Constructing an ADS-B Attack Baseline to Evaluate Security Measures

Hessa Bani Hammad
Computer Science, NYUAD
hbh253@nyu.edu

Advised by: Christina Pöpper, Secondary Grader: Yasir Zaki

ABSTRACT

Although the Automatic Dependent Surveillance-Broadcast (ADS-B) technology aims to improve the capabilities of conventional air traffic surveillance systems and improve the accuracy of detecting an aircraft , research has established that attacks on ADS-B are inexpensive and highly successful [4]. As such, the system lacks security and is susceptible to various attacks. Previous research has proposed methods to detect attacks on the ADS-B system [2, 3, 8–10] to improve its security and mitigate the associated risks. This paper focuses on machine-learning based algorithms to evaluate existing proposals for ADS-B spoofing detection attacks. An important step for this is deriving training and testing datasets that incorporate anomalous labeled messages (attacker modified). These datasets are then fed into three supervised classifiers for classification: Support Vector Machines (SVM), Naive Bayes and K-Nearest-Neighbor (KNN). The attack data sets represent GPS spoofing attacks with three different types of anomalies: random noise, vertical drift and different routes. The machine learning classifiers are successful in detecting both the random noise and different routes anomalies while performing less accurately for the vertical drift anomaly attack.

KEYWORDS

ADS-B, OpenSky, Air-Traffic Control, machine learning, anomaly detection, spoofing attacks, K-Nearest-Neighbor

This report is submitted to NYUAD's capstone repository in fulfillment of NYUAD's Computer Science major graduation requirements.

Capstone Project 2, Spring 2021, Abu Dhabi, UAE

Reference Format:
Hessa Bani Hammad. 2021. Constructing an ADS-B Attack Baseline to Evaluate Security Measures. In *NYUAD Capstone Project 2 Reports, Spring 2021, Abu Dhabi, UAE.* 10 pages.

1 INTRODUCTION

The increasing air traffic of our airspace requires additional and more developed air traffic management techniques. At the same time, civil aviation faces an increasing security risk of terrorist attacks and attacks of other forms, requiring protection and security measures. To do so and meet future demands, air traffic control (ATC) is moving from conventional air-traffic management systems to a system that uses calculated GPS information (NextGen in the US and SESAR in Europe) [4].

Conventional radar surveillance technologies can be classified as primary surveillance radars (PSR) or secondary surveillance radars (SSR) [4]. PSRs transmit high-frequency signals that are reflected by the target and provide information regarding the range, angular direction, velocity and even the size and shape of the target. On the other hand, SSR relies on transponders in the aircraft, which respond to interrogations from the ground stations. Responses include information about the accurate altitude, identification code and other technical problems. SSR requires the full cooperation from the aircraft to function properly [4].

To obtain a higher surveillance rate and better accuracy as compared to PSRs and SSRs, a new air traffic surveillance system that relies on a satellite-based navigation system has been developed, known as the Automatic Dependent Surveillance-Broadcast (ADS-B) technology. In ADS-B, an aircraft continuously determines its own position using a Global Navigation Satellite System (GNSS) and then broadcasts it periodically using ADS-B transceivers on-board the aircraft [7] over a radio frequency to ground stations or other air-crafts in proximity [5]. Therefore, one of the main advantages of ADS-B is that it eliminates the need for expensive and inaccurate PSR and SSR [5]. As a result, the situational awareness of pilots and air traffic controllers significantly

improves while reducing the costs of air traffic surveillance [5], making ADS-B a preferred system.

While ADS-B aims to enhance the capabilities of conventional air traffic surveillance systems and improve the accuracy of detecting an aircraft, the system has many vulnerabilities. The issue with the ADS-B system is that it is susceptible to various radio frequency attacks on both the logical and physical layers, which introduces many security weaknesses to the system [7]. Additionally, a security issue regarding the ADS-B system is that the messages are broadcasted as unencrypted plaintexts [1], which makes them susceptible to a range of attacks, thus compromising the security of civil aviation. This imposes a significant problem as in 2010, the Federal Aviation Administration (FAA) published a final rule mandating that by 2020 all aircraft should be equipped with ADS-B [4] and most aircraft manufacturers target a complete equipage by 2020 [4].

Moreover, given the technical progress made in the past decades and the availability of low-cost software-defined radio security attacks, the ADS-B system is thus compromised and this alongside its mandatory deployment in 2020 shows that there is a growing concern about the topic. Even though researchers and scholars have proposed a number of methods and countermeasures regarding the security of ADS-B, it is evident that there is not one feasible and optimal solution when considering countermeasures that have completely no or slight impacts on the current ADS-B software and hardware [5].

In this capstone project, an ADS-B attack baseline is constructed to evaluate security measures. This is done by deriving training and testing datasets for three different GPS spoofing attacks: random noise, different route and vertical drift. The datasets were then fed into the machine learning classifiers to classify whether a given message is fake or legitimate.

2 RELATED WORK

Several methods have been proposed to detect different attacks on the ADS-B system [2, 8–10]. These methods include both machine learning based approaches and other techniques. These approaches can be deployed to mitigate the attacks on the ADS-B system.

Machine Learning Based In [10], the authors propose a Deep Neural Network based spoofing detector to detect spoofing attacks that target ADS-B ground stations. In this attack, the ground-based or aircraft-based attacker manipulates the International Civil Aviation Organization (ICAO) address, a unique identifier for each aircraft in the ADS-B messages to fake the appearance of non-existent aircraft or disguise as a legitimate aircraft. The introduced SODA (a two stage deep neural network based spoofing detector) consists of a message classifier and an aircraft classifier. The approach is successful in detecting ground-based spoofing attacks with

Table 1: Related Works

Title	Classification	Summary
Detecting ADS-B Spoofing Attacks using Deep Neural Networks.	Machine learning based	SODA, a deep neural network based spoofing detector consisting of a message and aircraft classifier to detect ground and aircraft based spoofing attacks is introduced.
A localization approach for crowdsourced air traffic communication networks.	Machine learning based	A grid based localization approach using a combination of the k-Nearest Neighbor algorithm and the expected time differences of arrival of ATC signals to estimate the origin of aircraft based on their wireless communication.
Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages.	Machine learning based	A Long Short-Term Memory (LSTM) encoder-decoder algorithm to detect spoofed or manipulated ADS-B messages
Intrusion detection for airborne communication using PHY-layer information.	Non-machine learning based	Intrusion detection system using statistical testing of received signal strength to detect false data injection attacks in 40 seconds.

a probability of 99.3%. It outperforms other machine learning based approaches like XGBoost, Logistic Regression, and Support Vector Machines.

The machine learning model that is applied in [10], is a deep neural network model, while this capstone project explores three different machine learning models: K-Nearest-Neighbor, Naïve Bayes and Support Vector Machines to classify malicious ADS-B messages. Additionally, the authors in [10] introduce SODA, a deep neural network based spoofing detector for ADS-B which consists of a message classifier and aircraft classifier to detect the ground based and aircraft based spoofing attacks.

In the first stage, the message classifier examines each message and labels it as malicious or non malicious. For those that are considered non-malicious, they are processed by the aircraft classifier which predicts the ICAO address based on the physical layer features and compares the prediction against the claimed address to detect aircraft spoofing attacks, thus making it a two stage classifier, which is different from the goal of this capstone project.

The paper focuses on three different types of ADS-B spoofing attacks: message of IQ data replay attack, ghost aircraft injection attack and aircraft spoofing attack. The first two attacks are launched by a ground-based attacker, while the last one is launched by an aircraft-based attacker. Contrastingly, this capstone exploration focuses on three different types of GPS spoofing attacks: injecting random noise in the positional values (latitude and longitude) of an aircraft, changing the legitimate locations of aircraft by a given km and bearing angle and lastly applying a vertical shift on the altitude values of an aircraft.

In order to successfully train the message classifier, the authors in [10] constructed an ADS-B receiver (ground stations) to collect real ADS-B messages in an open area. Then, they built a SDR based ADS-B spoofer to emulate ADS-B spoofing attacks. In this way, a collected dataset was built to feed into the model. In this capstone project, an existing dataset of real ADS-B data collected by the OpenSky Network [6], a large-scale ADS-B sensor network for research is used. The OpenSky network continuously collects air traffic control data from thousands of aircraft. This data is received and streamed to OpenSky over the Internet. Moreover, the attack dataset in this capstone project is created manually by injecting random noise, shifting the locations of ADS-B messages by a given km and bearing angle as well as vertically shifting the altitude value of the ADS-B messages.

In another paper [9], a K-Nearest Neighbor (k-NN) based Localization Approach for Crowdsourced Air Traffic Communication Networks is proposed. The authors argue that the current methods of air traffic localization such as multilateration are insufficient for modern crowdsourced air traffic networks with random, unplanned deployment geometry, thus the need for another approach. The (k-NN) based localization approach that is proposed is verified by its aircraft location accuracy, surveillance coverage and the verification of false position data. The (k-NN) based localization method localizes aircrafts based on their wireless signals using a combination of the k-Nearest Neighbor algorithm and the expected time difference of arrival (TDoA) of ATC signals to estimate their origins. The proposed approach improves detection speed. The authors study the effectiveness of verifying air traffic control data by showing that verifying legitimate and false flight data using the (k-NN) localization approach is quicker and can provide improved localization of the origin of false data injections on or near the ground.

The authors in [9] develop a new grid-based method to localize aircraft based on their wireless communication through a new combination of the k-Nearest Neighbor algorithm and the expected time differences of arrival of ATC signals to estimate their origin. They utilize the k-Nearest Neighbor algorithm which is also used in this capstone project, in a different way as it is combined with the expected time differences of arrival of ATC signals to estimate the origin of the signal. The main goal in [9] is to verify air traffic control data by verifying legitimate and false flight data using the localization method while in this capstone project, three different machine learning classifiers are used to classify legitimate ADS-B messages from injected ones. In [9], the authors take into consideration deviated location attacks which are mainly GPS spoofing attacks. Similarly in this capstone project, three different GPS spoofing attacks are considered (injecting random noise, changing the flight route and vertical drift). Changing the flight route attack in this capstone project is similar to the attack model in this paper. Furthermore, the authors use real-world flight data collected from OpenSky, which is something that is also followed in this capstone project.

In [2], a Long Short-Term Memory (LSTM) encoder-decoder algorithm is proposed to detect spoofed or manipulated ADS-B messages sent by an attacker or compromised airplane. Using this model, an aircraft can autonomously evaluate received ADS-B messages and identify deviations from the legitimate flight path (anomalies). To evaluate the performance of the learned LSTM encoder decoder model, the authors injected three types of anomalies including introducing random noise (multiplying the values of the message attributes of the ADS-B messages), modifying the altitude of an aircraft and replacing a segment of the ADS-B messages of the tested flight with a segment of messages from a different (legitimate) route. The outlined approach is successful in detecting injected attacks using a deep learning model.

In this capstone project, a similar approach to that of in [2] is used to introduce anomalies. While the authors in [2] use an LSTM encoder-decoder algorithm for modeling flight

routes by analyzing sequences of legitimate ADS-B messages and identifying deviations (anomalies), this capstone project identifies anomalies in legitimate ADS-B messages using three different machine learning classifiers. In order to evaluate the performance of the learned modes (LSTM encoder-decoder), the authors introduce anomalies by injecting three types of anomalies: random noise, different route and gradual drift. In this capstone project, a similar approach is followed. The authors in [2] generate random noise by multiplying the original values of the message attributes of the ADS-B messages with a randomly generated floating number between 0 and 2. This approach is also followed in this capstone project where anomalies are introduced by multiplying the latitudinal and longitudinal values by a randomly generated floating number between 0 and 2. Moreover, the authors introduced anomalies by replacing a segment of the ADS-B messages of the tested flight with a segment of messages from a different (legitimate route). On the other hand, in this capstone anomalies are generated by shifting the original latitude and longitudinal values of the ADS-B messages by a fixed bearing angle and a fixed distance in kms. Additionally, the authors introduce anomalies as a gradual drift in the altitude feature by modifying the altitude of a segment of messages by continuously raising/lowering the altitude by an increasing multiplier of 400 feet. On the contrary, this capstone project introduces a vertical drift anomaly where the value of the altitude was raised/lowered by an increasing multiplier of 100 meters.

3 METHODOLOGY

To successfully detect anomalous ADS-B messages and evaluate existing proposals of anomaly detection, an existing dataset of real ADS-B data collected by the OpenSky Network [6], a large-scale ADS-B sensor network for research is used. The OpenSky network continuously collects air traffic control data from thousands of aircraft. This data is received and streamed to OpenSky over the Internet.

3.1 Dataset

Collected dataset: The OpenSky Network ADS-B dataset has been used to create three attack datasets. In this capstone project, only GPS spoofing attacks are represented. Three different types of anomalies as part of the GPS spoofing attack are created.Anomalies in the data are injected in three different ways: (1) random noise, (2) altering the positional data (latitude and longitude) by a specified distance and bearing, and (3) vertically drifting the altitude of an aircraft. The collected ADS-B data is in the form of a CSV file containing 15245 messages, where each row represents the reception of one aircraft position report. The messages are mainly intercepted from European countries where ADS-B sensors are mostly deployed. The geographical areas include Germany, The United Kingdom, France, Belgium, Norway and Switzerland. The collected ADS-B data consists of 8 attributes. Table 2 describes the features of the ADS-B flight data with their corresponding data types.

Table 2: ADS-B Data Features

Feature	Type
Aircraft ID	Integer
Timestamp	Float
Latitude	Float
Longitude	Float
Barometric altitude	Float
Geometric altitude	Float
Signal strength measurements from each of the sensors (RSS)	Integer
The number of sensors receiving the signal	Integer

3.2 Attack Data Generation

Injected dataset: Since the collected ADS-B data represents legitimate aircraft messages, under the assumption that no attack on ADS-B was happening at that geographic area at that time, there is a need to create “fake" (attacker modified) ADS-B messages to represent the attack data sets.Therefore, based on the collected (legitimate) dataset, three different datasets with attack data are created. The three different attack datasets represent GPS spoofing attacks, where an attacker injects false location data or alters the positional data of an aircraft.

There are different approaches of creating attack data sets which can be achieved through the manipulation of different features in the ADS-B messages. Aircraft spoofing attacks involve an aircraft-based attacker (malicious aircraft) which attempts to hide as a known or trusted aircraft by spoofing the ICAO address. Other attacks involve modifying the RSS values of the ADS-B messages.

Random noise: To create the random noise attack data set, the original dataset of legitimate ADS-B messages from the OpenSky network is modified to include injected false longitudinal and latitudinal position values. This simulates a GPS spoofing attack where an attacker alters the position data of an aircraft. To create the attack data set, I generate anomalies by adding random noise to the both the longitudinal and latitudinal values as proposed in [2]. The original latitude and longitude values are multiplied by a randomly generated floating number between 0 and 2. This introduces random noise to the positional data. Moreover, this results

up to 2,600 km drifts between the original locations and the spoofed locations. When multiplied by the randomly generated number, the positional data (longitude and latitude) are altered and thus new locations are introduced which generates anomalous locations to the original legitimate dataset.

Different route: To create the different route attack datasets, the original dataset of legitimate ADS-B messages from the OpenSky network is modified by replacing the original positional values (latitude and longitude) in a systematic manner. To do this, the original latitude and longitude coordinates were shifted by a fixed distance in kms and a fixed bearing angle. The bearing of an aircraft refers to the horizontal direction from a specified location, measured from true north or any other reference point through 360 degrees. In other words, it is the horizontal angle between the direction of an object and another object, or between it and that of true north. By shifting the positional values by a fixed distance and a fixed bearing angle, the ADS-B messages are shifted to a new location, thus changing the route of an aircraft. Instead of intercepting the ADS-B messages in the European geographical area, the messages are shifted to a new geographical area depending on the fixed distance and bearing angle. This represents a more sophisticated attack than just injecting random noise in the data as it shifts the original latitudinal and longitudinal values by a set distance which changes the original route as opposed to introducing random anomalies in random locations.This approach also has the benefit that it would not be detectable by simply checking big jumps in consecutive location reports which makes it harder to identify. This approach is similar to that of [2], where the injected anomalies are created by replacing a segment of the ADS-B messages with a segment of messages from a different (legitimate) route. However, in the generated attack dataset, the legitimate messages are replaced by false positional values, hence producing a fake flight route.

Vertical Drift: To create the vertical drift attack data set, the original dataset of legitimate ADS-B messages is modified by altering the geometric altitude value. To create this attack data set in a systematic approach, the value of the altitude was raised/lowered by an increasing multiplier of 100 meters, which is also a similar approach to that in [2], where the authors generated anomalies by modifying the altitude of a segment of messages by continuously raising/lowering the altitude by an increasing multiplier of 400 feet (the first message is increased/decreased by 400 feet, the second message will be increased/decreased by 800 feet and so on). This attack is a more sophisticated one, in the sense that it attempts to affect the air-space view by adding vertically-drifted ADS-B messages which are less observable. Also, this attack may result in collisions in the air, introducing noise to the ADS-B messages.

3.3 Adversary Model

By selecting these three different types of anomalies as part of the GPS spoofing attack, two types of attackers are represented. The first attacker is a naive adversary for both the random noise and the route change anomalies. The main goal of this attacker is to add noticeable noise to the airspace to reduce the legitimacy of the ADS-B system and interrupt the traffic management. A total of 15245 ADS-B messages are considered. The modified datasets consisting of the three different types of anomalies with legitimate and altered ADS-B messages are used for both training and testing purposes with a 80/20 split for the training and testing data. Figure 1 shows the process flow of the procedure to classify the ADS-B messages.

3.4 Machine Learning Classification Models

To evaluate previous proposals of attack detection and compare them, this capstone project evaluates the detection of three types of GPS spoofing attacks using three different supervised machine learning classifiers to assess the best approach of detecting GPS spoofed attacks and exploring which type of anomaly can be detected easily.

K-Nearest-Neighbor The k-Nearest Neighbor Algorithm (KNN) is used to classify labeled ADS-B messages. The KNN algorithm is a method for classifying objects based on their close proximity. The k-nearest-neighbor classifier is based on the Euclidean distance between a test sample and the specified training samples. Suppose we are given a new test sample and we want to find the categorical class that it belongs to. The idea in k-Nearest-neighbor is to identify k samples in the training set whose independent features are similar to the new test and to use the k samples to classify the test sample into the categorical class it belongs to, which is "real" or "fake" in the context of this experiment. Then, we would look for samples in the training dataset that are in close proximity to the test sample to categorize it using the labelled training datasets. To compute the distance between the samples based on the independent features of the data, we compute the Euclidean distance between the test sample and the training sample which is:

$$d(x, u) = \sqrt{\sum_{i=1}^{n} (x_i - u_i)^2}$$

Where x represents a new data point and u represents all the other data points in the dataset and n represents the number of features in the data. It is important to select the right value of k for the data and to do so, we run the KNN algorithm several times with different values of k and

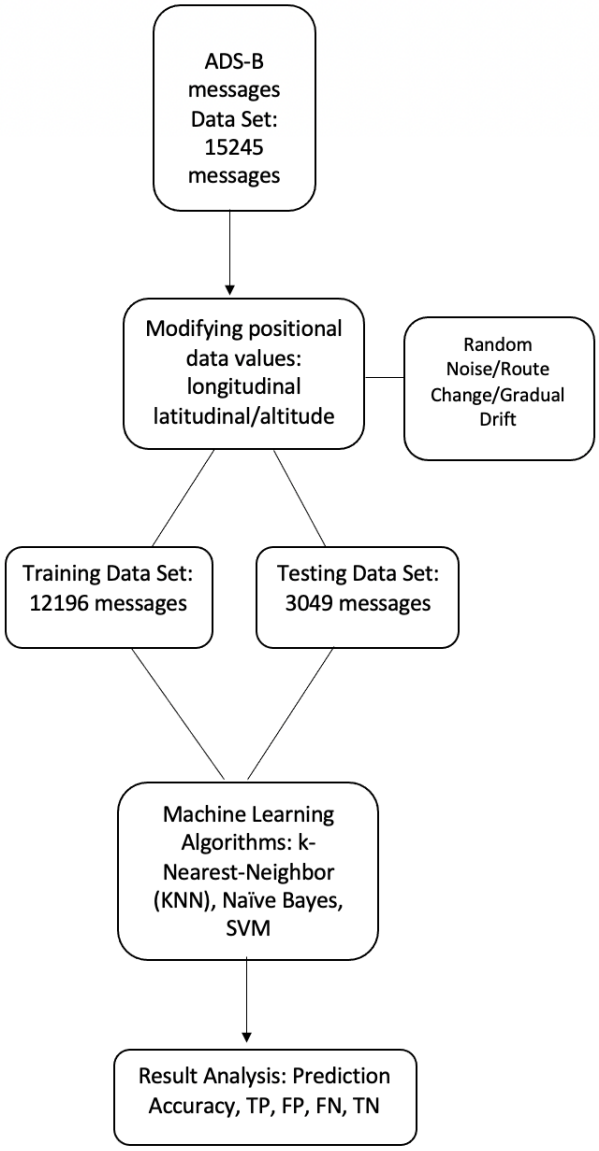

Figure 1: General process flow of the classifier for the three different types of anomalies

choose the k value that reduces the number of errors while maintaining the algorithm's ability to make predictions.

Naïve Bayes The Naïve Bayes classifier is based on the Bayes rule of conditional probability. It utilizes all the attributes in the data and analyzes them individually and independent of each other. Simply, it assumes that the presence of a feature is unrelated to the presence of any other feature. Bayes theorem calculates the posterior probability $P(c|x)$

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$$P(c|X) = P(x_1|c)P(x_2|c)...P(x_n|c)P(c)$$

of the target variable given the predictor (data features). The value c representing the target variable and x representing the data attribute to be classified. After calculating the posterior probability, the class (target variable) with the highest posterior probability is the outcome of the prediction.

Support Vector Machines (SVM) Support Vector Machines are supervised learning methods used for classification and regression. The main goal of SVMs is to find a hyper-plane in an N-dimensional space (N number of features) that classifies the data points. Support vectors are data points that lie closer to the hyper-plane and influence the position and the orientation of the hyper-plane. These support vectors are then used to maximize the margin of the classifier. To classify the data samples in SVM, the output of the linear function is taken into consideration, if its greater than 1, then the data sample is classified into a class, if its -1, the data sample is identified with the other class. The margin between the data points and the hyper-plane is maximized using a loss function.

4 EVALUATION

The three machine learning classifiers were trained using the training data set which included the labels of "fake" and "legitimate" data with the values 1 representing a fake message and 0 representing a legitimate message. The attack dataset created for the three different types of anomalies consisted of 15, 245 ADS-B messages with a 80/20 split for the training and testing data. The training data consisted of 12,196 ADS-B messages, while the testing data consisted of 3,049 ADS-B messages.

Random noise: In the random noise attack dataset, the following experimental results shows that the k-nearest-neighbor classifier performs the best in classifying anomalous positional ADS-B data messages with a prediction accuracy of 97.76%. The performance of the three classifiers is summarized in Table 3, where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positive, FN is the number of false negatives and the prediction accuracy of the classifier is the accuracy in percentage of the classifier being able to predict if a certain ADS-B message is fake or legitimate.

Different route: For the different route anomaly, the legitimate ADS-B messages are replaced by a different "fake" aircraft route. In the legitimate dataset, the intercepted ADS-B messages are mainly coming from the European geographical area. For example, most of the messages are intercepted from Germany, Austria, Switzerland, Italy, France and the

Table 3: Random Noise Anomaly Results

Machine Learning Algorithm	TP	FP	FN	TN	Precision	Recall	Prediction Accuracy
k-Nearest Neighbors	1520	38	30	1461	0.975	0.98	97.76%
Naïve Bayes	1505	46	11	1430	0.97	0.99	96.12%
Support Vector Machines	987	492	525	1045	0.66	0.65	66.6%

United Kingdom indicating aircraft flying in that route. By displacing the locations, an attacker can alter the flight route of the aircraft, indicating that an aircraft is flying over a specific geographical area while it is originally flying in a different route. In other words, the attacker masks the legitimate locations of the aircraft. To simulate this attack, four different attack datasets are created with modified distances and bearing angles. To represent more stealthy attacks, displacing the distance by certain kms (displacing the latitude and longitude) by specified kms is not enough. Therefore, the positional values (latitude and longitude) are displaced by a distance and a bearing angle. The bearing angle shift allows the positional values to be shifted in a specific degree clockwise from the north, instead of just displacing them horizontally. The positional values are then completely shifted to new geographical areas.

90°bearing: The first two datasets displace the legitimate flight route by a bearing angle of 90°. The first dataset displaces the original positional values by 2000 kms while the second displaces them by 4000 kms. Figure 4 depicts the generated "fake" locations on the map. As seen in Figure 4 the original ADS-B messages have been shifted from the European air space eastwards to the Black Sea. The distance is increased in the second dataset while keeping the bearing angle the same to see whether the displacement of the positional values affects the accuracy of the machine learning classifiers in predicting whether a given ADS-B message is legitimate or fake. Figure 3 shows the displaced "fake" locations by displacing the original ADS-B messages by 4000 kms. As seen in Figure 3, the flight route is changed from the European airspace to the Caspian sea and Turkmenistan and parts of northern Iran. Table 4 and Table 5 show the experimental results of classifying the anomalous ADS-B data by shifting the positional values by 2000 and 4000 kms respectively. Evaluating the results, the three machine learning classifiers are successful in classifying the changed route attack. Even when the distance is increased to 4000 kms, the classifiers are still successfully able to classify the ADS-B messages as "fake" as seen in Table 4 and Table 5.

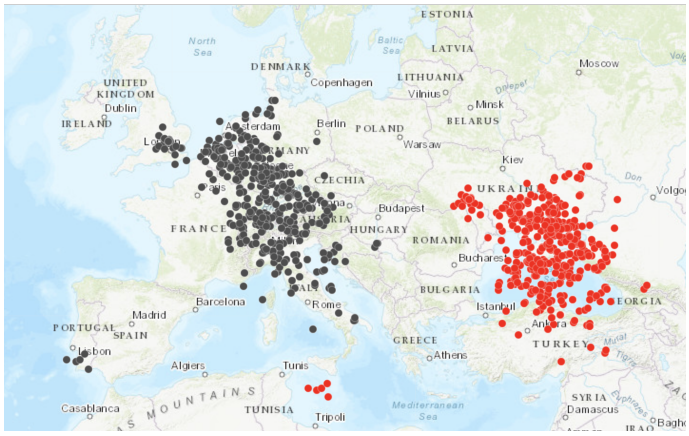

Figure 2: Displaced locations by 2000 kms and 90°bearing

Table 4: Different Route Anomaly Results (2000 km and 90°bearing

Machine Learning Algorithm	TP	FP	FN	TN	Precision	Recall	Prediction Accuracy
k-Nearest Neighbors	1570	25	31	1423	0.98	0.98	98.5%
Naïve Bayes	1369	46	54	1580	0.96	0.96	96.7%
Support Vector Machines	1679	23	26	1321	0.98	0.98	98.3%

180°bearing: The second two datasets displace the legitimate flight route by a bearing angle of 180°. Therefore, the original ADS-B messages have been shifted 180°from their original positional values (latitude and longitude). The bearing angle has been increased from the first two attack datasets (90°bearing) to explore whether changing the angle and distances has an effect on the accuracy of the machine learning classifiers in classifying the anomalous ADS-B messages. The third dataset displaces the original ADS-B messages by 2000 kms while the fourth displaces them by 4000 kms. Figure 4 depicts the generated "fake" locations on the map. As seen in Figure 4, the original ADS-B messages are shifted from the European airspace (where they have been originally intercepted), southwards towards parts of Spain and the Northern part of Algeria with some locations shifted

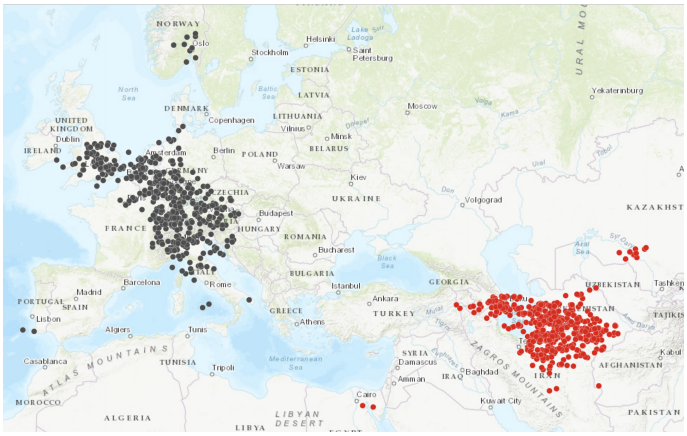

Figure 3: Displaced locations by 4000 kms and 90°bearing

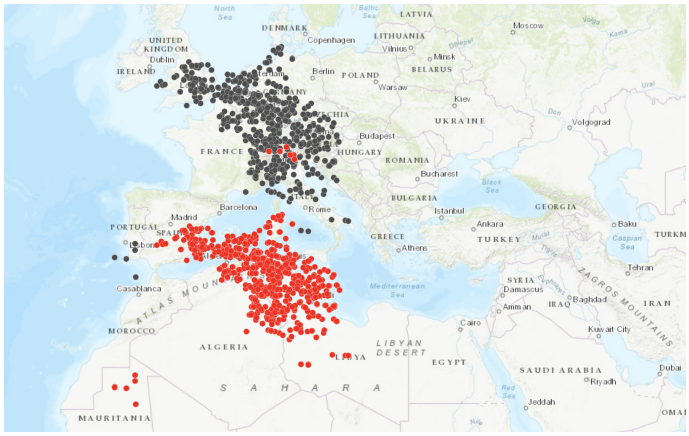

Figure 4: Displaced locations by 2000 kms and 180°bearing

Table 5: Different Route Anomaly Results (4000 km and 90°bearing

Machine Learning Algorithm	TP	FP	FN	TN	Precision	Recall	Prediction Accuracy
k-Nearest Neighbors	1489	15	24	1521	0.98	0.98	98.7%
Naïve Bayes	1345	48	71	1585	0.96	0.96	96.1%
Support Vector Machines	1724	12	39	1234	0.98	0.98	95.6%

to Libya. The distance is increased in the third dataset while keeping the bearing angle the same to see whether the displacement of the positional values affects the accuracy of the machine learning classifiers in predicting whether a given ADS-B message is legitimate or fake. Figure 5 shows the displaced "fake" locations by displacing the original ADS-B messages by 4000 kms. As seen in Figure 5, the flight route is changed from the European airspace to West Africa, specifically in Niger, Mali, Nigeria and Chad. Table 6 and Table 7 show the experimental results of classifying the anomalous ADS-B data by shifting the positional values by 2000 and 4000 kms respectively. Evaluating the results, the three machine learning classifiers are successful in classifying the changed route attack. Even when the bearing angle is increased to 180°and the displaced distance is increased up to 4000 kms, the classifiers are still successfully able to classify the ADS-B messages as "fake" as seen in Table 6 and Table 7.

Table 6: Different Route Anomaly Results (2000 km and 180°bearing

Machine Learning Algorithm	TP	FP	FN	TN	Precision	Recall	Prediction Accuracy
k-Nearest Neighbors	1621	10	17	1401	0.98	0.98	99.5%
Naïve Bayes	1537	16	38	1458	0.96	0.96	98.3%
Support Vector Machines	1421	17	14	1597	0.98	0.98	98.5%

Table 7: Different Route Anomaly Results (4000 km and 180°bearing

Machine Learning Algorithm	TP	FP	FN	TN	Precision	Recall	Prediction Accuracy
k-Nearest Neighbors	1479	15	26	1529	0.98	0.98	98.6%
Naïve Bayes	1569	45	35	1400	0.97	0.97	97.3%
Support Vector Machines	1345	69	32	1603	0.95	0.97	96.6%

Vertical drift: In the vertical drift anomaly, the attack dataset was created by modifying the value of the altitude and therefore, only one feature is taken into account when fed into the

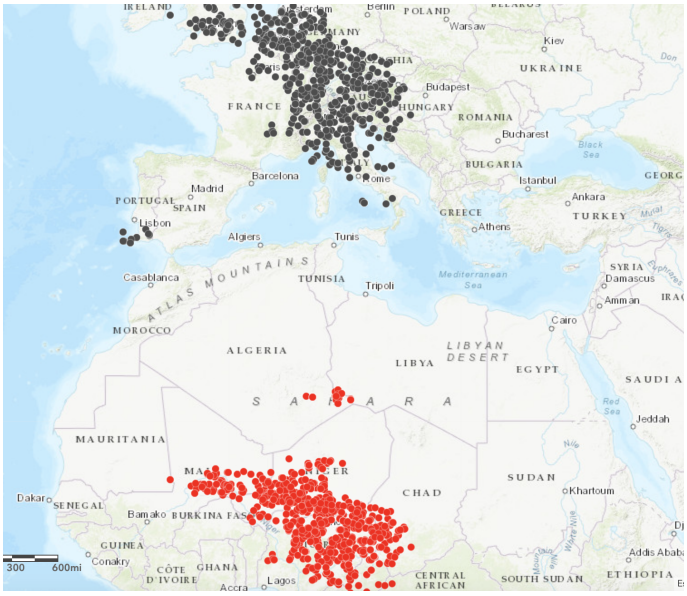

Figure 5: Displaced locations by 4000 kms and 180°bearing

three different machine learning classifiers. As opposed to both the random noise anomaly and the route change attack where the latitude and longitude values are the two features taken into account when fed into the machine learning classifiers, the gradual drift attack data set has only feature which is the altitude. Because there is only one feature that is affected in this attack, the detection accuracy of the machine learning classifiers significantly decreased. Table **??** summarizes the performance of the three machine learning classifiers in detecting the attack.

Table 8: Vertical Drift Anomaly Result

Machine Learning Algorithm	TP	FP	FN	TN	Precision	Recall	Prediction Accuracy
k-Nearest Neighbors	1321	521	504	703	0.71	0.72	66.3%
Naïve Bayes	1180	397	740	732	0.74	0.61	62.7%
Support Vector Machines	1100	542	758	649	0.66	0.59	57.3%

Examining the results of the GPS spoofing attacks in each of the three different types of anomalies: random noise, different route and the vertical drift, it is evident that the K-nearest neighbor algorithm outperform the other two machine learning classifiers (Naïve Bayes and Support vector machines). One reason to why k-nearest neighbors (KNN) outperforms Support vector machines (SVM) is that the training data is much larger than the number of features. Additionally, KNN is effective in case of large numbers of training examples, which is the case in this capstone. SVM will usually outperform KNN when there are large features and lesser training data.

Moreover, the machine learning classifiers were able to successfully classify the ADS-B messages in the two types of anomalies: random noise and different route as two features were taken into consideration: the latitudinal value and longitudinal value of an ADS-B message. The machine learning classifiers performed poorly in the vertical drift anomaly due to the single feature (altitude) that was taken into consideration for classification.

4.1 Additional Related Works

Other/Combined Approaches Other non-machine learning based methods have been proposed to detect different attacks on the ADS-B system. These approaches can be explored further as future work/extension to this capstone project. In [8], the authors developed an intrusion detection system that uses statistical testing of received signal strength (RSS) patterns to detect false data injection attacks with a single receiver in 40 seconds. The authors argue that given the current state of the ADS-B's lack of security, there is an urgent need for transparent countermeasures and thus an anomaly detection system can provide the base for defense-in-depth mechanisms. In other words, detection systems can provide the first steps to creating viable countermeasures. The authors analyze different features based on statistical tests (Pearson correlation coefficient, autocorrelation coefficient and antenna detection). Then, they combine them into a unified approach using one-class anomaly detection which uses several different machine learning classifiers with 5-fold cross validation to create a stronger intrusion detection system. The intrusion detection system (IDS) is then validated using real-world data from the authors' OpenSky sensor network. The authors assume that the attacker injects a ghost aircraft collected at an earlier time and replayed or created it from scratch. Additionally, the attacker uses different RSS patterns to inject false data, therefore using hypothesis testing, the intrusion detection system can classify the collected RSS sample as legitimate or fake. Because an attacker is stationed on the ground, the measurements of the injected ADS-B messages are unlikely to match legitimate RSS samples as they would be constant over time as opposed to an aircraft that is covering distances of miles in relation to the receiver. The combined detection method was able to

accurately detect all attackers with the Parzen classifier performing the best, followed by K-Means, Minimax, Minimum Spanning Tree and the k-Nearest Neighbors classifiers.

5 CONCLUSION

Considering the technical progress made in the past decades, such as the availability of low-cost software-defined radios and the different possible attacks on ADS-B messages, it is evident that the ADS-B system is vulnerable to security attacks and requires a baseline to evaluate the proposed countermeasures fairly. This capstone project evaluated the accuracy of classifying anomalous ADS-B data by using machine learning models through simulating a GPS spoofing attack scenario by creating three different types of anomalies: random noise, changed route and gradual drift in altitude.

REFERENCES

[1] Andrei Costin and Aurélien Francillon. 2012. Ghost in the Air (Traffic): On insecurity of ADS-B protocol and practical attacks on ADS-B devices. *Black Hat USA* (2012), 1–12.

[2] Edan Habler and Asaf Shabtai. 2018. Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages. *Computers & Security* 78 (2018), 155–173.

[3] Mohsen Riahi Manesh, Mahdi Saeedi Velashani, Elias Ghribi, and Naima Kaabouch. 2019. Performance Comparison of Machine Learning Algorithms in Detecting Jamming Attacks on ADS-B0 Devices. In *2019 IEEE International Conference on Electro Information Technology (EIT)*. IEEE, 200–206.

[4] Matthias Schäfer, Vincent Lenders, and Ivan Martinovic. 2013. Experimental analysis of attacks on next generation air traffic communication. In *International Conference on Applied Cryptography and Network Security*. Springer, 253–271.

[5] Matthias Schäfer, Martin Strohmeier, Vincent Lenders, Ivan Martinovic, and Matthias Wilhelm. 2014. Bringing up OpenSky: A large-scale ADS-B sensor network for research. In *IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*. IEEE, 83–94.

[6] Matthias Schäfer, Martin Strohmeier, Vincent Lenders, Ivan Martinovic, and Matthias Wilhelm. 2014. Bringing up OpenSky: A large-scale ADS-B sensor network for research. 83–94. https://doi.org/10.1109/IPSN.2014.6846743

[7] Martin Strohmeier, Vincent Lenders, and Ivan Martinovic. 2014. On the security of the automatic dependent surveillance-broadcast protocol. *IEEE Communications Surveys & Tutorials* 17, 2 (2014), 1066–1087.

[8] Martin Strohmeier, Vincent Lenders, and Ivan Martinovic. 2015. Intrusion detection for airborne communication using PHY-layer information. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*. Springer, 67–77.

[9] Martin Strohmeier, Vincent Lenders, and Ivan Martinovic. 2016. A localization approach for crowdsourced air traffic communication networks. *arXiv preprint arXiv:1610.06754* (2016).

[10] Xuhang Ying, Joanna Mazer, Giuseppe Bernieri, Mauro Conti, Linda Bushnell, and Radha Poovendran. 2019. Detecting ADS-B Spoofing Attacks using Deep Neural Networks. In *2019 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 187–195.