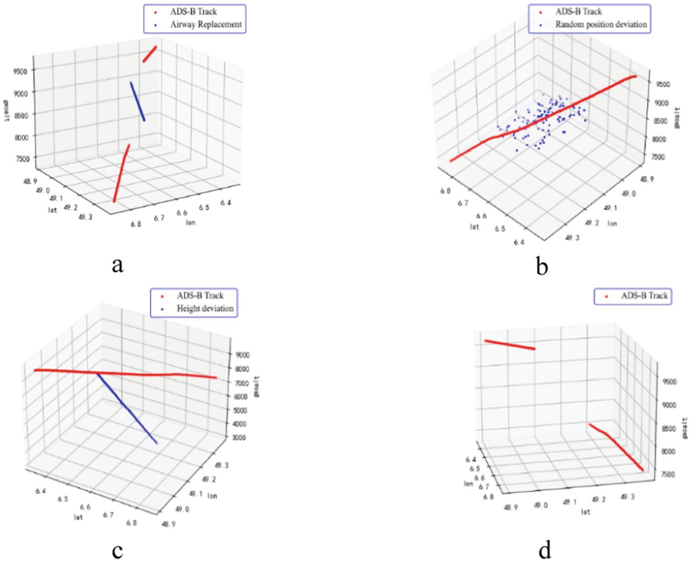


Fig. 1. Types of ADS-B data exceptions. **a** route replacement, **b** random position offset, **c** height offset, **d** DOS attack

time. In this way, the state can be selectively updated, and the important part of the state can be transmitted to the future stage, so that LSTM has the ability of long-term memory information, It works as follows:

$$f_t = \sigma_g\left(W_f x_t + U_f h_t + b_f\right) \tag{1}$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \tag{2}$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \tag{3}$$

$$c_t = f_t * c_{t-1} + i_t * \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \tag{4}$$

$$h_t = o_t * \sigma_c(c_t) \tag{5}$$

x_t represents the LSTM cell input vector; f_t represents the activation vector of the forget gate; i_t represents the activation vector of the input gate; o_t represents the activation vector of the output gate; h_t represents the hidden state vector, which is also the output vector of the LSTM cell; c_t represents the cell state vector; W and U represents the weight matrix and Bias vector parameter, that is, what needs to be learned in the training process; σ_g represents the sigmoid function; σ_c represents the tanh function, the LSTM structure is shown in Fig. 2.

When using LSTM models for long sequence predictions, prediction errors accumulate over time. Order to solve this problem, this paper uses a model called Long

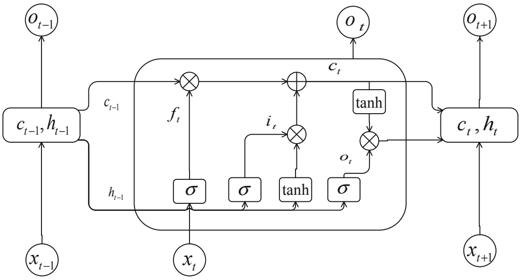


Fig. 2. LSTM network

Short-Term Memory Encoder-Decoder [12], whose structure is shown in Fig. 3. The encoder and decoder are two independent parts. Each part consists of one or more LSTM networks. In the model, the sequence is input to the encoder and outputs an internal state, which is used as the output of the decoder to predict the state at the current moment by the input of the previous.

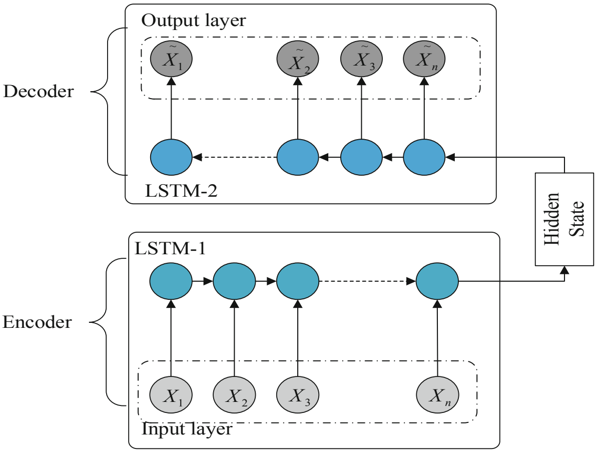


Fig. 3. Long short-term memory encoder–decoder

2.4 SVDD Classifier

The key to anomaly detection of ADS-B data lies in how to distinguish abnormal values from normal values. The most common method for classification is to judge according to the set threshold and the calculation of similarity, but the artificial threshold setting has certain limitations. The classifier obtains the classification boundary through the characteristics of the data itself, which has better adaptability as a threshold. The abnormal data detection only includes normal and abnormal data, which belongs to the binary classification problem.

SVDD (Support Vector Data Description) is an unsupervised machine learning method that can solve binary classification problem [13]. SVDD mainly maps the normal feature data $x \in R^{n \times d}$ (n represents the number of samples, d represents the feature dimension) into a high-dimensional feature space, and find a hypersphere that can contain the target samples. The hypersphere is minimized while including the target samples as much as possible, and the non-target samples are excluded from the hypersphere, as to achieve the purpose of distinguishing between normal and abnormal classes. As shown in Fig. 4, the hypersphere is used as the decision boundary, and the points on the boundary of the sphere are used as support vectors. When the distance between the test sample and the center of the sphere is less than the radius of the sphere, it is judged as a normal sample, and when the distance is greater than the radius of the sphere, it is judged as an abnormal sample.

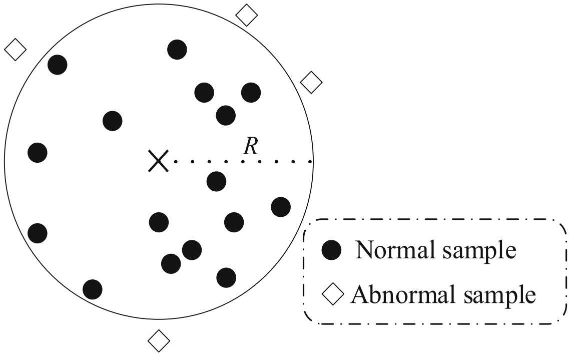


Fig. 4. SVDD hypersphere

Order to construct the smallest volume hypersphere, the optimization problem that SVDD needs to solve is shown in Eqs. 6 and 7:

$$\min_{a,R,\xi} R^2 + C \sum_{i=1}^{n} \xi_i \tag{6}$$

$$s.t. \|\phi(x_i) - a\|^2 \leq R^2 + \xi_i, \xi_i \geq 0, \forall i = 1, 2, \ldots, n \tag{7}$$

R represents the radius of the hypersphere, a represents the center of the hypersphere, ξ Represents the relaxation factor, and C represents the penalty coefficient for weighing the hypersphere volume.

Combined with the Lagrange multiplier method, the original optimization problem is transformed into a Lagrange dual problem:

$$\min_{\alpha_i} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j K\left(x_i, x_j\right) - \sum_{i=1}^{n} \alpha_i K(x_i, x_i) \tag{8}$$

$$s.t. 0 \leq \alpha_i \leq C, \sum_{i=1}^{n} \alpha_i = 1 \tag{9}$$

α_i represents the Lagrangian coefficient corresponding to the sample, $K\left(x_i, x_j\right) = \langle \phi(x_i), \phi(x_j) \rangle$ represents the kernel function used to map samples into a high-dimensional space. Solve to obtain the Lagrangian coefficient α_i corresponding to all samples, and the samples that satisfy the conditions in Eq. 9 are the support vector x_v.

The center a and radius R of the hypersphere are derived from the sample set S composed of support vectors:

$$a = \sum_{i=1}^{n} \alpha_i \phi(x_i) \tag{10}$$

$$R = \sqrt{K(x_v, x_v) - 2\sum_{i=1}^{n} \alpha_i K(x_v, x_i) + \sum_{i=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_j K\left(x_i, x_j\right)} \tag{11}$$

3 ADS-B Abnormal Data Detection Algorithm Based on LSTM-ED-SVDD

3.1 Model Building

The ADS-B anomaly detection model based on LSTM-ED-SVDD is mainly divided into two parts: time series data reconstruction and reconstruction error classification. Using the sliding window mechanism, the data of the next moment is constructed from the data of the previous moments. The difference between the reconstructed data and the original data is used as the sample of the classifier, and finally it is determined whether the reconstructed data is an outlier according to the trained classifier, The general framework of LSTM-ED-SVDD model is shown in Fig. 5.

3.2 Model Training

The data used for the training of the anomaly detection model is the flight data of 10 flights belonging to the same route. The flight time of each flight is about one hour, the amount of data is 3000 ~ 4000, and the division ratio of training set and test set is 7:3. The training set is used to train the anomaly detection model, and the divided test set data is also normal flight data, so it cannot be used for testing directly, and anomaly data simulation is required.

According to the analysis of ADS-B message attack type and flight safety impact factor, five basic features are screened out. In this paper, ADS-B feature data at each time is used as a feature vector includes longitude, latitude, ground velocity, vertical velocity and heading:

$$M_{t_i} = [lon, lat, geo\_alt, vel, ver\_rate, true\_track]^T$$

All feature vectors form a feature sequence:

$$M = \{M_{t_1}, M_{t_2}, \cdots M_{t_i}\}$$

In this model, min-max normalization is used to limit each feature data to [0, 1]. The normalized eigenvector $M_{t_i}^{'}$ is shown in Eq. 12.

$$M_{t_i}^{'} = \frac{M_{t_i} - \min\left(M_{t_i}\right)}{\max\left(M_{t_i}\right) - \min\left(M_{t_i}\right)} \tag{12}$$

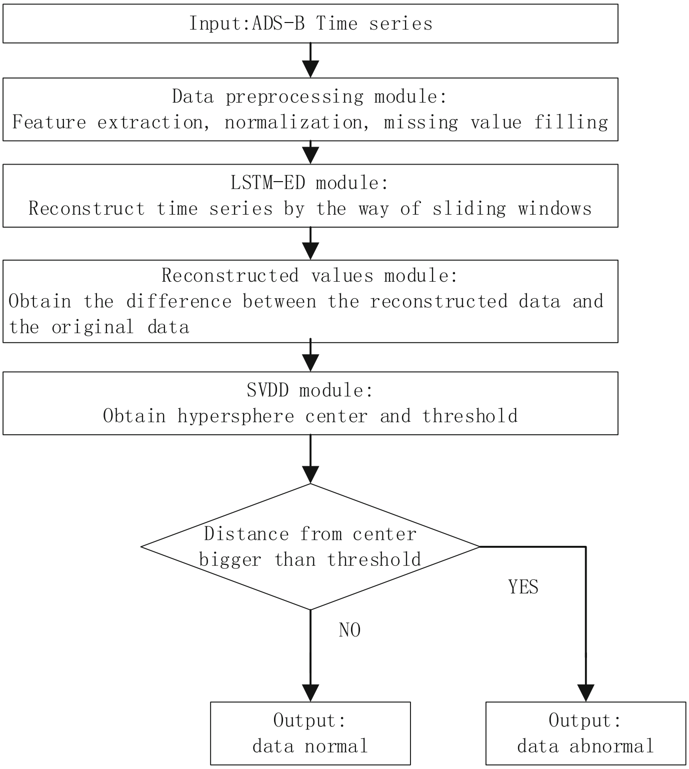

Input:ADS-B Time series
Data preprocessing module:
Feature extraction, normalization, missing value filling
LSTM-ED module:
Reconstruct time series by the way of sliding windows
Reconstructed values module:
Obtain the difference between the reconstructed data and the original data
SVDD module:
Obtain hypersphere center and threshold
Distance from center bigger than threshold
YES
NO
Output:
data normal
Output:
data abnormal

Fig. 5. Model framework

After the training data is preprocessed, the encoder in the LSTM-ED model learns the change trend in the normal historical data. Sliding window mechanism in the time series learning is the key to reconstructing and amplifying the abnormal data. This paper adopts the data sliding window mechanism. The window uses two sliding endpoints to identify the time series of specified length, as shown in Fig. 6. When new data is added in the window, the old data is deleted, that is, the data within a range is used to predict the next long data. According to the experimental calculation of the accuracy of reconstructed data with different window sizes, the window length is gradually increased from the window length of 2, and finally the length of the sliding window is determined to be 8, and the sliding step is set to 1.

The number of training epoches of the LSTM-ED network is 50, and the training batch size is 20. The change of the loss function value of the model during the training process is shown in Fig. 7. With the continuous increase of the number of training iterations, the value of the loss function gradually tends to be stable. The exact value of the training set can reach 0.928.

The training samples D used for SVDD classification are the difference sequence between the data reconstructed by LTSM-ED and the original data.

$$D = \{D_{t_1}, D_{t_2}, \cdots, D_{t_i}\},$$

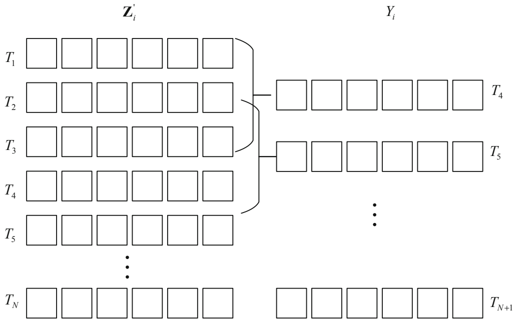
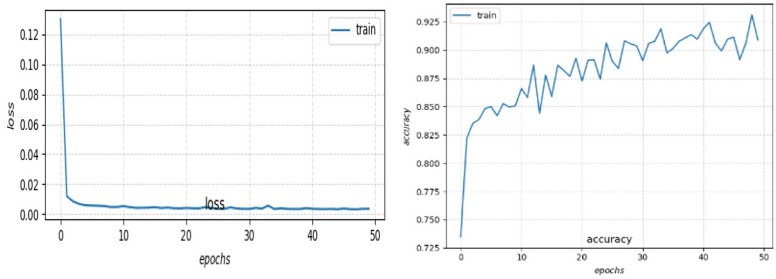


Fig. 6. Sliding window



a b

Fig. 7. Model training. **a** The change of loss value, **b** The change of accuracy

The sample data constructed by the difference sequence does not contain labels, and the samples need to be labeled before entering the classifier training. The samples used for training are all normal data, so the label needs to be added as 1. The classifier uses the Gaussian kernel function for model training. The model training mainly includes two parameters: The penalty coefficient C represents the tolerance to error. If its value is too large or too small, it will lead to the deterioration of the generalization ability of the model; The Gaussian kernel parameter gamma determines the distribution of the original data after mapping to the new feature space. The larger the gamma value, the fewer the support vectors, and the smaller the value, the more the support vectors. The number of support vectors will affect the speed of model training and prediction. According to the parameters of the training model, the grid search method is used to determine, that is, the punishment coefficient and Gaussian kernel parameters in a given range are arranged and combined, and all possible combinations generate a "grid".The accuracy of the model training corresponding to parameters shown as Table 1.

The Gaussian kernel function maps the difference samples to a high-dimensional space to find a hypersphere. The grid search method obtains the penalty coefficient of the optimal model C = 0.15, the Gaussian kernel parameter g = 10, and the radius of the hypersphere obtained by training is 0.221.

Table 1. Accuracy of model training with different parameters under Gaussian Radial Basis Kernel function

Penalty coefficient C	Gaussian kernel parameter g	Model training accuracy
0.15	1	78.4%(±0.525)
0.15	1	80.4%(±0.477)
0.15	0.5	95.4%(±0.142)
1	0.5	69%(±0.570)
1	0.1	88.6%(±0.513)
4	1	77.1%(±0.564)
4	0.2	87%(±0.449)
10	0.15	99.7%(±0.006)
10	1	91.7%(±0.214)
10	10	55.2%(±0.569)

3.3 Model Testing

In practice, it is difficult to obtain abnormal data after ADS-B is attacked. The test set used for model checking is the abnormal data constructed by simulation. Among the 300 pieces of flight data, the 100th to 200th pieces are abnormal data, and the rest are normal data. As a binary classification problem, the actual ADS-B data will be divided into normal classes or abnormal classes, and the data of the test set after passing through the anomaly detection model will be divided into four cases: The actual normal ADS-B data is accurately classified and predicted to be normal TP, the actual normal ADS-B data is misclassified and predicted to be abnormal FP, the actual abnormal ADS-B data is misclassified and predicted to be normal FN, and the actual abnormal ADS-B data is correct Classification predicted as abnormal TN.

- True positive rate TPR: It indicates the proportion of normal data that the classifier correctly attributes to all actual normal data.
- False positive rate FPR: indicates the proportion of abnormal data wrongly classified as normal class by the classifier to all abnormal data.
- Accuracy: Indicates the probability that the prediction is correct in the entire test data set.
- Precision: Indicates the ratio of actual normal data to predicted normal data.
- Recall: Represents the probability that the actual normal data is predicted correctly.

The anomaly detection judgment in this model is to calculate whether the distance from the feature data at each moment to the center of the hypersphere is greater than the trained sphere radius 0.221. The detection result of each type of attack consists of two parts, where the left graph represents the distance from each sample in the test set to the center of the sphere, where the red line represents the classification threshold (the radius of the hypersphere for training), and the points above the threshold line are abnormal samples, the normal samples at points below the threshold line. On the right is a receiver operating characteristic (ROC) curve formed by true positive rate TPR and false positive rate FPR. The area AUC under the curve represents the data probability of true normal class or true abnormal class predicted in the anomaly detection model. The closer the AUC of the model is to 1, the higher the accuracy of the model prediction.

1. Random position offset: The attack detection result is shown in Fig. 8. Among the 100 simulated abnormal data, 89 were correctly predicted as abnormal, and 195 of the 200 normal data were predicted as normal. The accuracy of the model test was 94.67%, the recall rate of the model was 97.5%, and the accuracy rate of the model was 94.67%.
2. High attack: The attack detection results are shown in Fig. 9. 96 of the 100 abnormal data in the middle are correctly predicted as abnormal, and 196 of the 200 normal data are predicted as normal. The accuracy of the model test is 97.33%, the recall rate of the model is 98%, and the accuracy rate of the model is 98%.
3. Route replacement: The attack detection results are shown in Fig. 10. 94 of the 100 abnormal data in the middle are correctly predicted to be abnormal, and 198 of the 200 normal data are predicted to be normal. The accuracy of the model test is 97.33%, and the recall rate of the model is Recall. is 99%, and the model accuracy is 97.0588%.
4. DOS attack: The attack detection results are shown in Fig. 11. 92 of the 100 abnormal data in the middle are correctly predicted as abnormal, and 198 of the 200 normal data are predicted as normal. The accuracy of the model test is 96.667%, the recall rate of the model is 99%, and the accuracy rate of the model is 96.165%.

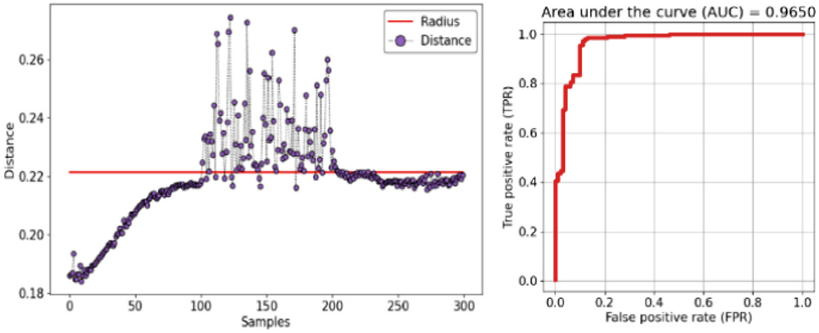


Fig. 8. Random offset detection result

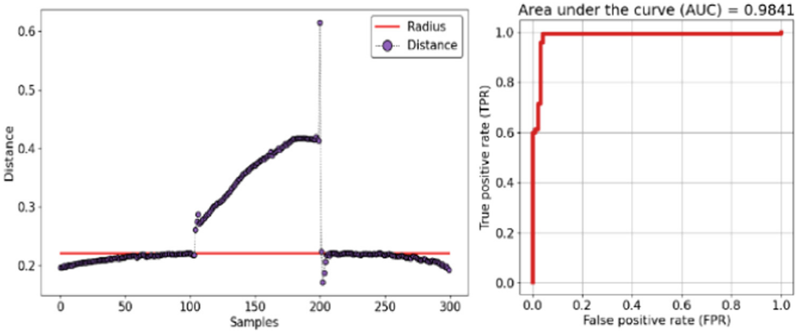


Fig. 9. Height attack detection result

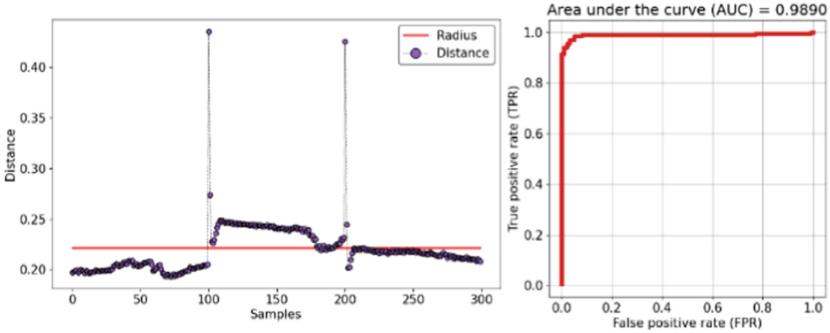


Fig. 10. Route replacement test result

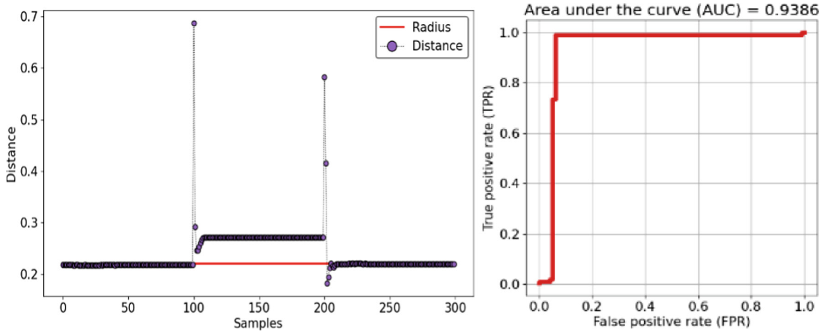


Fig. 11. Dos attack test result

4 Conclusions

This paper proposes an ADS-B anomaly data detection algorithm that combines the LSTM-ED model and the SVDD classifier. First, LSTM-ED is used to reconstruct the ADS-B data. Then, SVDD is used to solve the adaptive problem of threshold. The performance of the LSTM-ED-SVDD model is verified by simulating four common attacks.

In the future, we will continue our work in the following three aspects. First, we plan to use the radio software to simulate the abnormal data after the attack, which is more conducive to the verification of the accuracy of the model. Second, the model in this paper is anomaly detection for fixed routes, which lacks certain generality in practical engineering applications. Third, it is still in the analysis and research stage of offline data. With the development of deep learning and explainable artificial technology, the problems of optimal detection model and real-time detection remain to be solved.

Funding. Innovation and Entrepreneurship Training Program for Chinese College Students, Grant Number: S202110624266, Science and technology planning project of Sichuan province, Grant Number: 2021YFG0180.

References

1. Strohmeier, M., Lenders, V., Martinovic, I.: Security of ADS−B: State of the Art and Beyond. DCS (2016)
2. Mccallie, D., Butts, J., Mills. R.: Security analysis of the ADS-B implementation in the next generation air transportation system. Int. J. Critical Infrastruct. Protect. **4**(2), 78–87 (2011)
3. McCallie, D., Butts, J., Mills, R.: Security analysis of the ADS-B implementation in the next generation air transportation system. Int. J. Crit. Infrastruct. Prot. **4**(2), 78–87 (2011)
4. Costin, A., Francillon, A.: Ghost in the air (Traffic): On insecurity of ADS-B protocol and practical attacks on ADS-B devices. Black hat USA, pp. 1–12 (2012)
5. Magazu III, D.: Exploiting the automatic dependent surveillance-broadcast system via false target injection (2012)
6. Schäfer, M., Lenders, V., Martinovic, I.: Experimental analysis of attacks on next generation air traffic communication. In: International Conference on Applied Cryptography and Network Security. Springer, Berlin, Heidelberg, pp. 253–271 (2013)
7. Strohmeier, M., Schäfer, M., Lenders, V., et al.: Realities and challenges of next gen air traffic management: the case of ADS-B. IEEE Commun. Mag. **52**(5), 111–118 (2014)
8. Strohmeier, M., Lenders, V., Martinovic, I.: On the security of the automatic dependent surveillance-broadcast protocol. IEEE Commun. Surv. Tutor. **17**(2), 1066–1087 (2014)
9. Li, T., Wang, B.: Sequential collaborative detection strategy on ADS-B data attack. Int. J. Critical Infrastruct. Protect. (2018)
10. Strohmeier, M., Schafer, M., Fuchs, M., et al.: OpenSky: a Swiss army knife for air traffic security research. In: 2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC). IEEE (2015)
11. Yu, Y., Si, X., Hu, C., Zhang, J.: A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput. **31**(7), 1235–1270 (2019)
12. Habler, E., Shabta,i A.: Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages. Comput. Secur. **78**(sep.), 155–173 (2017)
13. Tax, D.M.J., Duin, R.P.W.: Support vector data description. Mach. Learn. **54**(1), 45–66 (2004). https://doi.org/10.1023/B:Mach.0000008084.60811.49