Intrusion Detection for Airborne Communication using PHY-Layer Information

Martin Strohmeier[1], Vincent Lenders[2], Ivan Martinovic[1]

[1]University of Oxford, Oxford, United Kingdom
[2]armasuisse, Thun, Switzerland

Abstract. With passenger and cargo traffic growing rapidly world-wide, and unmanned aerial vehicles (UAV) poised to enter commercial airspaces, a secure next generation of air traffic management systems is required. Recent articles in the academic and hacker community highlight crucial security challenges faced by integral parts of these next generation protocols, with the most dangerous attacks based on classic message injection. In this article, we analyze the possibility and effectiveness of detecting such attacks on critical air traffic infrastructures with a single receiver based on physical layer information. Using hypothesis testing and anomaly detection schemes, we develop an intrusion detection system (IDS) that can accurately detect attackers within 40 seconds.

1 Introduction

The air traffic load has experienced tremendous growth over the last decade. The reported average number of registered flight movements over Europe is around 26,000 per day. Large European airports may spike to more than 1,500 daily takeoffs and landings. This tendency is still increasing and forecasts assume that movements will nearly double between 2009 and 2030. With growing adoption of unmanned aerial vehicle technology for civil applications, we may even expect an additional boost in overall air traffic over the coming years.

The Automatic Dependent Surveillance-Broadcast (ADS-B) protocol is a crucial part of the procedural improvements of the next generation of air traffic management. In less dense airspaces above large unpopulated areas such as in Canada, Australia, or the Atlantic Ocean, ADS-B is already the only means of air traffic surveillance today. With single sensors providing a coverage radius of up to 400 km, the system offers not only high accuracy but is also very cost-efficient. Both of these features are strong drivers of a quick adoption and the use of ADS-B will be mandatory by 2017 in Europe and 2020 in the US. However, the protocol is also widely considered to be insecure by hacker and academic communities and by practitioners because of its lack of authentication. Consequently, recent high-profile cases of aircraft incidents such as the disappearance of Malaysian aircraft MH370 or hijacked emergency signals created a lot of speculation about insecure air traffic control (ATC) protocols [3,8].

Due to decade-long roll out and planning times for new protocols and related prohibitive costs, there is currently no upgrade on the horizon which could address the security flaws of ADS-B in the foreseeable future. Taking the former

into account, there is an urgent need for separate, transparent countermeasures that do not require modifications to the current ADS-B systems but can significantly improve the real-world security of the protocol.

In this paper, we make the following contributions:

- We develop an IDS based on physical layer measurements to detect false-data injection attacks into ATC networks in less than 40 seconds without additional cooperation by the aircraft or infrastructure overhead.
- We analyse different features based on statistical tests and combine them into a unified approach using one-class anomaly detection.
- We validate our system against real-world data from our OpenSky sensor network and simulated attackers conducting message injection attacks.

Related Work There are several works that use statistical testing of received signal strength (RSS) patterns to detect attackers in wireless networks but to the best of our knowledge, this work is the first to apply such techniques in the unique aircraft domain. The works most similar to ours are [2] and [12]. In [2], the authors consider attacks on RSS-based wireless localization systems in WiFi and ZigBee. Their models use statistical hypothesis testing to detect significant deviations from the expected RSS readings of the landmarks used for localization. While we also utilize statistical tests in our IDS, the aircraft location problem is different since an attacker does not use signal strength to change the outcome of the localization but directly injects messages with false data. In [12], the authors use RSS patterns to detect the spoofing of a MAC address, which shares some similarities to spoofing identities with ADS-B. They analyze antenna diversity and use it to improve on the examined detection algorithms.

In contrast to LANs, in our work we exploit the location data encoded in ADS-B, and the large velocities and distances found in air traffic. While RSS is a difficult property in settings without line-of-sight (LOS) that are affected by multi-path, the LOS propagation of air traffic communication provides sound conditions for physical layer schemes. Furthermore, we go beyond statistical tests and apply an anomaly detection approach that can integrate arbitrary features.

2 Overview of ADS-B Security Concerns

In this section, we give a short overview of the ADS-B protocol, its known security flaws and non-technical considerations about potential solutions.

The ADS-B Protocol

Currently rolled out into all major airspaces, and mandatory by 2017 (Europe) / 2020 (USA), ADS-B is a satellite-based replacement of traditional primary and secondary surveillance radar systems. Aircraft use onboard satellite navigation (e.g., GPS) to fetch their own position and velocity; these and other relevant data are periodically transmitted by the *ADS-B Out* subsystem. The broadcasted messages are processed by ATC ground stations, and in the future also by other aircraft close by, if equipped with *ADS-B In* (see Fig. 1 for an illustration).

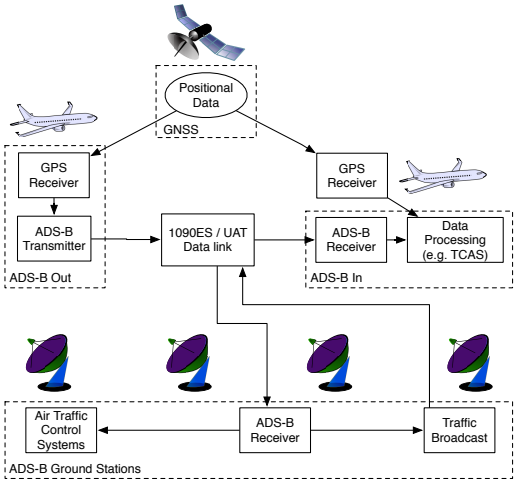

Fig. 1. ATC system architecture. [14] The position provided by the global navigation satellite system (GNSS) is processed by the aircraft and broadcasted through the ADS-B Out system alongside other situational information. ATC ground stations and other aircraft (via ADS-B In) receive these messages over the two possible data links, 1090 Extended Squitter (1090 ES) or Universal Access Transceiver (UAT).

Security Overview In recent years, ADS-B's susceptibility to radio frequency attacks has generated a lot of attention in hacker circles [4,6], the mainstream media [15], and among academic researchers [7,10]. It has been shown that an attacker can easily record and analyze the unencrypted ADS-B messages. Worse, an adversary actively interfering with ATC communication poses a severe threat to aviation safety. As adversarial action on the ADS-B data link can also impact the traffic collision avoidance system (TCAS), it is crucial to deploy countermeasures promptly to facilitate widespread deployment of the protocol.

When the ADS-B protocol was designed in the early 1990s, precise manipulation of radio frequency communication was possible only for powerful military adversaries. The required cost and engineering knowledge were considered too prohibitive to add security mechanisms to the protocol. With the recent advent of cheap, accessible software-defined radios and specialized hardware for the reception of ATC communication, the threat model has shifted considerably. Today, typical wireless attacks such as eavesdropping, jamming and modification, insertion and deletion of messages are feasible for anyone with widely available off-the-shelf hard- and software (see, e.g., [4,7,10]). For a full overview of such attacks and their potential impact, and also possible ways to address these vulnerabilities, see [13]. Here, we focus on the insertion of fake data into radar systems as detailed in the next section. Crucially, all proposed countermeasures require either upgrades to the protocol or a large number of sensors to facilitate physical layer defenses such as passive localization. These characteristics make them unsuitable in many scenarios due to some non-technical considerations:

Legacy requirements A viable security design for ADS-B must not require changes to the existing protocol, or additional cooperation from the aircraft. This legacy requirement is common to slow-changing industries such as aviation. ADS-B, for example, has been in development since the early 1990s and is only now being deployed, more than two decades later. Hence, countermeasures against ADS-B attacks need to work alongside the current system without disrupting it.

Cost effectiveness Cost is considered a main driver for the adoption of new ATC protocols. Conventional radar technologies are both more expensive to deploy and experience much higher maintenance cost compared to ADS-B. The International Civil Aviation Organization (ICAO) specifies the technological cost of operating traditional radar techniques to monitor an en-route airspace at $6-14 million, while ADS-B surveillance comes in significantly cheaper at $380,000 [5]. The ability to rely solely on ADS-B data would be very cost effective. This is a crucial argument, especially considering the massive investments already made during the development of ADS-B. Countermeasures requiring a large number of stations also negate this cost advantage and ignore the reality of ATC deployments in Canada, Australia or over oceans, where single sensors cover a radius up to the radio horizon of about 400 km.

The Case for Intrusion Detection

As argued in [13] and [14], we believe that given the current state of the ADS-B roll out, there is a strong need for transparent countermeasures as cryptographic means are not a feasible option in the medium term due to the requirements discussed above. Air traffic management as a critical infrastructure system has many characteristics of supervisory control and data acquisition (SCADA) systems. Cardenas et al. [1] note that threats on these systems need to be dealt with by defense-in-depth mechanisms and anomaly detection schemes. They argue an adversary may hide the *specific* exploits but cannot conceal their ulterior goals and intentions. Indeed, there must be a noticeable adverse effect to the *physical* system (i.e., the management of air traffic), otherwise the attack may even be ignored, e.g., when somebody is simply relaying live ADS-B data.

As such physical effects are achieved through injection of malicious data which does not match the expected behaviour, an anomaly detection system can help with the discovery of the attacker and provide the base for defense-in-depth mechanisms. A high rate of attack detection is at the heart of any such system where non-detection might cause disastrous consequences. However, in the real world low false positive rates are just as crucial. While they can normally be sorted out by using voice communication with the aircraft, constant nagging and false alarms can potentially have an adverse effect on overall system safety.

3 Modeling False-Data Injection Attackers

In the following, we describe the model that an attacker uses to inject false data into an ADS-B target receiver. The injection of false data provides the basis

of most of the attacks on the ADS-B system as discussed in the literature [7]. Executed correctly, they are subtle but have devastating effects on the system.

We assume that the attacker injects a ghost aircraft, either collected at an earlier time and replayed, or created from scratch. In both cases, we assume a non-naive attacker who has sufficient knowledge to inject valid-looking messages that are well-formed with reasonable content, withstanding a superficial check.

This means the attacker creates correctly formatted ADS-B messages, covering the expected types (position, velocity, identification) in valid sequential orders and spacings according to the standard specification [9]. We also assume the attacker uses a legitimate ICAO address and reasonable flight parameters (e.g., believable altitude and speed) to create a valid-looking aircraft that cannot be distinguished from a real one using standard ATC procedures.

Signal Strength We model the attacker's use of different RSS patterns using a single antenna. We assume all attackers are more or less stationary on the ground attacking specific sensors in transmission distance, i.e., we do not consider UAVs. Weather effects on RSS have proven negligible for our use case [14].

- **Attacker 1:** This attacker uses a straight-forward constant sending strength, resulting in a Gaussian distribution due to the noisy nature of the channel. Without loss of generality, we assume the standard settings of a typical software-defined radio with a 100 mW power output and a distance of 500 m to the sensor under attack. This creates a signal with a RSS of about -65 dBm at the receiver; the standard deviation of the random noise is 3.5 dB.
- **Attacker 2:** The RSS is a random variable X, within the limits of the hardware. To simulate a random stationary non-adjusting attacker, we assume the RSS received at the attacked sensor to be fully random within the typical values of legitimate aircraft (in our case, the 5%/95% percentiles are -75.60 dBm and -63.54 dBm, respectively).
- **Attacker 3:** This attacker adjusts the sending strength in an attempt to be in line with the position the injected messages are representing to the attacked sensors. More concretely, the attacker knows the position of the receiver with a maximum error of 1 km (mean: 500 m) on which he bases the calculation of the distance to the claimed flight positions.

Our goal is to get an accurate read of legitimate aircraft behavior, enabling us to detect all but the most knowledgeable, powerful and carefully carried out attacks by entities who have perfect knowledge of the IDS and its sensor locations.

4 Intrusion Detection

In this section, we describe the physical layer features that we select for our IDS and how we combine them in a unified detection approach. When receiving ADS-B messages from an aircraft, the ground station can measure and store the RSS. Due to the attacker's positioning on the ground, the measurements

of injected ADS-B messages are highly unlikely to match the RSS of legitimate samples. Furthermore, they should be comparably constant over time compared to aircraft covering distances of hundreds of miles in relation to the receiver. Using standard hypothesis testing, an IDS can judge the probability whether a collected RSS sample stems from a legitimate aircraft or not.

Pearson Correlation Coefficient In physical space, we calculate the Pearson correlation coefficient ρ between the distance (derived from the position claim in the ADS-B messages) and the RSS. Path loss suggests a strong negative relationship in legitimate flights, while an injection attacker who does not adjust the sending strength in line with the claimed distance should show no correlation. Formally, we test the null hypothesis $\mathcal{H}_0$ stating that there is no association between the two variables in the population against the alternative hypothesis $\mathcal{H}_A$, stating that there is a negative association between the two variables in the population:

$$\mathcal{H}_{0:}\ \rho = 0 \tag{1}$$

$$\mathcal{H}_A : \rho < 0 \tag{2}$$

We consider a sample where $\mathcal{H}_0$ is rejected at the 99% significance level a legitimate flight sample and an attack if the hypothesis is accepted.

Autocorrelation Coefficient In signal space, we use the autocorrelation coefficient (ACF) to identify attackers that are stationary and/or do not adapt their sending strength. Autocorrelation is the cross-correlation of a signal with itself. It can be used to show that a time series is not random, but instead exhibits significant correlations between the original observations and the same observations shifted backwards by a lag τ. The ACF helps to find repeated patterns such as periodic signals in a noisy channel. Formally, we test the null hypothesis $\mathcal{H}_0$ which states that there is no autocorrelation $R(\tau)$ in the population against the alternative hypothesis $\mathcal{H}_A$, saying that there is a a positive autocorrelation:

$$\mathcal{H}_{0:}\ R(\tau) = 0 \tag{3}$$

$$\mathcal{H}_A : R(\tau) > 0 \tag{4}$$

We run these tests for lags 1 to 8 and take their mean to create a single measure for finding autocorrelation significant at the 1% level. We again consider a sample where $\mathcal{H}_0$ is rejected at the 99% significance level a legitimate flight sample and an attack if the hypothesis is accepted.

Detection of Multiple Antennas Legitimate ADS-B-equipped flights send alternatingly using two separate antennas, one on top of the aircraft and one on the bottom, as specified in [9]. This setup creates a behavior that a sophisticated attacker needs to mimic. Fig. 2 shows an example of the distinctive RSS patterns. To exploit this feature, we divide the full RSS time series into their two antenna subparts according to their time slots and compare various features that

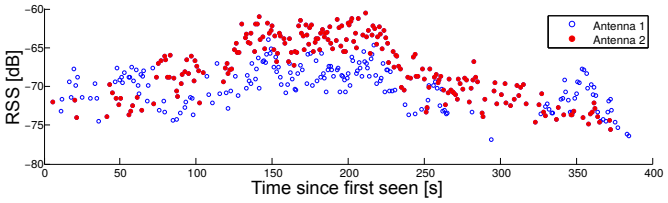

Fig. 2. RSS samples of a flight's two separate antennas.

show only on the newly created time series. For example, with 300 samples per flight, we found a difference of around 1.8 dBm ($\sigma = 1.4$) in the means of the two antennas in our sample data. A single-antenna attacker, who does not adapt his sending power to mimic two antennas, is expected to exhibit no significant difference between the RSS of messages in alternating time slots. Based solely on RSS time series, we can identify other differences between a single-antenna user (i.e., an anomaly that would most likely be caused by an attacker) and messages sent out by commercial aircraft:

- The ACF of the divided antenna time series falls much faster than the one by a single-antenna attacker.
- Even lags (2, 4, 6, 8) of the combined ACF are greater than odd lags.
- Similarly, the ACF for a lag of 1 is typically higher for the separated antennas, while for an attacker divided and combined ACF are similar.

Furthermore, we found that separating the antennas first vastly improves the results of the correlation features discussed in this section.

Combined Anomaly Detection We combine our features in a one-class classification problem. One-class classifiers try to separate one class of data, the target data, from the rest of the feature space. Our target class is a well-sampled class of aircraft behavior based on collected RSS data. The outlier class is unknown and online target samples are used at the time of learning. The process creates an n-dimensional classifier, where n is the number of features. For new samples, this classifier decides if they fit into the expected space or if they are rejected (i.e., classified as an anomaly worth investigating).

5 Experimental Design

First, we analyze the effectiveness of our selected features on their own, using standard hypothesis testing before we combine them with a machine learning approach to create a more robust IDS. We employ the MATLAB toolkits Dd_Tools and PRTools[1] to create data descriptions of our air traffic data. We define one-class datasets based on legitimate data collected with an ADS-B sensor and use various one-class classifiers to create descriptions which include the data.

[1] See http://prlab.tudelft.nl/david-tax/dd_tools.html and http://prtools.org

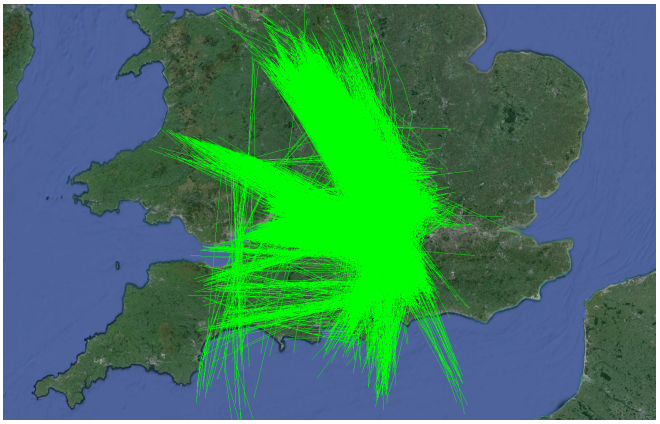

Fig. 3. Visualization of the 7,159 flight trajectories used for our anomaly analysis.

Data We used a data sample consisting of 7,159 flights, each flight with 200 or more received messages, collected over 24 hours and visualized in Fig. 3. The data collection was conducted with an OpenSky sensor installed at the top of our lab building. OpenSky is a participatory sensor network that collects raw ADS-B message data and stores them in a database for further research [11].

For our anomaly detection approach, we test several different classifiers with 5-fold cross validation and the fraction of outliers in training set to zero (i.e., all training samples are accepted as legitimate). While the training sets are drawn from our collected sample of legitimate flights only, the separate test sets for each attacker have an added 2% of falsely-injected data (amounting to 143 flights) to be detected by the classifier. To verify our models and test our IDS, the RSS patterns of the attackers are simulated as described in Section 3.

6 Results

Table 1 shows the results of the examined detection approaches. The hypothesis tests each detect attackers 1 and 2 with more than 99% probability. Especially the autocorrelation feature proves to be accurate, with few legitimate flights misclassified as false positives (0.1%). As expected, both tests fail to detect the more sophisticated attacker 3. To counter this, we analyze the distinct antenna characteristics, which detects over 90% of all three attackers with a false positive rate of 3.9%. On its own, the antenna method requires 300 messages to become reliable enough, as aircraft may move in ways that can obfuscate their antenna features in the short run.

With the combined classifier, we can accurately detect all attackers 1 and 2 without false negatives and one single false positive (less than 0.01%), using a small RSS sample of 200 messages. At the standard rate of 5.4 ADS-B messages per second, this allows detection in under 40 seconds, assuming no message loss. Even with a typical loss of 30% [14], this can be achieved in less than one minute.

Table 1. Effectiveness of the examined detection approaches. We used 7,159 legitimate flights and 143 simulated attackers for every class, with 200+ messages per flight. The percentages show the average detection rates over 5-fold cross validation.

Detection Rate [%]	attacker 1	attacker 2	attacker 3	legit flights (FPs)
Pearson	99.8	99.9	0.2	18.6
Autocorrelation	99.6	99.4	0.3	0.1
Antenna Detection	92.6	94.0	95.5	3.9
Combined Detection	100	100	98.8	<0.01

As illustrated in Fig. 4 a), attacker 3 who easily deceives the individual hypothesis tests, can be *too* good. He would need to introduce additional randomness and patterns similar to the spoofed airplanes to fall within the expected data range. This demonstrates the strength of the anomaly detection approach where the precise type of anomaly need not be known in advance. The results may see further improvement through the collection of more samples. This naturally increases the confidence of the system and improves detection results at the cost of slower reaction times.

Fig. 4 b) shows the results of the comparison between various tested classifiers, depending on the number of samples. The Parzen classifier performs best, having the lowest number of misclassified attackers. It is followed by K-Means, but the Minimax, Minimum Spanning Tree and k-Nearest Neighbors classifiers also achieve a near-zero false negative rate as 200 samples are collected, still significantly improving on pure hypothesis testing.

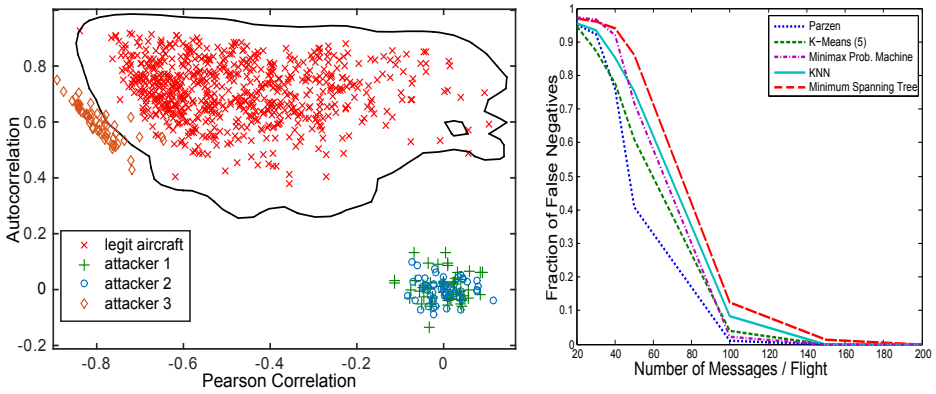

Fig. 4. a) 2D-Parzen classifier example with 200 collected samples. Red crosses are legitimate flight samples. Attacker 1 and 2 are entirely classified as anomaly here, while attacker 3 creates few false positives. **b)** Complete classifier comparison with 5-fold cross validation. Joint false negative rates for attackers 1 + 2.

7 Conclusion & Future Work

In this article, we proposed an IDS for false-data injection attacks on the ADS-B protocol used in air traffic control. We provided a threat model for such injection attacks on ADS-B and developed an IDS based on RSS measurements. We validated our system against real-world data from our OpenSky sensor network and found that the Parzen classifier performed best in our sample. In future work, we plan to analyze more sophisticated attackers, additional features to deal with them such as the angle of arrival, and the long-term stability of our system.

References

1. Cardenas, A.A., Amin, S., Lin, Z.S., Huang, Y.L., Huang, C.Y., Sastry, S.: Attacks against process control systems: risk assessment, detection, and response. In: Proceedings of the 6th ACM symposium on information, computer and communications security. pp. 355–366. ACM (2011)
2. Chen, Y., Xu, W., Trappe, W., Zhang, Y.: Attack detection in wireless localization. In: Securing Emerging Wireless Systems, pp. 1–22. Springer (2009)
3. Clayton, M.: Malaysia Airlines Flight MH370: Are planes vulnerable to cyber-attack? Christian Science Monitor (Mar 2014)
4. Costin, A., Francillon, A.: Ghost in the Air (Traffic): On insecurity of ADS-B protocol and practical attacks on ADS-B devices. In: Black Hat USA (2012)
5. ICAO: Guidance Material on Comparison of Surveillance Technologies (GMST). Tech. Rep. September (2007)
6. Kunkel, R.: Air Traffic Control Insecurity 2.0. In: DefCon 18 (2010)
7. McCallie, D., Butts, J., Mills, R.: Security analysis of the ADS-B implementation in the next generation air transportation system. International Journal of Critical Infrastructure Protection 4(2), 78–87 (Aug 2011)
8. Moran, N., De Vynck, G.: WestJet Hijack Signal Called False Alarm. Bloomberg (Jan 2015)
9. RTCA Inc.: Minimum Operational Performance Standards for 1090 MHz Extended Squitter Automatic Dependent Surveillance – Broadcast (ADS-B) and Traffic Information Services – Broadcast (TIS-B). DO-260B with Corrig. 1 (2011)
10. Schäfer, M., Lenders, V., Martinovic, I.: Experimental analysis of attacks on next generation air traffic communication. In: Applied Cryptography and Network Security. No. 7954 in LNCS, Springer (Jun 2013)
11. Schäfer, M., Strohmeier, M., Lenders, V., Martinovic, I., Wilhelm, M.: Bringing Up OpenSky: A Large-scale ADS-B Sensor Network for Research. In: ACM/IEEE International Conf. on Information Processing in Sensor Networks (2014)
12. Sheng, Y., Tan, K., Chen, G., Kotz, D., Campbell, A.: Detecting 802.11 MAC layer spoofing using received signal strength. In: INFOCOM 2008. The 27th Conference on Computer Communications. IEEE. IEEE (2008)
13. Strohmeier, M., Lenders, V., Martinovic, I.: On the Security of the Automatic Dependent Surveillance-Broadcast Protocol. Communications Surveys & Tutorials, IEEE PP(99) (2014)
14. Strohmeier, M., Schäfer, M., Lenders, V., Martinovic, I.: Realities and Challenges of NextGen Air Traffic Management: The Case of ADS-B. Communications Magazine, IEEE 52(5) (May 2014)
15. Zetter, K.: Air traffic controllers pick the wrong week to quit using radar. Wired (Jul 2012)