Test Data Generation for False Data Injection Attack Testing in Air Traffic Surveillance

Aymeric Cretin
FEMTO-ST Institute, UBFC, CNRS
Besançon, FRANCE
aymeric.cretin@femto-st.fr

Alexandre Vernotte
FEMTO-ST Institute, UBFC, CNRS
Besançon, FRANCE
alexandre.vernotte@femto-st.fr

Antoine Chevrot
FEMTO-ST Institute, UBFC, CNRS
Besançon, FRANCE
antoine.chevrot@femto-st.fr

Fabien Peureux
FEMTO-ST Institute, UBFC, CNRS
Besançon, FRANCE
fabien.peureux@femto-st.fr

Bruno Legeard
FEMTO-ST Institute, UBFC, CNRS
Smartesting Solutions and Services
Besançon, FRANCE
bruno.legeard@femto-st.fr

***Abstract*—The ADS-B — Automatic Dependent Surveillance Broadcast — technology requires aircraft to broadcast their position and velocity periodically. The protocol was not specified with cyber security in mind and therefore provides no encryption nor identification. These issues, coupled with the reliance on aircraft to communicate on their status, expose air transport to new cyber security threats, and especially to FDIAs — False Data Injection Attacks — where an attacker modifies, blocks, or emits fake ADS-B messages to dupe controllers and surveillance systems. This paper is part of an ongoing research initiative toward the generation of FDIA test scenarios and focuses on the test generation activity, i.e. providing the mechanisms to alter existing ADS-B recordings *as if* an attacker had tempered with the communication flow, in order to improve the detection capabilities of surveillance systems. We propose a set of alteration algorithms covering the taxonomy of FDIA attacks for ADS-B previously defined in the literature. We experiment this approach by generating test data for an AI-based FDIA detection system [9]. Experimental results show that the proposed approach is straightforward to generate the initial situations used to validate the detection system. Moreover, it provides a efficient way to easily generate sophisticated alterations that were not picked up by the detection system.**

***Keywords*—AI testing, test data generation, false data injection attacks, air surveillance.**

I. Introduction

The world of air transport is facing new challenges as the traffic load keeps growing steadily[1]. With an increasingly congested airspace, Air Traffic Control (ATC) needs surveillance technologies that can support the increasing constraints in terms of simultaneously handled aircraft as well as positioning accuracy. The Automatic Dependent Surveillance-Broadcast (ADS-B) protocol is currently being rolled out in an effort to reduce costs and improve aircraft position accuracy [22]. Communication via ADS-B consists of participants broadcasting their current position and other information periodically (a.k.a. a beacon) in an unencrypted message [15].

The fundamental technological changes in ATC needed to support increasing traffic, which consist of a shift from independent and non-cooperative surveillance technologies to dependent and cooperative ones, have rendered the ATC community unable to foresee the new emerging threats related to cyber security. The ADS-B protocol was not designed with security in mind since securing ADS-B communication was not a high priority during its specification. As a consequence, anyone with the right equipment can listen and emit freely. For instance, there is a market for equipping private aircraft with ADS-B transponders using a smartphone and a dongle[2]. The complete freedom of ADS-B both in emission and reception makes it vulnerable to spoofing, and more precisely to a class of attack called FDIA — False Data Injection Attack — which purpose is to covertly emit meticulously-crafted fake surveillance messages in order to dupe ATC controllers into thinking, for instance, that some aircraft is dangerously approaching a building, while in reality it is flying normally.

Although it is not the only means for Aircraft tracking — other protocols are also used in conjunction of radar technologies —, ADS-B plays a central role in the current shift regarding aircraft position, collected from radar systems to GNSS [4]. It is so central in fact that it has become a mandatory brick of air traffic surveillance and any observed problem will ground all aircraft in the area[3]. Hence there is a strong need to improve its overall security. Nevertheless, because of the inherent properties of the protocol, current solutions for securing ADS-B communications are only partial or involve an unbearable cost [24].

Instead, ATC should be made more secure by strengthening its logic, but the ability to differentiate attacks from real critical situations still remains a challenge to be tackled by the ATC community. Indeed, multiple integrity checks or detection approaches are rolled out or under study. Those solutions are new and need to be deeply tested. To the best of our knowledge, there is no direct previously published work addressing the topic of assessing the efficiency of this kind

[1] http://www.boeing.com/commercial/market/current-market-outlook-2017/

[2] https://www.uavionix.com/products/skybeacon/

[3] https://hackaday.com/2019/06/09/gps-and-ads-b-problems-cause-cancelled-flights/

of security. The most related work is a framework proposed by Barreto et. al. [3], which allows for the simulation of an entire air traffic environment (aircraft, radio relay, network infrastructure, etc.). They perform FDIAs in the simulated environment to evaluate the attack impact on each component. The approach provides substantial information on how components react to an FDIA. Still, implementing all network behaviours of a scenario requires a lot of effort and the approach does not allow for the concretization of the simulated attacks on actual ATC software.

The contribution presented in this paper is part of an ongoing research initiative about FDIA testing that ultimately led to the creation of a testing framework called *FDI-T* [6] (False Data Injection Testing framework). This framework allows ATC experts to simulate FDIAs by creating, modifying and deleting recorded legitimate ADS-B messages in a fruitful, scalable and productive manner. The generated test scenarios can be executed on ATC systems in order to evaluate their resilience against potential security and safety anomalies related to FDIAs. They can also be used as test data to evaluate machine learning based detection systems, as well as improve their detection rate. The paper focuses on the alteration mechanisms that take as input one or several air traffic recordings and a list of alteration directives in order to simulate the presence of FDIAs in air traffic surveillance communications.

The paper is organized as follows: Section II briefly provides a basis for common concepts and current practices regarding air traffic surveillance as well as the key aspects to test such systems, especially regarding test scenarios based on FDIAs. Afterwards, Sect. III introduces the proposed approach to perform FDIAs on ADS-B messages and describes the automated process supporting the method. Section IV details the various algorithms used to apply alteration directives on ADS-B messages and thus explains how the generation of attack scenarios is automated. Section V demonstrates the ability of the proposed approach to generate test data to assess the detection capabilities of a machine learning technique from the literature [9]. Finally, Sect. VI recaps the major contributions of this paper, and suggests directions for future work.

II. Background and Research Objective

A. ADS-B

Communication via ADS-B consists of aircraft using a Global Navigation Satellite System (GNSS) to determine their position and broadcasting it periodically without solicitation (a.k.a beacons or squitters), along with other information obtained from on-board systems such as altitude, ground speed, aircraft identity, heading, etc. Ground stations pick up on the squitters, process them and send the information out to the ATC system. The ADS-B data link is generally carried on the 1090MHz Extended Squitter (1090ES), the same frequency used by Mode S, although there is a new data link standard (UAT – Universal Access Transceiver) dedicated to protocols such as ADS-B, but it requires new hardware and is not very common at the moment. ADS-B is therefore a cooperative (aircraft need a transponder) and dependent (on aircraft data) surveillance technology, which constitutes a fundamental change in ATC. It means for instance that not only ground stations with antennas positioned at the right angle and direction can receive position information. Aircraft can now receive squitters from other aircraft, which notably improves cockpit situational awareness as well as collision avoidance. For instance, the second generation of the Traffic Alert and Collision Avoidance System (TCAS-II) is based on ADS-B data.

Its introduction also provides controllers with improved situational awareness of aircraft positions in En-Route and TMA (Terminal Control Area) airspaces, and especially in NRAs (Non Radar Areas). It theoretically gives the possibility of applying much smaller separation minima than what is presently used with current procedures (Procedural Separation) [1]. Indeed, ADS-B offers position accuracy of 0.05 NM and velocity accuracy of 19.4 NM/h, with updates once to twice second. Concretely, ADS-B performance requirements were designed to allow an aircraft lateral separation from 90 to 20 NM and longitudinal separation from 80 to 5 NM in NRAs, and 5 to 3 NM in covered areas [25]. ADS-B has the advantage of being a much cheaper technology as it has minimal infrastructure requirements. For instance, an ADS-B receiver can easily be bought online for a few hundreds euros[4]. As mentioned in the previous paragraph, ADS-B has a much greater accuracy and update rate, with a smaller latency. The major drawback of the technology lies in its lack of encryption and authentication, which is discussed in the following section.

B. False Data Injection Attacks

Extensive research can be found in the literature discussing the cyber security of surveillance communications [18], [24], [25], [28]. The progressive shift from independent and non-cooperative technologies (PSR) to dependent and cooperative technologies (ADS-B) has created a strong reliance on external entities (aircraft, GNSS) to estimate aircraft state. This reliance, along with the introduction of air-to-ground data links via Modes A/C/S and the broadcast nature of ADS-B, has brought alarming cyber security issues.

FDIAs were initially introduced in the domain of wireless sensor networks [12]. A wireless sensor network is composed of a set of nodes (i.e. sensors) that send data report to one or several ground stations. Ground stations process the reports to reach a consensus about the current state of the monitored system. A typical scenario consists of an attacker who first penetrates the sensor network, usually by compromising one or several nodes, and thereafter injects false data reports to be delivered to the base stations. This can lead to the production of false alarms, the waste of valuable network resources, or even physical damage. Active research regarding FDIAs has been conducted in the power sector, mainly against smart grid state estimators [7], [11]. It shows that these attacks

[4]https://flywithscout.com

may lead to power blackouts but can also disrupt electricity markets [26], despite several integrity checks.

FDIAs also exist in the domain of air traffic surveillance. Because surveillance relies on the information provided by aircraft's transponders to ground stations, aircraft transponders are equivalent to nodes from a wireless network, and ground stations are equivalent to base stations. Although in the ATC domain, there is no real effort to penetrate the sensor network, as all communications are unauthenticated and in clear text. Still, performing FDIAs on surveillance communications is no simple task: it requires a deep understanding of the system, its protocol(s) and its logic, to covertly alter (by injecting falsified squitters and deleting genuine ones) the consensus reached by the ground station regarding the air situation picture. These attacks are much more complex to achieve than e.g., jamming, because the logic of the communication flow must be preserved and the falsified data must go unnoticed.

The means of the attacker to conduct FDIAs against ADS-B communications have already been detailed in previous work [14], [23]. Considering the attacker has the necessary equipment, he can perform three malicious basic operations:

(i) *Message injection* which consists of emitting non-legitimate but well-formed ADS-B messages.
(ii) *Message deletion* which consists of physically deleting targeted legitimate messages using destructive or constructive interference. It should be noted that message deletion may not be mistaken for jamming, as jamming blocks all communications whereas message deletion drops selected messages only.
(iii) *Message modification* which consists of modifying targeted legitimate messages using overshadowing, bit-flipping or combinations of message deletion and message injection.

The above three techniques allow for the execution of several attack scenarios [18] that can be categorized in a taxonomy:

- **Ghost Aircraft Injection.** The goal is to create a non-existing aircraft by broadcasting fake ADS-B messages on the communication channel.
- **Ghost Aircraft Flooding.** This attack is similar to the first one but consists of injecting multiple aircraft simultaneously with the goal of saturating the air situation picture and thus generates a denial of service of the controller's surveillance system.
- **Virtual Trajectory Modification.** Using either message modification or a combination of message injection and deletion, the goal of this attack is to modify the trajectory of an aircraft.
- **False Alarm Attack.** Based on the same techniques as the previous attack, the goal is to modify the messages of an aircraft in order to indicate a fake alarm. A typical example would be modifying the squawk code to 7500, indicating the aircraft has been hijacked.
- **Aircraft Disappearance.** Deleting all messages emitted by an aircraft can lead to the failure of collision avoidance systems and ground sensors confusion. It could also force the aircraft under attack to land for safety check.
- **Aircraft Spoofing.** This scenario consists of spoofing the ICAO number of an aircraft through message deletion and injection. This could allow an enemy aircraft to pass for a friendly one and reduce causes for alarm when picked up by PSR.

One can sense the potential for disaster if one of these attack scenarios was to be executed successfully. It is of the utmost importance that none of the scenarios represent a real threat to such a critical infrastructure with human lives on the line. However, because of the inherent properties of the ADS-B protocol, current solutions for securing ADS-B communications are only partial or involve an unbearable cost [24]. Therefore, ATC systems must become robust against FDIAs, i.e. being capable of automatically detecting any tempering with the surveillance communication flow while being able to maintain the infrastructure in a working state.

C. FDIA detection using Machine Learning

Detecting FDIAs among regular data is critical to avoid the aforementioned risks. Machine Learning techniques have contributed significantly to improving the detection of anomalies in many domains. Detecting abnormal cells or markers on medical imagery [17] [19], using logging information to detect intruders on a network [27], or detecting traffic inconsistencies in bus trajectory data [10] are few of many applications of irregularity disclosure using Machine Learning techniques.

Regarding anomaly detection in ADS-B data, several experiments are found in the literature. Using signal discrepancies through different receiver sensors [13] show the ease of access to physical data through personal antennas. On the other hand, platforms like Opensky-Network [20] helped the training of models using the logical aspects of ADS-B [21] [9].

Although, a major drawback of using Machine Learning techniques for anomaly detection is the lack of pre-existing altered data, whether to use it for training or testing. While contributions exist using unsupervised models on unbalanced dataset with good detection scores [16] [5], supervised learning will often yield underperforming results. Hence, there is a need of harmonizing dataset to explore new models and improve existing ones. Alteration mechanisms as proposed in this paper would enable such goals on ADS-B data.

D. Research Objective and Questions

It is critical to make sure that FDIA detection systems are properly and thoroughly tested. Such a testing campaign needs large sets of test data in the form of air traffic recordings, where ideally half of the recordings present an anomaly, i.e. an FDIA that should be detected. The creation of FDIA test scenarios in the ATC domain can be very complex. This requires at the same time altering or creating false data (e.g., creating ghost aircraft) while ensuring the consistency of all data. It is also necessary to ensure the widest possible coverage w.r.t. the taxonomy of attacks. The creation of a set of algorithms capable of generating FDIA test scenarios thus

aims to considerably increase the feasibility of in-depth testing of the efficiency and effectiveness of FDIA detection systems. We have defined the following research objective based on these observations:

Create a set of alteration algorithms to automate the production of synthetic FDIA scenarios for the testing of supervised learning-based detection models.

From the above research objective, we have identified 2 research questions that are expressed and developed below. We refer to these questions in Section V, during the evaluation of the approach.

RQ1 **To what extent is it possible to automate the generation of FDIA scenarios that cover the ADS-B taxonomy of attack?**

Although they are all based on the same weakness – i.e. the injection of false data – each scenario of the taxonomy has its subtleties. For instance, it should be possible to easily alter aircraft's trajectory while ensuring realism of the modified trajectory throughout time. In turn, if realism (including computation of each latitude/longitude coordinates) is abstracted from the user in exchange of design primitives (e.g., by defining way-points and time of passage), then realism and the computation of a new trajectory shall be done algorithmically. Moreover, algorithms must integrate what makes the protocol and its domain of application specific.

RQ2 **To what extent is the approach relevant to generate synthetic data for testing an AI model detecting abnormalities in ADS-B data?**

In other cyber security domains (typically, in IT), huge historical databases of past attacks and demonstration of vulnerability exploitation are available on the internet (e.g., the National Vulnerability Database of the NIST[5]). It is usually straightforward to reproduce these attacks or simulate them on a system under test, especially when there also exist databases of exploits (e.g., exploit-db[6]). In the ATC domain however, there is no history of False Data Injection Attacks. At least none that was reported to the general public, and therefore the only option to obtain test data is to generate synthetic anomalies. Our approach aims to generate complex FDIA scenarios that could contribute to more efficient ML-based detection systems. We experimented the proposed approach by generating test data sets and submitted them to Habler et al.' auto-encoder model [9], results are presented in Section V.

III. Alteration Process Overview

The current approach aims to generate altered versions of an original recording in order to feed FDIA detection systems that shall be confronted to abnormal situations. A recording is a set of messages emitted by real aircraft and stored in a file through an antenna or created from data collected with online providers such as OpenSky Network[7].

[5]https://nvd.nist.gov/

[6]https://www.exploit-db.com/

[7]https://opensky-network.org/

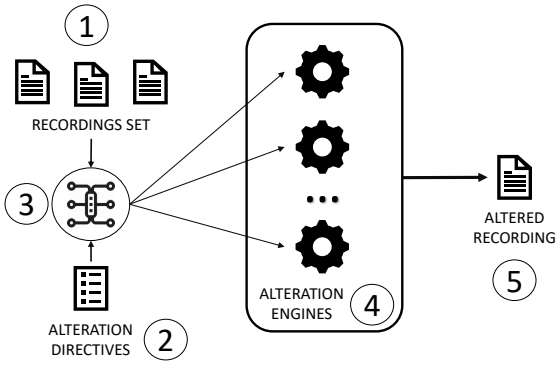

Fig. 1: Alteration process overview

Figure 1 depicts the alteration process. An altered recording is produced as output from a set of original recordings and a set of alteration directives. There are five items in this process that shall be detailed:

① **A set of original recordings** that contains recordings that can be altered or used as a specific parameter of an alteration directive (e.g., a replay attack needs one recording to perform the alteration on, and one (source) recording to inject in the first one. The source recording is specified in an alteration directive).

② **A list of alteration directives** that constitutes the specification of the test scenario. Indeed, the alteration process relies on a list of alteration directives to be performed on the original recording. An alteration directive, depending on the type of alteration, has several parameters, such as a time window (*when* the alteration takes place in the recording), a list of targeted aircraft, a list of way-points in case of virtual trajectory modification for instance, etc.

③ **The switcher** iterates the list of alteration directive, and for each one calls the corresponding alteration engine based on the type of the alteration specified in the current alteration directive.

④ **The alteration engines** are used to generate an altered recording from an original one. There is an alteration engine for each type of alteration, the engines are following the same structure of methods, as further explained in Section IV. All the algorithms used by these engines are discussed

⑤ **An altered recording** is outputted from the process. Its closeness to the original recording depends on the number of targeted aircraft and on the scope defined in the alteration directives.

A test scenario is formalized as $S = (r, D)$ with r as the targeted recording picked among the set of recordings, and D as a list of alteration directives.

Regarding the format of messages taken as input, our prototype accepts surveillance data in the SBS format[8]. SBS messages contain twenty two fields. The ten first fields contain the reception time stamp and static properties about the emitter

[8]http://woodair.net/sbs/Article/Barebones42_Socket_Data.htm

aircraft or the flight e.g., its ICAO 24-bit address[9], its aircraft ID[10].

The last twelve fields contain dynamic properties of the emitter aircraft, i.e. properties that evolve over time such as altitude and ground speed. There are eight types of message, depending on the types of property that are communicated. For instance, type 3 messages contains position information: altitude, latitude, longitude, while type 4 messages contain velocity information: ground speed, vertical rate, and heading. It should be noted that, for the alteration engines' description, we ignore the fact that there are various types of message, and consider that messages contain all information. This considerably reduces the complexity of the algorithms without loosing essential information.

IV. Alteration Engines

The section present all type of alteration directives and their corresponding algorithms. The figure 2 depicts step by step the workflow shared between all algorithms.

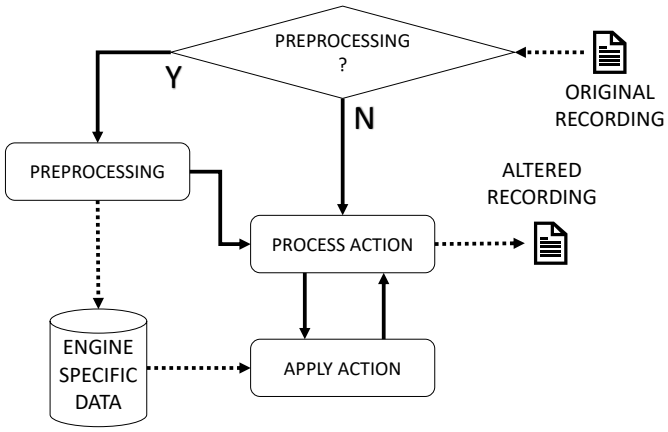

Fig. 2: Alteration process of the engines

The first step consists of determining if a preliminary analyse is required by the current algorithm. Indeed, some algorithms must extract information from the recording that will be used later to produce altered messages. For example, with the trajectory modification, the dedicated engine needs to generate interpolation functions from the latitude, longitude and altitude of the targeted aircraft. Then these functions will be queried to produce altered values latitude, longitude and altitude and therefore alter the initial trajectory of the aircraft.

The workflow is mainly based on the *process action* and *apply action* steps. The *process action* method iterates the recording, and checks if specific conditions are verified (e.g., the message is in the scope of the alteration and if the emitter aircraft is targeted by the alteration). If it is the case, the message is sent to the *apply action* method that applies the supplied alteration directive to the message. Then the resulting message, if any, is returned to the *process action* step and written into the altered recording.

Each alteration engine implements a specific version of the pre-analyse, *process action* and *apply action* methods. We detail each of them below.

A. Property modification

Although most alterations consist in some way of changing the properties of received surveillance messages, this alteration allows users to be quite precise in the message properties they want to modify. It makes it possible to perform *False Alarm* attacks and *Aircraft spoofing*, as both consists of changing the value of a single property: squawk code and ICAO address, respectively.

Property modification is represented as an alteration directive $dir_{prop} = (s, t, P)$ where s is the duration of the attack defined as a time interval relative to the recording, t is a list of targeted aircraft, and P is a non-empty set of property value changes. A property value change is represented as a triplet $p = (i, v, o)$ where i is the property identifier (e.g., altitude or ground speed), v v is a value , and o specifies how v shall be employed to modify the property's initial value. $p.o$ can be of for types: REPLACE, OFFSET, NOISE, and DRIFT. If $p.o = REPLACE$, then v replaces the property's initial value. If $p.o = OFFSET$, then v is added to the property's initial value. If $p.o = NOISE$, then a random value ranging $0 - v$ is added the property's initial value. Finally, if $p.o = DRIFT$, then v plus the sum of the previous drift is added to v.

The *process action* method for the Property modification alteration is described in Algorithm 1. It iterates the messages of the initial recording and for each message, if it was sent by a targeted aircraft within the alteration time frame (line 2), then the message as well as the alteration directive are sent over to the *apply action* method, which returns the altered version of the message (line 3). Otherwise, the initial message is preserved as is (line 4). Finally, the obtained message (altered or preserved) is added to the resulting recording al_rec (line 7), and once all messages have been processed, the resulting recording is outputted (line 9).

The algorithm that implements the *apply action* method for the Property Modification algorithm is presented in Algorithm 2. It iterates the list of property value changes (line 2), and for each property, it checks in line 3 if the directive's property value change is an offset. If it is an offset type of alteration, then the message resulting property value is the addition of the initial value of the property and the property value contained in the directive (line 4). Otherwise, the message initial value is replaced by the directive's property value (line 6). Note that the dp variable represents the drift progress and, as such, it is only used if $dir.p.o = DRIFT$.

B. Aircraft Disappearance

The objective of this alteration is to hide aircraft i.e. delete messages of certain aircraft at a certain time. It is represented as an alteration directive $dir_{del} = (s, t, n)$ where s is the

[9]https://www.iomaircraftregistry.com/flight-operations/flight-operations/icao-24-bit-aircraft-address-mode-s-coding/

[10]https://www.faa.gov/licenses_certificates/aircraft_certification/aircraft_registry/releasable_aircraft_download/

Algorithm 1: *process action* method for Property Modification

Input: *rec* \ list of genuine surveillance messages
dir \ alteration directive
Result: al_rec \ list of genuine and altered messages

$dp \leftarrow 1$
foreach $old\_msg \in rec$ **do**
  if $old\_msg.ts \in dir.s \wedge old\_msg.icao \in dir.t$ **then**
    $new\_msg \leftarrow applyAction(old\_msg, dir, dp)$
    $dp \leftarrow dp + 1$
  else
    $new\_msg \leftarrow old\_msg$
  end
  $al\_rec \leftarrow al\_rec \cup \{new\_msg\}$
end
return al_rec

Algorithm 2: *apply action* Method for Property Modification

Input: *msg* \ genuine surveillance message
dir \ alteration directive
dp \ drift progress
Result: al_msg \ altered message

$al\_msg \leftarrow copy(msg)$
foreach $p \in dir.P$ **do**
  if $p.o = OFFSET$ **then**
    $al\_msg.getParam(p.i).value \leftarrow al\_msg.getParam(p.i).value + p.v$
  else if $p.o = REPLACE$ **then**
    $al\_msg.getParam(p.i).value \leftarrow p.v$
  else if $p.o = NOISE$ **then**
    $al\_msg.getParam(p.i).value \leftarrow al\_msg.getParam(p.i).value + rand(0, p.v)$
  else if $p.o = DRIFT$ **then**
    $al\_msg.getParam(p.i).value \leftarrow al\_msg.getParam(p.i).value + p.v \times dp$
  end
end
return al_msg

duration of the attack defined as a time interval relative to the recording, t is the targeted aircraft and n is the number of deleted consecutive messages e.g. if $n = 0$ then all messages are deleted, while if $n = 3$ for instance, then only three out of four messages are deleted. This is certainly the simplest form of alteration, as it solely about dismissing messages originating from certain aircraft, during a certain time frame. The algorithm for this alteration lies in the *process action* method, where messages are written in the resulting altered recording only if the sending aircraft is not targeted and the message was not sent within the specified time frame, i.e. $\neg(msg.icao \in dir.t \wedge msg.ts \in dir.s)$.

C. Virtual Trajectory Modification

Altering the trajectory of aircraft is a much more complex problem than simply modifying property values at a certain time. This should be done realistically, i.e. with regards to aircraft physical characteristics, as it would be easily detected otherwise. Dynamic properties of aircraft can be formalized as continuous functions of property values related to time, which mimic real-life aircraft physical behaviour as closely as possible. A good candidate for this is interpolation as it "fills the gap" between each pair of consecutive values. We opted for the Akima interpolation [2] as local interpolation technique as local interpolation is not subjected to Runge's phenomenon, i.e. a problem of oscillation at the edges of an interval that occurs over a set of equally spaced interpolation points [8] (as opposed to global interpolation). Another benefit of this technique is, because it only uses the neighbouring points for its calculation, a faster computation of approximation functions. Because recordings can be substantial in volume (a 30min recording from one sensor may contain around 150000 squitters), a fast interpolation method certainly contributes to the approach's scalability. It should be noted that interpolation is a form of approximation and as such there is a certain share of uncertainty in the calculated property values, i.e. interpolation divergence. While this level of uncertainty would not be acceptable to gain sufficient confidence in the efficiency of FDIA detection systems critical applications, this is certainly a good fit to demonstrate the capabilities of our approach to rely on a third party "aircraft simulation module" to generate realistic aircraft trajectories.

Virtual Trajectory Modification is represented as an alteration directive $dir_{vtm} = (s, t, \Omega)$, where Ω is a nonempty set of way-points. A way-point is defined as $\omega = (lat, lon, alt, ts)$, i.e. 3D coordinates and a time of passage.

The algorithm for trajectory modification is two-fold. First, as a pre-analyse step, it must populate interpolation functions that account for the whole aircraft trajectory, taking into account the part of the trajectory that should be removed and replaced by the supplied way-points. Second, during the *process action* step, it iterates the recording, and for each targeted message, it replaces its property values with the one obtained by querying the interpolation functions. Note that this implementation does not create additional messages, it only edits existing messages. Therefore, if the altered trajectory is longer that the initial one (i.e. more distance is traveled) then the aircraft's ground speed must be augmented accordingly since travel time $dir.t$ is not adjustable (as it is bound to the messages). This can lead to erroneous situations where aircraft fly at an impossible speed (either too high or too low) to travel the altered distance within the fixed time $dir.t$.

Population of the interpolation function, as part of the pre-analyse step, is presented in Algorithm 3. We consider a data structure called *Traj* that contains three interpolation functions (for latitude, longitude and altitude), and an identifier (an ICAO). The algorithm iterates the recording, and for each message emitted by targeted aircraft, if it was not sent

Algorithm 3: Pre-analyse method for Trajectory Creation

Input: *recording* \ list of genuine surveillance messages
dir \ alteration directive
Result: trajs \ list of aircraft trajectories

```
1  updated ← zeros(dir.t.size)
2  foreach msg ∈ recording do
3      if msg.icao ∈ dir.t then
4          if ¬(msg.ts ∈ dir.s) then
5              trajs(msg.icao)
6                  .addPos(msg.lat, msg.lon, msg.alt, msg.ts)
7          else if updated(msg.icao) = 0 then
8              foreach ω ∈ dir.Ω do
9                  trajs(msg.icao)
10                     .addPos(ω.lat, ω.lon, ω.alt, ω.ts)
11             end
12             updated(msg.icao) ← 1
13         end
14
15     end
16 end
17 return trajs
```

within $dir.s$, the received position is added to a *Traj* instance associated to $dir.t$ (lines 4–6). Internally, each value is added to their corresponding interpolation functions. When the first message (of a certain icao) that was sent within $dir.s$ is iterated (line 7), then the algorithm iterates $dir.\Omega$ and adds each way point ω to the *Traj* instance (lines 9–10). Finally, the trajectory is marked as altered (line 12), and all subsequent messages within $dir.s$ are ignored. In other words, all positions of a given trajectory are added to a Traj instance, except the positions that are within the alteration time window $dir.s$, which are replaced by way-points supplied in $dir.\omega$. The result is a list of *Traj* instances, one for each targeted aircraft.

Once all targeted aircraft are associated with a *Traj* instance, the *process action* method performs the alteration. It iterates the recording, and for each message from targeted aircraft sent within $dir.s$, it calls the *apply action* method. The latter replaces property values of the supplied message according to dir, as follows:

- For latitude, longitude, and altitude properties, the initial values are replaced with interpolated values using the *Traj* instance.
- For ground speed, track, and vertical rate properties, new values are computed based on the interpolated position values. To compute ground speed of a certain message m for instance, the algorithm first computes the horizontal distance (i.e. ignoring altitude) between two near points of the trajectory, one taken 5 seconds before m and another one 5 seconds after m. The obtained distance, divided by the time it took to travel it (10 seconds), gives the ground speed.

D. Ghost Aircraft Creation

In this attack, the attacker creates a fake track from scratch, implying that in this case, fake messages must be created and inserted into the target recording. Ghost Aircraft Creation is represented as an alteration directive $dir_{gac} = (s, t, P, \Omega)$ where Ω is a nonempty set of way-points. Its implementation relies on the trajectory modification engine. First, a pre-analyse method is called to create *Traj* instances, where the interpolation functions are solely populated by the directive's way-points $dir.\Omega$, to be sent to the *process action* method, shown in Algorithm 4. Knowing that ADS-B messages are sent every 400 to 600ms, the *process action* method generates empty messages, starting at $dir.s.start$ (line 1), randomly increasing their time of sending accordingly (line 6), and until it has reached the end of the alteration time window $dir.s.end$ (line 2). Then, to populate each message, the method calls the *apply action* method of the Virtual Trajectory modification engine, using the list of *Traj* instances.

Algorithm 4: *apply action* Method for Ghost Aircraft Creation

Input: *rec* \ list of genuine surveillance messages
dir \ alteration directive
Result: al_rec \ resulting recording

```
1  mts ← dir.s.start
2  while mts <= dir.s.end do
3      msg ← newMsg(dir.P)
4      msg.ts ← mts
5      msg ← applyAction(msg)
6      mts ← mts + rand(0.4, 0.6)
7  end
8  return al_rec
```

E. Ghost Aircraft Flooding

The initial definition of this attack consists of suddenly creating a lot of ghost aircraft, thus supposedly saturating the Recognized Aircraft Picture (RAP – i.e. what the controller sees). However, this has proven to be quite straightforward for detection systems to recover from this type denial of service. We propose instead to slightly modify the definition of the attack, to be virtual trajectory modification flooding. The goal is to suddenly generate many different trajectories for a targeted aircraft, as if the aircraft was being split in multiple pieces, thus saturating the detection systems with many conflicting messages.

Aircraft flooding is represented as an alteration directive $dir_{gaf} = (s, t, z, k, \alpha, v, d)$ where:

- s is a time window, where $s.end$ is necessarily the end of the recording.
- z is the number of fake trajectories to create;
- k is the maximum offset value in terms of distance between the fake trajectories and the original trajectory;

- α is a probability distribution ranging 0.1 – 1 that determines, for each fake trajectory, how close to k its divergence will be;
- v is the speed of divergence, i.e. how fast the fake trajectories are offset from the original trajectory. it is expressed as a duration and specifies how much time it should take for the fake trajectory to reach its divergence bound, i.e. $k \times \alpha$;
- d is the direction of divergence of a given fake trajectory; it is a probability distribution that takes its value from a horizontal 90deg cone originating from the aircraft nose and centered on the aircraft track line.

Each fake trajectory is thus created by slowly drifting horizontally from the original trajectory toward direction $dir.d$, up to $dir.k * \alpha$, at speed $dir.v$. Once the maximum offset has been reached, then the trajectory mimics the original, while preserving its horizontal position offset.

Algorithm 5: Trajectory creation Algorithm for flooding

Input: *initraj* \ original *Traj* instance (up to $dir.s.start$)
dir \ alteration directive
Result: trajs \ set of rogue trajectories

1 **for** $i \leftarrow 1$ ***to*** $dir.z$ **do**
2 $\alpha_i \leftarrow sample(dir.\alpha)$; $ts \leftarrow dir.s.start$
3 $d_i \leftarrow sample(dir.d)$; $traj_i \leftarrow initraj$
4 **while** $ts \leq dir.s.end$ **do**
5 $pCoef \leftarrow min((ts - dir.s.start)/dir.v), 1)$
6 $step \leftarrow pCoef * \alpha_i * dir.k$
7 $lat \leftarrow fulltraj.interp(\text{"}lat\text{"}, ts) + step +$
8 $cos(fake\_ac.d)$
9 $lon \leftarrow fulltraj.interp(\text{"}lon\text{"}, ts) + step +$
10 $sin(fake\_ac.d)$
11 $traj_i.addPos(lat, lon,$
12 $fulltraj.interp(\text{"}alt\text{"}, ts), msg.ts)$
13 $ts \leftarrow ts + dir.$
14 **end**
15 $trajs \leftarrow trajs \cup \{traj_i\}$
16 **end**
17 **return** $trajs$

A fake trajectory is represented as a *Traj* instance, built during the pre-analyse phase, and shown in Algorithm 5. We first assume that a *Traj* instance *initraj* has been created, containing all positions of the targeted aircraft (i.e., one position per message) up until $dir.s.start$. Then, for each fake trajectory, the method samples $dir.\alpha$ to determine the divergence magnitude of the trajectory (line 2), and creates the trajectory by duplicating *initraj* (line 3).Next, it creates at regular intervals the way-points that draw the fake trajectory, starting from $dir.s.start$ until the end of the recording. At each interval ts, the algorithm computes the divergence progression percentage $pCoef$ (line 5), which consists of dividing the time distance from the beginning of the alteration to the current interval, by the specified duration of the progression $dir.v$. If $dir.v$ is greater than the current time distance, then it means the trajectory is not done diverging from the initial one, and $pCoef$ is under 100%. If $dir.v$ is lower, then divergence has reached its maximum and $pCoef$ is at 100%. Then the euclidean distance from the initial position $step$ is computed by multiplying $pCoef$ to the sampled value of α and the maximum offset k (line 6). Finally, the actual latitude and longitude values are obtained using $step$, trigonometric functions *sin* and *cos*, and the sampled direction of divergence k_i (lines 7–8). The algorithm eventually returns all fake trajectories $trajs$.

The last step is to convert the fake trajectories into ADS-B messages to be injected to the initial recording. For each message sent by the targeted aircraft within $dir.s$, and for each fake trajectory, the message is duplicated, and its dynamic properties are replaced by values obtained from sampling the interpolation functions of the corresponding *Traj* instance, in a similar fashion as the *process action* of Virtual Trajectory Modification.

F. Replay

Although this type of attack is not part of the taxonomy, we take it from recent discussions with experts that replay attacks represent a very serious threat as it abstracts itself from realism issues. A typical example of such an attack would be terrorists who collected ADS-B messages of a regular flight on a certain day, then a few days later, hop in on the plane, hijack it and physically temper with the ADS-B transponders to make them send out the messages they previously recorded. This could allow terrorists to change course of flight without being noticed immediately. In this scenario, there is no need to compute realistic altered values to dupe detection mechanisms, since the fake ADS-B messages that are emitted come from a genuine source. Therefore, replay attacks can certainly bypass FDIA detection systems that rely on realism analysis.

Replay is represented as an alteration directive $dir_{replay} = (s, t, r)$ where r is a source recording from which to extract the targeted aircraft in t.

The first step is to extract messages from the source recording. This is done as a first pre-analyse phase by iterating the recording's messages and checking whether the emitting aircraft is present in $dir.t$ and if the current message was sent within the directive's time frame. If that is the case, then the message is marked to be replayed. The list of all the messages that were marked is eventually returned.

The second step is about adjusting the timestamp of the extracted messages so that they can be correctly inserted in the target recording. For each message msg to be replayed, the algorithm assigns a new time stamp that relates to the target recording, as follows:

$msg.ts \leftarrow first\_ts + msg.ts - first\_rep\_ts + offset$

where $first\_ts$ is the time stamp of the first message of the target recording, $first\_reo\_ts$ is the time stamp of the first message extracted for replay, and $offset$ is a value in second expressing where in the target recording the previously extracted messages should be inserted.

Algorithm 6: *process action* Method for Replay attack

Input: reps \ list of messages picked up for replay
rec \ source recording
Result: al_rec \ resulting recording

```
1  cur ← rec.first
2  rp ← reps.first
3  while cur ≠ ∅ do
4      while reps ≠ ∅ ∧ cur.ts > rp.ts do
5          al_rec ← al_rec ∪ {rp}
6          reps ← reps − {rp}
7          rp ← reps.next
8      end
9      al_rec ← al_rec ∪ {cur}
10     cur ← rec.next
11 end
12 while reps ≠ ∅ do
13     al_rec ← al_rec ∪ {rp}
14     reps ← reps − {rp}
15 end
16 return al_rec
```

The last step is the actual merging between the target recording and the list of extracted messages, as perform by the *process action* method detailed in Algorithm 6. The method iterates the target recording's messages (lines 3–11), and if the timestamp of the first of the extracted messages cur is smaller than the timestamp of the current message rp, then rp is put in the resulting recording al_rec (line 5) while being removed from the list of extracted messages (line 6). Otherwise, cur is added to the resulting recording. If the target recording has been iterated and the list of extracted messages is not empty, the remaining extracted messages are appended to al_rec (lines 12–14).

Looking back at the first research question defined in Section II-D:

RQ1 **To what extent is it possible to automate the generation of FDIA scenarios that cover the taxonomy of attack?**

We believe this question can be answered positively, as we were able to provide a dedicated algorithm for each type of attack from the taxonomy. Moreover, we introduced a new type of attack, *Replay*, and we improved on the definition of an existing attack, *Ghost Aircraft Flooding*, to become *Trajectory Modification Flooding* as it is an FDIA much more complicated to detect.

V. Experimentation

ATC constitutes one of the most critical infrastructures on the planet and as such, it is not possible to demonstrate the use of FDI-T in real situation by feeding falsified data to a deployed FDIA detection system and reporting results in a publicly available paper.

Therefore, we opted instead to confront our tool to a detection approach from the literature [9]. First, we elaborate on the constituents of the model and their characteristics. Then we present the dataset that were used to train and test the model, and especially how the proposed approach can allow to generate the anomalies presented in [9] as well as a more complex anomaly of our own making. Finally we discuss the experimentation results and propose an answer to the two research questions

A. The model

The authors propose a Machine Learning method to discriminate malicious messages from regular ones. They use an LSTM-based encoder-decoder, a deep learning architecture. The model is decomposed in two different sub-models:

① **The encoder** which learn to create a representation of its input. In this experimentation, it is based on LSTM cells that have been proven relevant in time series analysis. [cite] It cannot learn alone as it has no way to check if its output are correct by itself. That is the reason why it is always coupled with a decoder for its training.

② **The decoder** will take the encoder's output, namely the vectorized representation of the original data and will decode it to fit the expected output. In the case of Habler et. al' model, the output is actually the input. That is why, this kind of encoder-decoder is called an auto-encoder. The decoder will have a similar architecture, if not identical, to the one used in the encoder.

In this experiment, several complete, isolated flights (from take off to landing) are taken and sliced in windows of 15 messages with a stride of 1. The encoder with these windows and checked the cosine similarity between the input and the output, which are supposed to be as similar as possible. Once the model has been trained with normal data, an anomaly threshold is set based on the reconstruction score of our architecture. If trained properly, the model will have trouble to auto-encode malicious data as they differ from regular ones and the threshold set previously will be exceeded, resulting on a detected anomaly.

B. Dataset

To train the data and test it, we used several flights collected from the OpenSky database [20] using our fork from the Traffic library [citation et footnote] to use raw messages. These were all European flights as it was easier to get full flight thanks to the ADS-B cover Europe benefits. We tried to get as many flights morphologies as possible, e.g., Madrid-Moscow which comes with significantly different routes depending on the weather due to its long journey.

After training the model on legitimate data, we created with our approach alteration scenarios based on the anomalies defined in Habler's paper used to validate their approach. For each anomaly, we present below the corresponding alteration scenario and its associated directives using the notation given in Section III, completed with the following notation:

- R_{origin}: a recording containing one flight from one European city to another.

- R_{source}: a recording containing one flight from one European city to another. Where $R_{source} \neq R_{origin}$
- a_1: the aircraft contained into R_{origin}.
- a_2: the aircraft contained into R_{source}.
- t_n: a time duration in seconds that cannot exceed R_{origin}'s duration, i.e. $t_n < t_{n+1}$.
- **Random noise (RND)** - anomalies are generated by adding random noise to messages. To do so, values from original messages are multiplied with a random number between zero and two. RND requires a scenario using a *Property Modification*. According to the Section IV-A it is possible to formalize the scenario S_{RND} as following:

S_{RND}	$=$	$(R_{origin}, \{dir_1\})$
With		
dir_1	instance of	dir_{prop}
$dir_1.s$	$=$	$[t_0, t_1]$
$dir_1.t$	$=$	$\{a_1\}$
$dir_1.P$	$=$	$\{p_1\}$
p_1	$=$	$(ALTITUDE, 2, NOISE)$

- **Different Route (Route)** - anomalies replace a whole segment of the ADS-B messages with a different route taken from OpenSky. Route requires a scenario that combines a *Replay* and an *Aircraft Disappearance* directive. According to Sections IV-F and IV-B it is possible to formalize the scenario S_{Route} as following:

S_{Route}	$=$	$(R_{origin}, \{dir_2, dir_3\})$
With		
dir_2	instance of	dir_{replay}
$dir_2.s$	$=$	$[t_0, t_1]$
$dir_2.t$	$=$	$\{a_2\}$
$dir_2.r$	$=$	R_{source}
dir_3	instance of	dir_{del}
$dir_3.s$	$=$	$[t_0, t_1]$
$dir_3.t$	$=$	$\{a_1\}$
$dir_3.n$	$=$	1

- **Gradual Drift (DRIFT)** - anomalies simulate an altitude drift. The altitude messages on the attacked time window are all raised/lowered by an increasing/lowering multiple of n feet. So if the first message is lowered by a hundred feet, the second will be lowered by two hundreds, etc. DRIFT requires a scenario using a *Property Modification*. It is possible to formalize the scenario S_{DRIFT} as following:

S_{DRIFT}	$=$	$(R_{origin}, \{dir_4\})$
With		
dir_4	instance of	dir_{prop}
$dir_4.s$	$=$	$[t_0, t_1]$
$dir_4.t$	$=$	$\{a_1\}$
$dir_4.P$	$=$	$\{p_2\}$
p_2	$=$	$(ALTITUDE, 25, DRIFT)$

- **Velocity Drift (VEL)** - anomalies are gradual drift applied to the velocity feature in knot. VEL requires a scenario using a Property Modification. It is possible to formalize the scenario S_{VEL} as following:

S_{VEL}	$=$	$(R_{origin}, \{dir_5\})$
With		
dir_5	instance of	dir_{prop}
$dir_5.s$	$=$	$[t_0, t_1]$
$dir_5.t$	$=$	$\{a_1\}$
$dir_5.P$	$=$	$\{p_3\}$
p_3	$=$	$(SPEED, 1.0, DRIFT)$

- **Message blocking (BLK)** - only the first message out of every five consecutive messages is preserved, to simulate a denial-of-service attack or a network delay mode. BLK requires a scenario using an *Aircraft Disappearance* directive. It is possible to formalize the scenario S_{BLK} as following:

S_{BLK}	$=$	$(R_{origin}, \{dir_6\})$
With		
dir_6	instance of	dir_{del}
$dir_6.s$	$=$	$[t_0, t1]$
$dir_6.t$	$=$	$\{a_1\}$
$dir_6.n$	$=$	4

It was possible to generate all the anomalies as presented in the publication associated to the detection model. In addition, to demonstrate the added value of the presented approach, we introduce an anomaly that was not part of the model's validation data set. It consists of a trajectory modification **(TRJ)** of a hundred kilometers throughout a long period of time e.g., at least for 60% of the flight.

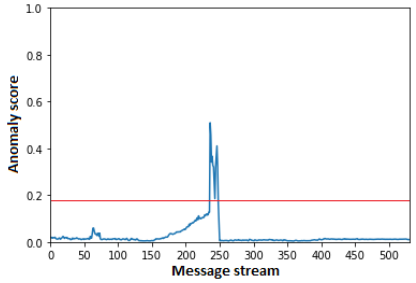

Fig. 3: Example of a drift down anomaly score with a false data injection. The line refers to the threshold above which an alert is raised.

C. Results

Out of the five anomalies we recreated, we obtained similar results to Habler's on four of them. Figure 3 is a graphical visualization of the drift down anomaly on a flight from Moscow to Madrid. We can clearly see the pike between the message 180 and 250 going above the anomaly threshold, defined in a similar fashion as the reference experimentation. Similar

results can be observed on the 3 other detected anomalies. However, we did not manage to obtain a comparable pike on the **BLK** anomaly, thus not triggering any anomaly flag. One assumption for this result would be the difference between our data and the ones used in the original experiment. For instance, results can differ depending on the time gap between messages. If this gap is bigger, then deleting 4 messages out of 5 means that the aircraft goes dark for a longer period of time, easing the detection. Overall, the false data injections we recreated using FDI-T seems to be detected the same way as handcrafted ones.

As for our generated anomaly **TRJ** on a flight between Oslo and Paris, results are presented on Figure 4. We find that the anomaly score goes way below the detection threshold set for the other five anomalies. It results to an unnoticed attack. However, we can also observe that the use of FDI-T to modify trajectories has the side effect of smoothing the curve, hence making it detectable for a human eye. One workaround would be to add a certain amount of noise to simulate more realistic data.

We show with **TRJ** that our tool is capable to create alterations that would be too cumbersome to create by hand as we are close to 5000 messages modified from the original. In addition, these crafted attacks are smart enough to not be detected by a model taken from the litterature, which exposes its limits on real contexts.

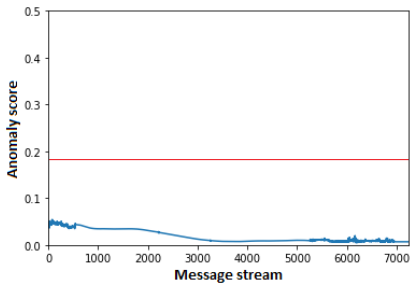

Fig. 4: The anomaly score for **TRJ** is way below than the drift down anomaly. One can observe the curve being smoothed out by the alteration

Base on the experimentation results, it is possible to provide an answer to the second research question:

RQ2 **To what extent is the approach relevant to generate synthetic data for AI Testing?**

Not only it was possible to generate the anomalies that were used to evaluate Habler et. al.'s model, but the presented approach made it possible to generate a more complex anomaly that was not picked up by the model, hence demonstrating that it is suitable for providing better synthetic test data sets for ML based detection systems and therefore contributing to their overall improvement.

VI. Conclusion

In the context of improving the cybersecurity in air surveillance communications, and especially with regards to the ADS-B protocol, this paper describes a novel test data generation approach to be part of an existing FDIA testing framework dedicated to ATC systems, called FDI-T. More precisely, the contribution concerns the various generation algorithms that were developed, referred to as called alteration engines, of which the specificities are extensively presented in this paper. The two objectives of the proposed approach is, first, to cover the taxonomy of ADS-B attack scenarios that has been described in the literature on multiple occasions, and second, to use the approach to create synthetic test data sets for the evaluation of machine learning based detection systems. We show how both objectives have been reached by successfully designing all classes of scenarios from the taxonomy using the various algorithms, and by using the approach to reproduce the test data for a detection system from the literature as well as generating new test data that helped unveil previously unknown weaknesses of the model. Among future work, we have started experimenting with our approach in the context of maritime surveillance, which is based on a communication protocol - AIS - similar to the ADS-B protocol. Among the ongoing work, we have started to experiment with our approach in the context of maritime surveillance, which is based on a communication protocol - AIS - similar to the ADS-B protocol.

VII. Acknowledgement

This work is part of an ongoing research initiative toward the generation of FDIA test scenarios partially carried out in partnership with the french company Kereval. This work is partially supported by a UBFC-ISITE-BFC Grant.

References

[1] EUROCAE Working Group 51. Safety, performance and interoperability requirements document for ads-b/nra application. Technical report, The European Organisation for Civil Aviation Equipment, 2005.

[2] Hiroshi Akima. A new method of interpolation and smooth curve fitting based on local procedures. 17:589–602, 1970.

[3] A. Barreto, M. Hieb, and E. Yano. Developing a complex simulation environment for evaluating cyber attacks. In *Interservice/Industry Training, Simulation, and Education Conf. (I/ITSEC)*, volume 12248, pages 1–9, 2012.

[4] P. Brooker. Sesar and nextgen: investing in new paradigms. *The Journal of Navigation*, 61(2):195–208, 2008.

[5] Nazly Rocio Santos Buitrago, Loek Tonnaer, Vlado Menkovski, and Dimitrios Mavroeidis. Anomaly detection for imbalanced datasets with deep generative models.

[6] A. Cretin, B. Legeard, F. Peureux, and A. Vernotte. Increasing the resilience of ATC systems against false data injection attacks using DSL-based testing. In *Proc. of the 8^{th} Int. Conf. on Research in Air Transportation (ICRAT'18), Doctoral Symp.*, pages 1–4, Spain, 2018.

[7] Gyorgy Dan and Henrik Sandberg. Stealth attacks and protection schemes for state estimators in power systems. In *Smart Grid Communications (SmartGridComm), 2010 First IEEE Int. Conf. on*, pages 214–219. IEEE, 2010.

[8] J F Epperson. On the runge example. *The American Mathematical Monthly*, 94(4):329–341, 1987.

[9] E Habler and A Shabtai. Using lstm encoder-decoder algorithm for detecting anomalous ads-b messages. *Computers & Security*, 78:155–173, 2018.

[10] X. Kong, X. Song, F. Xia, H. Guo, J. Wang, and A. Tolba. Lotad: long-term traffic anomaly detection based on crowdsourced bus trajectory data. *World Wide Web*, 21(3):825–847, May 2018.

[11] Yao Liu, Peng Ning, and Michael K Reiter. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security (TISSEC)*, 14(1):13, 2011.

[12] M. Ma. Resilience against false data injection attack in wireless sensor networks. In *Handbook of Research on Wireless Security*, pages 628–635. IGI Global, 2008.

[13] P. Malhotra, A. Ramakrishnan, G? Anand, L. Vig, P. Agarwal, and G. Shroff. Lstm-based encoder-decoder for multi-sensor anomaly detection. *arXiv preprint arXiv:1607.00148*, 2016.

[14] M. Riahi Manesh and N. Kaabouch. Analysis of vulnerabilities, attacks, countermeasures and overall risk of the automatic dependent surveillance-broadcast (ADS-B) system. *Int. Journal of Critical Infrastructure Protection*, 19:16 – 31, 2017.

[15] I. Martinovic and M. Strohmeier. Security of ads- b: State of the art and beyond. *DCS*, 2013.

[16] T. Maruthi Padmaja, Narendra Dhulipalla, Raju S. Bapi, and P. Radha Krishna. Unbalanced data classification using extreme outlier elimination and sampling techniques for fraud detection. pages 511–516. IEEE.

[17] H. R. Roth, L. Lu, J. Liu, J. Yao, A. Seff, K. Cherry, L. Kim, and R. M. Summers. Improving computer-aided detection using convolutional neural networks and random view aggregation. *IEEE Transactions on Medical Imaging*, 35(5):1170–1181, May 2016.

[18] M. Schäfer, V. Lenders, and I. Martinovic. Experimental analysis of attacks on next generation air traffic communication. In *Int. Conf. on Applied Cryptography and Network Security*, pages 253–271. Springer, 2013.

[19] Thomas Schlegl, Philipp Seeböck, Sebastian M. Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery, 2017.

[20] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, and M. Wilhelm. Bringing Up OpenSky: A Large-scale ADS-B Sensor Network for Research. In *Proc. of the 13th Int. Sym. on Information Processing in Sensor Networks*, IPSN '14, pages 83–94, Germany, 2014.

[21] Z. Shi, M. Xu, Q. Pan, B. Yan, and H. Zhang. Lstm-based flight trajectory prediction. In *2018 Int. Joint Conf. on Neural Networks (IJCNN)*, pages 1–8, July 2018.

[22] A Smith, R Cassell, T Breen, R Hulstrom, and C Evers. Methods to provide system-wide ADS-B back-up, validation and security. In *25th Digital Avionics Systems Conf.*, pages 1–7. IEEE, 2006.

[23] M. Strohmeier. *Security in Next Generation Air Traffic Communication Networks*. PhD thesis, Oxford University, 2016.

[24] M. Strohmeier, M. Schäfer, R. Pinheiro, V. Lenders, and I. Martinovic. On perception and reality in wireless air traffic communications security. *IEEE Transactions on Intelligent Transportation Systems*, 18(6):1338–1357, 2017.

[25] K. D Wesson, T. E Humphreys, and B. L Evans. Can cryptography secure next generation air traffic surveillance? *IEEE Security and Privacy Magazine*, 2014.

[26] Le Xie, Yilin Mo, and Bruno Sinopoli. False data injection attacks in electricity markets. In *Smart Grid Communications (SmartGridComm), First Int. Conf. on*, pages 226–231. IEEE, 2010.

[27] Chuanlong Yin, Yuefei Zhu, Jinlong Fei, and Xinzheng He. A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, 5:21954–21961, 2017.

[28] R. Zhang, G. Liu, J. Liu, and J P Nees. Analysis of message attacks in aviation datalink communication. *IEEE Access*, 2017.