Aircraft Fingerprinting Using Deep Learning

Alessandro Nicolussi, Simon Tanner, Roger Wattenhofer
ETH Zurich
Switzerland
{anicolus, simtanner, wattenhofer}@ethz.ch

***Abstract*—Aircraft periodically broadcast their position, identity and other information using the ADS-B protocol. This allows safe air traffic flow as ground stations and other aircraft can depend on the sent information. However, these messages are not authenticated or encrypted. Therefore, this system is vulnerable to attacks from Software Defined Radios (SDRs) and other transmitters. We propose a deep learning-based approach for fingerprinting of aircraft messages based on physical characteristics. This helps to verify the origin of an observed message.**

***Index Terms*—Aircraft fingerprinting, deep learning, radio frequency fingerprinting**

I. Introduction

Aircraft determine their exact position using satellite navigation and then broadcast it over a protocol called Automatic Dependent Surveillance - Broadcast (ADS-B). They additionally also send information such as their velocity and identification. These messages can then be received by nearby aircraft and on the ground by air traffic control.

However, ADS-B does not provide any means of message authentication and therefore the origin of an aircraft message cannot be verified. This means that an attacker using a Software Defined Radio (SDR) or another ADS-B transponder can inject false information without being detected. The attacker could create messages of aircraft that do not exist or even jam messages from an actual aircraft and then send modified messages that seem to originate from the aircraft. These attacks have become easily possible in recent years due to the wide availability of SDRs.

Unfortunately, improvements to the ADS-B protocol cannot be implemented quickly as new hardware would be required in each aircraft. Authentication or encryption schemes face additional challenges of key management and distribution [1]. Therefore, only approaches that do not need modified aircraft transponders can be implemented quickly to increase the security of ADS-B.

We propose a deep learning-based method for the identification based on the radio signal of the aircraft. To verify the origin of ADS-B messages, we extract different fingerprints from aircraft messages. Using these fingerprints, we examine whether we can determine the originating aircraft based on a single received ADS-B message.

The observed differences in the fingerprints of different aircraft can have different origins. The hardware design varies between transponder manufacturers and hardware versions and oscillators have different errors. Additionally, silicon-based hardware is known to be very complex. Small deviations in the manufacturing process are inevitable and therefore no two transponders are perfectly identical. This results in observable characteristics in the transmitted signals, such as different rise times/fall times, frequency errors, etc. We examine how such physical-layer characteristics can be used to tell individual transponders apart.

Aircraft transmit their identity and possibly other identifying data in the messages. As we cannot trust the data contained in the messages sent by the aircraft, we have to be careful not to give this information to the fingerprinting methods.

To record the data from actual aircraft, we use cheap and widely available RTL-SDR receivers. Flight tracking services such as FlightAware and Flightradar24 use the same receivers.

II. Related Work

Radio frequency fingerprinting techniques have been proposed to enhance the security in multiple wireless communication protocols. Reising et al. examined physical-layer fingerprinting for mobile phones from different manufacturers operating on the GSM standard [2]. The authors show the possibility of enhancing security in GSM communication by extracting physical-layer fingerprints from different sections of GSM signals.

Device fingerprinting has also been applied to identify Bluetooth transceivers [3]. The authors extracted fingerprints from the energy envelope of the transient signal and could distinguish between seven individual transceivers. These approaches require a high signal-to-noise ratio and are evaluated over short distances. The fingerprinting of aircraft communication however has to be performed over large distances with a wide range of reception conditions.

Radio frequency fingerprinting has also been considered for securing aircraft communication over ADS-B. The phase during a message has been used to classify the aircraft into one of seven classes depending on the shape of the phase pattern [4]. Also the received signal strength has been used to identify messages that do not originate from an observed aircraft [5]. Multilateration techniques have been proposed to localize aircraft and verify the origin of ADS-B messages [6], [7]. For an attacker equipped with a single spoofing device, multilateration makes it very hard to inject illegitimate aircraft messages as the attacker has to travel along the claimed path. However, a multi-device attacker performing a geographically distributed coordinated attack on the receiving base stations

The authors of this paper are alphabetically ordered.

can imitate any location as the injected message's origin. For the detection of coordinated attacks on aircraft multilateration systems, a different approach for fingerprinting has been considered [8]. The authors compare the physical-layer fingerprints determined at different receivers of the same aircraft message to each other. Therefore, they can determine whether the message has been sent by a single transmitter at the correct location. Also software-dependent features such as the interarrival time of messages can be used to cluster aircraft [9]. This however assumes that the considered messages come from the same transmitter.

There have also been suggestions to add authentication to ADS-B. Berthier et al. [10] suggest SAT, a new, backward-compatible replacement for ADS-B. Using their proposed SAT transponders, aircraft fitted with the corresponding equipment can receive and send both ADS-B and SAT messages. The SAT messages are authenticated using TESLA [11], a delayed key disclosure mechanism to achieve broadcast authentication without prior key exchange.

Message authentication approaches to secure the communication however need adapted transceivers in the aircraft or even modifications of the ADS-B protocol. These approaches are interesting for a future revision of the protocol but do not solve the current security threat.

Deep learning has successfully been applied to many areas of signal processing. It is used for tasks such as image classification [12] and generation [13]. It has also been applied to analyze and synthesize music [14]. Both, music signals and the radio frequency signals used in aircraft communication, are one-dimensional time series signals. Deep learning has also shown good results for many tasks in communication systems such as modulation recognition, channel estimation, etc. [15].

III. Background

Mode S is a secondary surveillance radar technique that allows a selective interrogation of aircraft to report information, such as altitude. A ground-based transceiver sends the interrogation to which the aircraft responds with a message sent at $1090\,\mathrm{MHz}$. The information is encoded using Pulse Position Modulation (PPM). A high signal level followed by a low signal level denotes a 1-bit, and vice-versa, a 0-bit. Each symbol has a duration of $1\,\mu\mathrm{s}$. A Mode S message starts with a preamble of $8\,\mu\mathrm{s}$ consisting of a fixed pattern of four pulses. In Mode S, two lengths of messages exist, with 56 and 112 bits.

To remove the requirement of selective aircraft interrogation, Automatic Dependent Surveillance - Broadcast (ADS-B) was introduced. Aircraft periodically broadcast information such as position, velocity, and identification over this protocol. ADS-B messages are sent using the Mode S Extended Squitter format which carries 112 bits. Figure 1 shows the message format. The first 8 bits are made up of 5 bits Mode S *downlink format (DF)* which is always 17 in the case of ADS-B, and 3 bits of *capability (CA)*. The 24 bits constituting the unique *ICAO address* for each aircraft are static and sent along with every message. Information about position, altitude, velocity, etc. is encoded in the 56 bits of *message field*. A 24-bit *cyclic redundancy check (CRC)* is employed to detect and account for transmission errors. In the USA, aircraft must be equipped with an ADS-B capable transponder since the beginning of 2020 depending on the airspace [16]. Also in the European Union, ADS-B will become mandatory in June 2020 for large aircraft [17].

DF CA	ICAO address	message field	CRC
8 bits	24 bits	56 bits	24 bits

Fig. 1: The ADS-B message format

ADS-B does not support any authentication or encryption of the sent messages. Therefore, many different attacks on the ADS-B system are possible [1]. We consider an attacker that is able to inject fake messages into a ground based receiver. We assume that a sophisticated attacker is able to send correctly formatted messages that also follow a believable aircraft path. We can therefore not rely on the message content for fingerprinting.

IV. Data Collection

To collect many messages from aircraft, we record ADS-B messages from multiple ground stations located across Switzerland into a database. Each ground station consists of a Raspberry Pi 3 with an attached RTL-SDR. This hardware setup consists of easily available and affordable components.

The RTL-SDR samples the signal with $2.4\,\mathrm{MS/s}$. The dump1090 [18] decoder software detects and decodes the messages sent from the aircraft. After decoding, the messages are forwarded to the database server together with the I/Q samples of each message.

The messages were recorded on 11 days. On each of the days, a network of six to seven receivers was recording for approximately three hours. If the same message is received by multiple receivers, the corresponding signal data recorded by all receivers is used. As training set we use data collected on the first six days, the test set is made up of the last five days. In both sets, we require at least 2000 received transmissions from each aircraft. This provides us with 274 common aircraft in both training and test data, with a total of three million messages over 11 days. The validation data set is randomly split off the training data.

V. Fingerprints

The goal of the fingerprinting approach is to identify the aircraft without trusting the ICAO address sent in each message. Additionally also other contents in the message might give away the identity of the aircraft. The positions of the aircraft might allow to identify the aircraft, as the same aircraft may always be on the same route over multiple days. Therefore, we employ only fingerprinting methods that are independent of the message content. We base the fingerprints on the phase and magnitude of the message preamble and on the phase over the whole message. The phase of Mode S messages is not specified, as only the magnitude is used to encode the data in the PPM scheme. Therefore, we expect the phase to be useful for distinguishing aircraft.

The observed differences in the fingerprints of different aircraft can have different origins. The hardware design might vary between transponder manufacturers and hardware versions, oscillators have different errors and manufacturing differences may also lead to individual fingerprints.

A. Preamble

The preamble of 8 µs lends itself to fingerprinting as it is common to every Mode S message and does therefore not contain any data. Additionally, the first five bits of each ADS-B transmission denote the same downlink format, which is always 17 for ADS-B messages. We can also use the raw signal data encoding these bits without leaking any identification information. We expect to see differences between aircraft in the phase values during the preamble as only the magnitude is specified by the pulse position modulation. Additionally also the exact position in time and the magnitude of the pulses might vary slightly.

B. Phase pattern

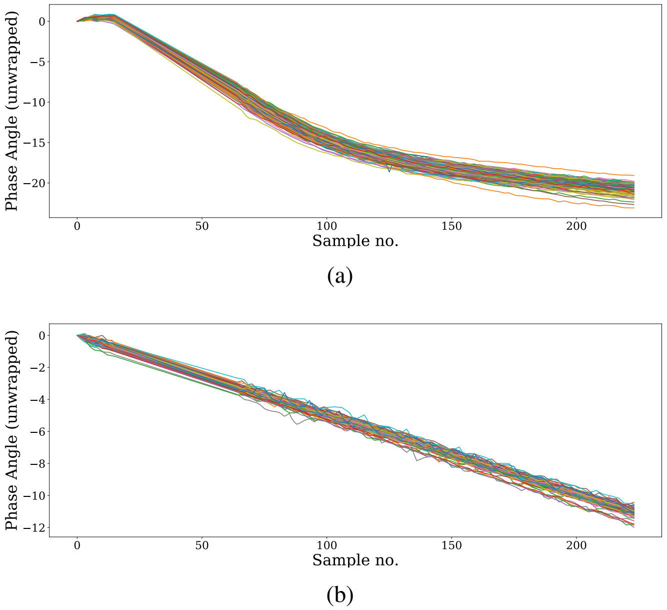
The preamble is relatively short compared to the whole message duration of 120 µs. Therefore, we expect to detect more physical-layer information in the whole transmission.

We consider the phase pattern over the entire message. The phase of the signal is only determined by the aircraft during the pulses. Between the pulses, the phase is random. Therefore, only the phase at the pulse positions should be evaluated. Otherwise, the message content is visible by considering the variance of the phase. The classifier might then recognize the aircraft by the ICAO address or other identifying content of the message, which may be spoofed by an attacker.

We sample the phase at each magnitude pulse and interpolate at non-pulse positions. This results in a sampling rate of 2 MS/s. Additionally, we completely remove and interpolate the phase samples at the position of the 24-bit ICAO address. Since the phase angle of the signal for every sample is in the range $(-\pi, \pi]$, we unwrap the phase to obtain a continuous phase pattern.

As aircraft are travelling at high speeds, the Doppler effect causes a frequency shift of up to 1000 Hz. This results in phase patterns with different slopes. With the position and velocity of the aircraft, we can calculate the Doppler shift with respect to each receiving ground station and can compensate it to obtain high quality fingerprints.

Figures 2a and 2b show collections of phase patterns of messages from two different aircraft. However, not for all aircraft the patterns are that easily distinguishable.



(a)

(b)

Fig. 2: Phase patterns of multiple messages from two different aircraft.

VI. Classification

All classifiers used for the fingerprinting are CNN-based and have been optimized for the individual fingerprints. As predictions, the classifiers output vectors where each element expresses the confidence of the classifier that a signal originated from a certain aircraft.

For the preamble fingerprinting, we upsample the raw signal by a factor of 20 from 2.4 to 48 MS/s. The network classifying based on preamble magnitudes consists of two convolutinal layers, each followed by a MaxPool layer with pool size of two. The two convolutional layers consist of 64 and 48 filters, respectively, with a kernel size of 64. Two dense layers with 512 and #(aircraft) units produce the categorical output. All but the last layer employ the ReLU activation, the last layer uses the Softmax activation function. During training, the Adam optimizer minimizes categorical crossentropy loss. The network classifying based on the phase during the preamble, uses a similar network architecture. The only difference is a kernel size of 28 instead of 64 for the convolutional layers.

For phase patterns, a CNN with four convolutional layers and one dense layer is used. Average pooling with a pool size

TABLE I: Classification metrics using preamble magnitudes

	precision	recall	F1-score
macro avg	0.173	0.171	0.156
accuracy: 0.171			

TABLE II: Classification metrics using preamble phase

	precision	recall	F1-score
macro avg	0.278	0.284	0.264
accuracy: 0.284			

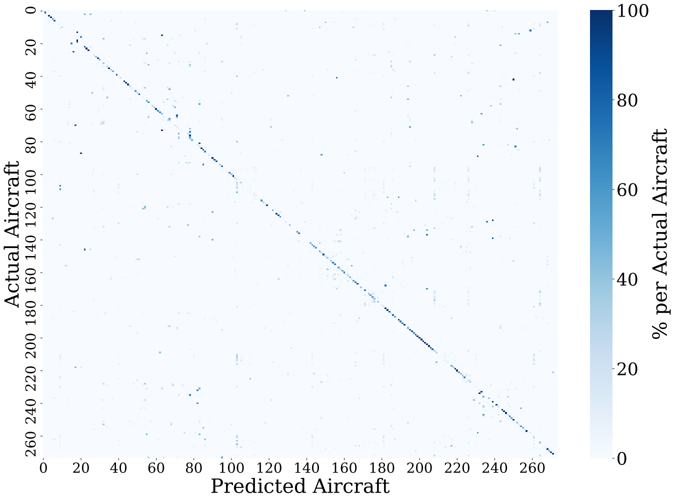


Fig. 3: Confusion matrix for aircraft classification using phase patterns

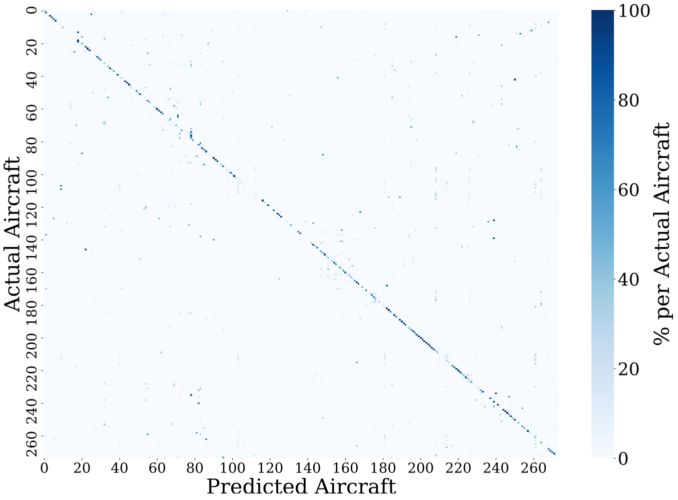


Fig. 4: Confusion matrix for aircraft classification when combining preamble phase and phase patterns

of two is applied before the first, third, and fourth layer. The first two convolutional layers consist of 32 filters, layers three and four consist of 64 filters, each with a kernel size of 64. The dense layer has #(aircraft) units and produces the output. LeakyReLU is used as activation for all but the last layer which uses the Softmax activation. The categorical crossentropy loss is optimized by the Adam optimizer.

VII. Evaluation

As evaluation metric, we consider accuracy on the test set, defined as the percentage of signals for which the correct aircraft was predicted. We undersample the test set to overcome class imbalance and obtain more balanced results. Otherwise aircraft that are over-represented in the test data contribute more to the classification accuracy.

A. Preamble

The preamble magnitude allows to expose information about the pulse positions which are not uniform among different aircraft transponders. Even at a low sampling rate of $2.4\,\mathrm{MS/s}$, we expect to still extract some characteristics about the exact pulse positions.

Among 274 aircraft, our CNN-based classifier achieves an accuracy of $17.13\,\%$ and an F1-score of 0.156 as we can see in Table I. This is still relatively high when we take into consideration the number of aircraft and the limited resolution of the recorded signals.

Using the same upsampling strategy as for the magnitude of the preamble, we extract the phase over the ADS-B transmission start.

As can be seen in Table II, with a classification accuracy of $28.4\,\%$ and an F1-score of 0.264, we can conclude that the phase contains considerably more information about transponder hardware than the magnitude of an aircraft signal.

B. Phase pattern

When we consider the phase pattern, i.e., the interpolated and unwrapped phase angle over an entire ADS-B transmission, we can see in Figure 3 that this fingerprint allows to very accurately classify certain aircraft transponders.

Even at a low sampling rate of $2\,\mathrm{MS/s}$, an accuracy of $36.7\,\%$ and an F1-score of 0.342 is achieved among signals from 274 individual aircraft (see Table III). The fingerprints vary slightly by day, as environmental conditions or uncontrollable factors impacting the transponder's oscillator on the test days do not perfectly reproduce the conditions on the training days. As certain aircraft have very similar phase patterns, they become indistinguishable.

C. Combining predictions

As different classifiers show different confidences for certain aircraft, we examine the usefulness of combining fingerprints. This is where the categorical predictions from each classifier are beneficial.

We achieve the best results by combining preamble phase and phase patterns over the whole message. With a classifica-

TABLE III: Classification metrics using phase patterns

	precision	recall	F1-score
macro avg	0.367	0.367	0.342
accuracy: 0.367			

TABLE IV: Classification metrics when combining preamble phase and phase patterns

	precision	recall	F1-score
macro avg	0.416	0.419	0.391
accuracy: 0.419			

tion accuracy of $41.9\,\%$ and an F1-score of 0.391 as shown in Table IV, we observe considerable improvements compared to using only phase patterns. The confusion matrix in Figure 4 is very similar to the one for only the phase patterns in Figure 3, but some of the outliers have disappeared.

This means among 274 aircraft observed over eleven days, for four out of ten received messages, the originating aircraft transponder can be determined using only physical-layer characteristics.

Combining the predictions of the different fingerprints leads to a more reliable classification. While it is not possible to reliably classify all aircraft, it is possible to classify many aircraft transponders accurately.

VIII. Conclusion

We have introduced a novel fingerprinting method for ADS-B messages using deep learning methods.

Our results show that it is possible to distinguish between 274 aircraft that fly by on multiple days with an accuracy of $41.9\,\%$. This result shows that the used fingerprints are stable over multiple days. Compared to [4] that classified messages into seven classes, our physical-layer fingerprints can distinguish many more transmitters.

The proposed physical-layer fingerprinting approach for single ADS-B messages could be used in combination with other security mechanisms such as higher-level fingerprinting techniques [9], e.g. using the interarrival times, or aircraft multilateration to build a complete security system for ADS-B.

References

[1] D. McCallie, J. Butts, and R. F. Mills, "Security analysis of the ADS-B implementation in the next generation air transportation system," *IJCIP*, vol. 4, no. 2, pp. 78–87, 2011. [Online]. Available: https://doi.org/10.1016/j.ijcip.2011.06.001

[2] D. R. Reising, M. A. Temple, and M. J. Mendenhall, "Improving intra-cellular security using air monitoring with RF fingerprints," in *2010 IEEE Wireless Communications and Networking Conference, WCNC 2010, Proceedings, Sydney, Australia, 18-21 April 2010*, 2010, pp. 1–6. [Online]. Available: https://doi.org/10.1109/WCNC.2010.5506229

[3] S. U. Rehman, K. W. Sowerby, and C. Coghill, "RF fingerprint extraction from the energy envelope of an instantaneous transient signal," in *Australian Communications Theory Workshop, AusCTW 2012, Wellington, New Zealand, January 30 - Feb. 2, 2012*, 2012, pp. 90–95. [Online]. Available: https://doi.org/10.1109/AusCTW.2012.6164912

[4] M. Leonardi and D. Di Fausto, "Secondary surveillance radar transponders classification by rf fingerprinting," in *2018 19th International Radar Symposium (IRS)*. IEEE, 2018, pp. 1–10.

[5] M. Strohmeier, V. Lenders, and I. Martinovic, "Intrusion detection for airborne communication using phy-layer information," in *Detection of Intrusions and Malware, and Vulnerability Assessment - 12th International Conference, DIMVA 2015, Milan, Italy, July 9-10, 2015, Proceedings*, ser. Lecture Notes in Computer Science, M. Almgren, V. Gulisano, and F. Maggi, Eds., vol. 9148. Springer, 2015, pp. 67–77. [Online]. Available: https://doi.org/10.1007/978-3-319-20550-2_4

[6] J. Krozel and D. Andrisani, "Independent ads-b verification and validation," in *AIAA 5th ATIO and16th Lighter-Than-Air Sys Tech. and Balloon Systems Conferences*, 2005, p. 7351.

[7] M. Schäfer, M. Strohmeier, V. Lenders, I. Martinovic, and M. Wilhelm, "Bringing up opensky: A large-scale ads-b sensor network for research," in *Proceedings of the 13th International Symposium on Information Processing in Sensor Networks*, ser. IPSN '14. Piscataway, NJ, USA: IEEE Press, 2014, pp. 83–94. [Online]. Available: http://dl.acm.org/citation.cfm?id=2602339.2602350

[8] D. Moser, P. Leu, V. Lenders, A. Ranganathan, F. Ricciato, and S. Capkun, "Investigation of multi-device location spoofing attacks on air traffic control and possible countermeasures," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking, MobiCom 2016, New York City, NY, USA, October 3-7, 2016*, 2016, pp. 375–386. [Online]. Available: https://doi.org/10.1145/2973750.2973763

[9] M. Strohmeier and I. Martinovic, "On passive data link layer fingerprinting of aircraft transponders," in *Proceedings of the First ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy, CPS-SPC 2015, Denver, Colorado, USA, October 16, 2015*, 2015, pp. 1–9. [Online]. Available: https://doi.org/10.1145/2808705.2808712

[10] P. Berthier, J. M. Fernandez, and J. Robert, "Sat : Security in the air using tesla," in *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)*, Sep. 2017, pp. 1–10.

[11] A. Perrig, R. Canetti, J. D. Tygar, and D. Song, "The tesla broadcast authentication protocol," *Rsa Cryptobytes*, vol. 5, no. 2, pp. 2–13, 2002.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[13] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv preprint arXiv:1710.10196*, 2017.

[14] H. Purwins, B. Li, T. Virtanen, J. Schlüter, S. Chang, and T. N. Sainath, "Deep learning for audio signal processing," *J. Sel. Topics Signal Processing*, vol. 13, no. 2, pp. 206–219, 2019. [Online]. Available: https://doi.org/10.1109/JSTSP.2019.2908700

[15] T. Wang, C. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep Learning for Wireless Physical Layer: Opportunities and Challenges," *CoRR*, vol. abs/1710.05312, 2017. [Online]. Available: http://arxiv.org/abs/1710.05312

[16] FAA, "Airspace," https://www.faa.gov/nextgen/equipadsb/research/airspace/, accessed 2020-03-01.

[17] EASA, "Seasonal technical communication," https://www.easa.europa.eu/sites/default/files/dfu/EASA_STC_NEWS_JUNE_2018.pdf, accessed 2020-03-01.

[18] O. Jowett, "Github - mutability/dump1090," https://github.com/mutability/dump1090, accessed 2020-03-01.