Secondary Surveillance Radar Transponders classification by RF fingerprinting

Mauro Leonardi*, **Davide Di Fausto****

*University of Rome "Tor Vergata"
Rome, ITALY
email: mauro.leonardi@uniroma2.it

**University of Rome "Tor Vergata"
Rome, ITALY
email: ddifausto@gmail.com

Abstract: *Secondary Surveillance Radars are Air Traffic Control systems used to obtain identity and altitude of the cooperative airplanes and, together with the Primary Radar, allow a safe air traffic flow. The Secondary Surveillance Radar protocol is also used for Traffic Alert and Collision Avoidance System and Automatic Dependent Surveillance Systems: in these applications, the aircraft transmit their own information (identity, position, velocity etc.) to any equipped listener for anti-collision and surveillance scope without the interrogation of the SSR Radar. The simple Secondary Surveillance Radar protocol doesn't provide any kind of authentication and encryption, making it vulnerable to many types of cyber-attacks. In the paper, it is proposed the use of the airplane/transmitter carrier phase as a feature to perform a classification of the aircraft and, therefore, distinguish legitimate messages from fake ones. The feature extraction process is described and different classification methods are tested by the use of real data.*

1. Introduction

Secondary Surveillance Radars (SSRs) are Air Traffic Control (ATC) systems used to obtain identity and altitude of the cooperative airplanes and, together with the Primary Radars, allow a safe air traffic flow. In this system SSR interrogates the aircraft requiring its identity and/or altitude; the aircraft, that is equipped with a device called *transponder*, replies with a message (called *reply*) containing the requested information. The last implementation of SSR (called *MODE S*) has introduced a protocol evolution, that is able to make selective interrogation and an unique identification code for each aircraft (called *ICAO Address*) is assigned[1]. This new protocol is also used for Traffic Alert and Collision Avoidance System (TCAS) and Automatic Dependent Surveillance-Broadcast Systems (ADS-B)[1]. In these applications, the aircraft transmit their own information (identity, position, velocity etc.) for anti-collision and surveillance scope without the interrogation of SSR Radar. Moreover, concerning the ADS-B system, it is considered one of the pillars of the Future Air Traffic Systems[2][3] and, nowadays, about 80% of the commercial aircraft are equipped with the ADS-B hardware[4].

978-3-7369-9545-1

The previously mentioned SSR reply is composed of a four pulses preamble and a data-block of 56 or 112 pulses where the requested information are coded with a 24-bit CRC[5][1]. Every message also contains the 24-bit ICAO Address[5]; in Figure 1 the reply format is reported. In case of ADS-B, this message is sent without interrogation (and it is called *squitter*) and contains, in addition to the ICAO address, various information, such as the on board computed position and the aircraft velocity.

This protocol was introduced in the 80s when security and cyber-attacks with RF manipulation of the communication weren't considered as easy as they are nowadays and it does not offer any encryption and authentication technique. Today the protocol security aspect must be reconsidered as illustrated in [6][7][8][9]. Possible attacks to Mode S channel are: ***Eavesdropping***, i.e. listening to the transmissions: it is impossible to be prevented without applying encryption and, of course, it is impossible to be detected; ***Jamming***, i.e. the intentional transmission of high power harmful signals in the RF channel in order to disable the airground communication; ***Message injection (or spoofing)***, i.e. the intentional transmission of signals with the same protocol but with misleading information; ***Message deletion*** by SSR reply Garbling: legitimate messages can be deleted or manipulated by the superposition of false message with higher power; ***Transponder not authorized substitution/ICAO address modification***. All these vulnerabilities may produce a security/safety risk.

In this paper we will focus on ***False Aircraft Injection***, that is the ability of the attacker to send false replies/squitters in the channel emulating the presence of one or more aircraft. This can be easily done also using low cost Software Defined Radio (SDR). To mitigate this problem we propose to classify the transponder's transmitters applying the so-called RF Fingerprinting. It consists on the identification (or classification) of a wireless device by the extraction of unique features embedded in the electromagnetic waves emitted by its transmitter. These unique features arise from randomness in the manufacturing process such as, for example, the presence of analog components in the transmission chain, different HW and SW implementation of the device, transmitter clock stability etc. Once particular features of the transmitter are identified it is possible to create a Database (or Library) of trusted aircraft/transponders (or of classes of trusted transponders) and check if the signal received from an airplane has the expected features (i.e. it is generated from the transmitter associated with the aircraft in the Database) and, if not, raise an alarm.

We will focus on the phase of transmitter carrier along the message and we will verify if this feature can be used for transponder fingerprinting. Afterwards, automatic classification techniques for this feature will be introduced and tested by the use of real data. In the next Section we will describe the phase signature feature and its extraction procedure. The feature characteristics and the identification of different aircraft classes are reported in Section 3 and, in Section 4, classification methods are introduced and evaluated using real data.

2. RF signature extraction

Consider the Mode S down-link format reported in Figure 1: the Pulse Position Modulation (PPM) implies that, neglecting the preamble, the Data-Block is always composed of $m = 112$ or $m = 54$ pulses with different time positions to encode the information to be transmitted (i.e. Manchester coding)[5][10]. By now we always use, without loss of generality, $m = 112$.

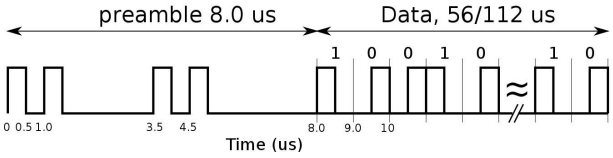


Figure 1: ADS-B/Mode S reply format.

The ICAO standards [5][10] allow the manufacturers to develop transmitting devices with some tolerances on various parameters and no restrictions exist concerning the carrier phase of the message due to the fact that all the information is coded on the signal amplitude. For these reasons the transmitted signal (considering only the data-block) $s_t(t)$ can be represented as follows:

$$s_t(t) = A(t) \left[\sum_{m=1}^{112} g\left(t - 2mT + c_m T + T/2\right) \right] \sin\left[2\pi\left(f_C + \delta f\right) t + \phi(t)\right] \tag{1}$$

where $A(t)$ is the message amplitude, $\{c_m\}$ is the bits sequence to be transmitted (composed of $m = 112\ bits$), f_C is the carrier frequency equal to 1090 MHz, T is the pulse width equal to 0.5 μsec, δf is the allowed jitter of the carrier frequency, $g(t)$ is a function that represents the real shape of the transmitted pulse considering the specification on rise and decay time. Finally $\phi(t)$ is the carrier phase. Assuming the use of a coherent receiver with IF sampling and the presence of Additive White Gaussian Noise (AWGN), the received signal become:

$$s_r(k) = s_r(kT_s) = \tag{2}$$
$$= A(kT_s) \cdot \left[\sum_{m=1}^{112} g\left(kT_s - 2mT + c_m T + T/2\right) \right] \sin\left[2\pi\left(f_{IF} + \delta f\right) kT_s + \phi(kT_s)\right] + n(kT_s)$$

where $n(kT_s)$ represents the noise and T_s is the sampling time. We have assumed equal to zero the propagation delay from the transmitter to the receiver only to simplify the notation.

Now, focusing only on the carrier phase behaviour inside the 112 μsec of transmitted Data-Block, it is possible to estimate 112 different phase values $\hat{\phi}_m$, one for each pulse of the ADS-B message using the following formula[11]:

$$\hat{\phi}_m = \arctan\left[\frac{\sum_K s_r(kT_S) \sin(2\pi\left(f_{IF} + \delta f\right) kT_S)}{\sum_K s_r(kT_S) \cos(2\pi\left(f_{IF} + \delta f\right) kT_S)} \right] \tag{3}$$

where m identifies the pulse and K the relative samples.

It must be noted that to perform this computation it is mandatory to know: (1) the time position of each pulse (it can be easily determined estimating the time of arrival of the message, with a preamble detection algorithm and then decoding the envelope of the received signal) and, (2) the central frequency of the message $f_C + \delta f$ (it may include also the Doppler frequency due to the airplane velocity) that can be estimated using any kind of frequency estimator such as, for example, $argmax(FFT(s_r(k)))$.

The proposed extraction process can be easily integrated into a classical Mode S receiver as illustrated in Figure 2: the classical envelope decoder, that provides the sequence of message bits, can be used to estimate the pulses time positions; besides a frequency estimator is used to estimate the residual carrier frequency of the signal. Finally, pulses time positions and residual carrier frequency are used to estimate the phase signature of the Mode S Reply/Squitter.

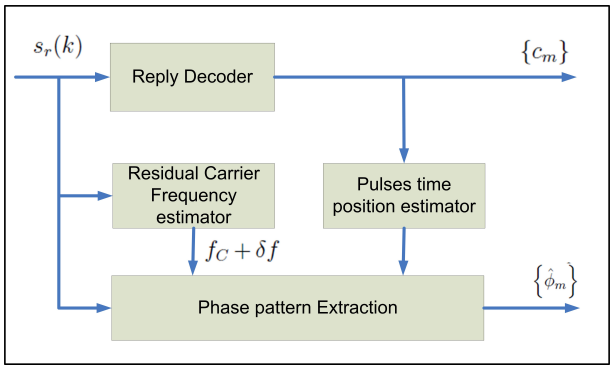


Figure 2: Proposed phase signature extraction block diagram.

Moreover, without loss of generality, all the phase values can be referred to the first one and a phase unwrapping procedure can be applied to obtain the sequence $\left\{\hat{\phi}_m\right\}$. We expect that $\left\{\hat{\phi}_m\right\}$ depends on the transmitter peculiarities (such as stability of the oscillator, phase noise, transmitter clock, etc.) varying for different transponders (e.g. different vendor, hardware version, Firmware version, etc). To verify these two characteristics, a measurement campaign with real data has been done.

3. Real Data Analysis

To verify if the sequence $\left\{\hat{\phi}_m\right\}$ can be used for fingerprinting, a measurement campaign with real data has been done using a Mode S receiver called *Transponder Data Recorder (TDR)*[12]. The TDR is a Mode S receiver composed of four independent linear channels and one logarithmic channel. Each receiving channel is connected to an element of an array antenna; the linear channels down-convert the signals to an intermediate frequency (IF) at 21.5 MHz and the logarithmic channel is based on the Analog Devices AD8313 log receiver with a base-band output.

The digital section is based on an NI platform with 6 AD converters with sample rate up to 100 Msamples/s. Pictures of the TDR elements are reported in Figure 3. The antenna has been

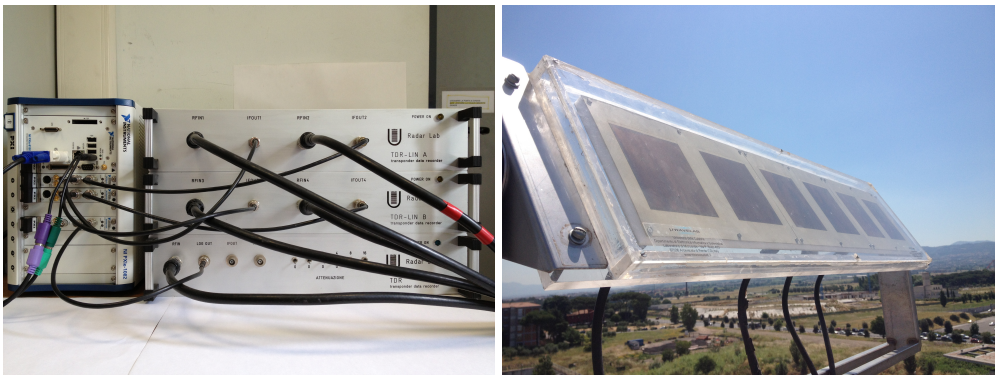


Figure 3: **TDR system pictures: on left the TDR receiver, on right the TDR antenna.**

installed on the Engineering Faculty roof for 4 consecutive days, receiving 660182 messages sent by 676 different aircraft. Examples of measured phase signatures, $\left\{\hat{\phi}_m\right\}$, are reported in Figure 4. For each graph, the phase signatures obtained from different messages coming from the same aircraft, are plotted. As expected, some airplanes don't have a particular phase signature and different messages are uncorrelated with each other (see Figure 4.(d)), but others have very specific signatures (see Figure 4.(a-b-c-e-f-g)). Seven different signature classes have been discovered by visual inspection: *(a) Linear, (b) Quadratic, (c) Oscillating, (d) Non-Coherent, (e) Mixed: quadratic+linear, (f) Mixed: linear+linear* and *(g) Wave* (see also [13]).

This result is important because, although standards and recommendations don't require any phase restriction, many real transponders use a precise oscillator to generate signals and different aircraft show different phase signatures. All the first two days received replies have been classified by inspection. Classification results for the first day are summarized in Table 1: the airplanes that don't change their phase signature during all the day are about the 70% of the total amount. Most of these belong to the *non coherent* class followed by the *linear* ones and the *quadratic* ones. Results for the second day are quite similar to the first one. Moreover it has been discovered that the 52.5% of common airplanes in the two days belong to the same class of the day before. Summing up, it can be affirmed that at least the 50% of the observed aircraft are classifiable within the proposed classes and don't change their phase signature for two consecutive days.

4. Transponder automatic classification

In our particular application, aircraft classification will be used to understand if the received signals are really generated from the expected aircraft or not. Many methods are proposed in the literature to solve this kind of problem[14][15]: each method has pros and cons with different performances for different applications. In our case we have approached two classification techniques: Neural Network (NN) and K-Nearest Neighbors (KNN) with a Principal Component

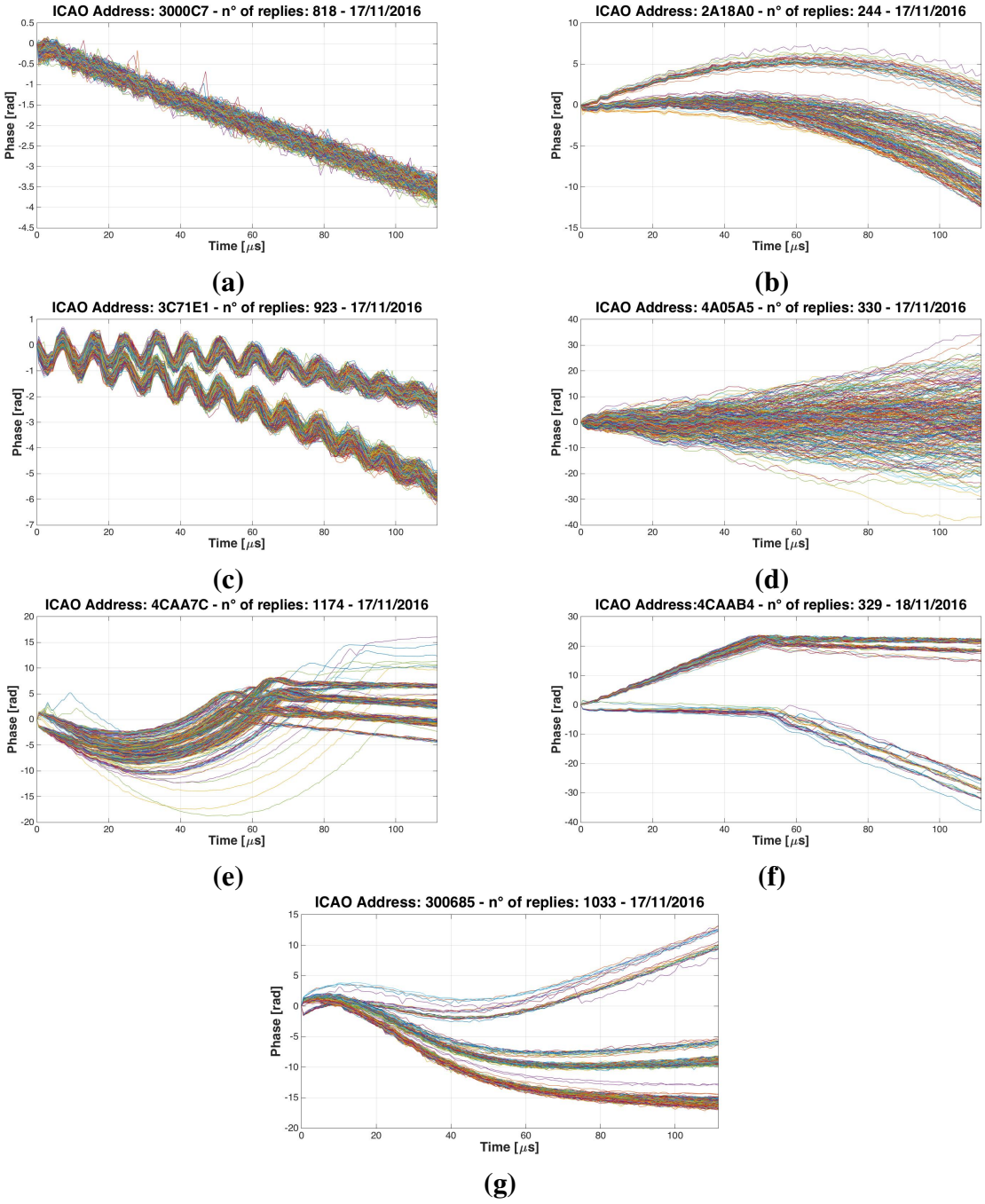


Figure 4: Phase signature examples of seven different aircraft: (a) Linear, (b) Quadratic, (c) Oscillating, (d) Non-Coherent, (e) Mixed: quadratic+linear, (f) Mixed:linear+linear and (g) Wave

Analysis preprocessing (PCA)[16]; in the following, the performance of these two techniques will be compared.

Whatever the chosen technique is, a pre-elaboration of the sequence $\left\{\hat{\phi}_m\right\}$ is needed to reduce its cardinality. First of all, it should be noted that the sequence $\left\{\hat{\phi}_m\right\}$ is not homogeneous distributed in the time because phase measures are made in the pulses that change their time

Table 1: Classification results for day 17/11/2016.

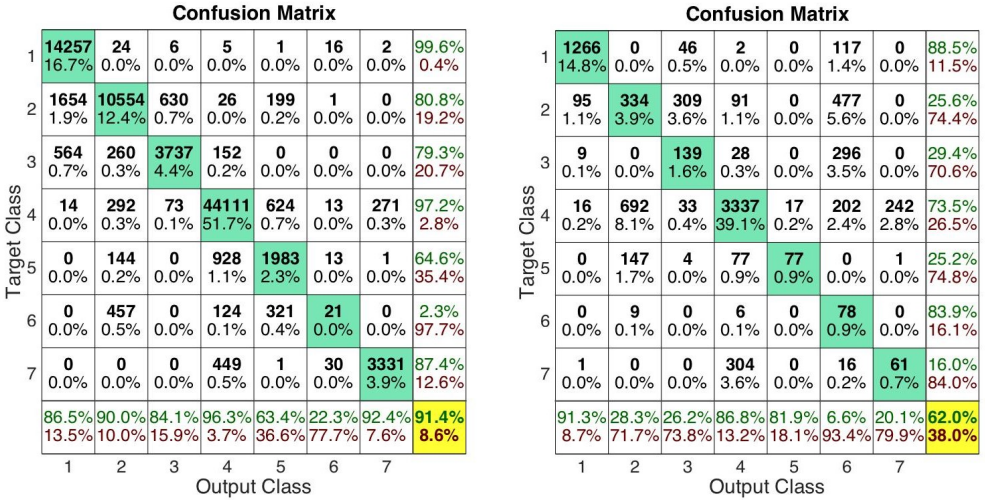
Calssification Results for day 17/11/2016		
Total number of replies	232888	
Classified replies	151226 (64%)	
Assigned class for day 17/11/2016		
Class	**n. of aircraft**	**% of aircraft**
1 - Linear	40	19.8
2 - Quadratic	27	13.4
3 - Oscillating	14	6.9
4 - Non Coherent	109	53.9
5 - Mixed: Quadratic+Linear	8	4.0
6 - Mixed: Linear+Linear	1	0.5
7 - Wave	3	1.5

position according to the PPM modulation. To overcome this problem an interpolation is used to evaluate the phase also in the time space where the pulse is not present, obtaining a new sequence $\left\{\hat{\phi}^m\right\}$ composed of 224 elements. After that, we propose two alternative approaches to reduce the sequence dimension: *Single Reply Analysis - SRA*: a simple 5:1 decimation is applied to the sequence obtaining 45 phase values (in this manner the classification can be done using a single reply); *Groups of Replies Analysis - GRA*: the second grade best-fitting polynomial, $y = ax^2 + bx + c$, is determined for each reply and, for every group of *N* consecutive replies of the same aircraft, the following six features are extracted : $E(a)$, $E(b)$, $E(c)$, $std(a)$, $std(b)$ and $std(c)$. This allows performing the classification every N replies relative to the same aircraft. A comparison of the classification performance of the two proposed classification techniques (NN and KNN+PCA) in case of *SRA* or *GRA* follows.

Concerning the Neural Network, a very simple Multi-Layer Feed-Forward NN[17] has been used. This Network is composed of three layers: an Input layer (45 elements in case of *SRA* and 6 elements in case of *GRA*); a 10 neurons Hidden layer; an Output layer composed of 7 elements that is the number of identified aircraft classes. The NN has been trained using the real data coming from the first day (151266 replies) and validated with the data of the second day (85289 replies). The *GRA* is performed grouping $N = 20$ consecutive replies. Concerning the KNN, it considers every input sequence (45 elements for *SRA* and 6 elements for *GRA*) as a point in a M-dimensional space (where M is the number of the input element). A classification model is created with the training data and, for every test, the M-dimensional Euclidean distance is calculated respect to every training data[16][18]; the algorithm finds the K training points nearest to the test data and assigns it to the class whose the most of those K training data belong. For our application $K = 3$ value has been used. To maintain low the computational cost of this technique (the higher M, the higher computational cost), data have been per-processed by Principal Component Analysis to reduce M, considering only the most relevant components, called Principal Components: these are computed as linear old components combination[19]. M has been chosen equal to 5 for *SRA* and 6 for *GRA*. The KNN+PCA classification model has

been trained and validated with the same data of the NN and the obtained classification results are shown in Figure 5 and in Figure 6 using the so called *Confusion Matrices*. They indicate how the chosen classification model is efficient for the considered dataset and for every different class which the analysed data have been divided into. Each matrix column represents the predicted class of the test data, while each row indicates the real class. All the input data in the diagonal (from top left square to bottom right one) are the right-classified data. Every instance that does not belong to this diagonal has been wrongly classified (the predicted class does not correspond to the real one). Finally bottom right square summarizes the total performances of the utilized classification model and technique.

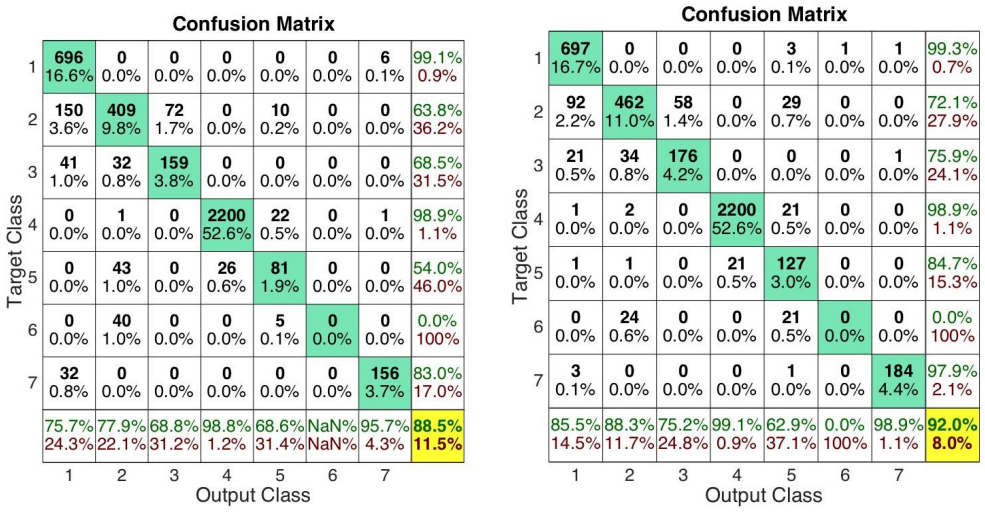
The results show that in case of *SRA* (reported in figure 5) the NN performances are higher than K-NN ones and that, using aircraft single replies, K-NN model cannot be considered efficient due to its overall performance, in term of correct classification probability equal to 62%. Conversely, using *GRA* approach (as shown in figure 6), K-NN method reaches 92% of correct



(a) **(b)**

Figure 5: Confusion matrices of Neural Networks classification (a) and PCA+KNN classification (b) for Single Reply Analysis

classification, i.e. better than NN performances (88.5%). These results show that K-NN model could be considered the right supervised classification technique for the *GRA* approach and NN could be considered the right one for the *SRA*. More in detail, both methods show excellent performances for some classes (class 1 Linear and class 4 Non-Coherent) but lower (or poor) results for classes 5 and 6. In the considered training data set, in fact, only few signals belong to these classes (see Table 1). This situation provides a non-well trained classification model for mixed classes. For this reason, we have tried to exclude the mixed classes from the classification process implementing also a 5 classes NN in order to obtain a performance improvement in term of correct classification. However, with this configuration, results show only small performances improvements (2%) and, in the author's opinion, it should be better to maintain the class number as higher as possible and, for future applications, to improve the training phase with a larger data-set.



(a) **(b)**

Figure 6: Confusion matrices of Neural Networks classification (a) and PCA+KNN classification (b) for Groups of Replies Analysis

Finally, it must be remembered that: (1) the NN ensures lower computational cost and memory usage respect to KNN because the algorithm, after the training phase, has to use only the NN weights and not all the training data used to compute the classification model as the K-NN does; (2) *SRA* with NN classification method performs a test every time a message is received and for this reason it is capable to provide a very low Time to Alarm that could be essential for some applications. Contrariwise *GRA* approach with KNN+PCA can be used together with other *group features* such as Time Difference of Arrival between consecutive messages coming from the same aircraft (e.g. as described in [20]). Concluding, considering that the performance of NN/SRA and KNN+PCA/GRA are quite similar (91.% and 92%), the choice must be done only selecting the one more appropriate to the type of approach (SRA or GRA).

5. Conclusion

This work shows that it is possible to classify Mode S transponders using the phase signature of the transmitted signal. It was discovered, by real data analysis, that many real transponders use a very stable oscillator that produces a specific signature and that, using some classification techniques, it is possible to distinguish at least 7 different classes of aircraft. More than the 50% of observed aircraft have a particular and representative phase signature that can be used for automatic classification of the transponder within one of the seven classes and the classification performances reach the 93% in terms of correct classification probability. In the author's opinion, the phase signature, jointly with other transmitted signals characteristics (e.g. carrier frequency stability, pulse shapes, message timings etc.) could be used for a more complex classification providing an improvement of air traffic security.

References

[1] M. Stevens, *Secondary Surveillance Radar*. Artech House, 1988.

[2] SESAR, "http://www.sesarju.eu/."

[3] NEXTGEN, "https://www.faa.gov/nextgen/."

[4] M. Strohmeier, "Large-scale analysis of aircraft transponder data," *IEEE Aerospace and Electronic Systems Magazine*, vol. vol. 32, pp. pp. 42–44, 2017.

[5] *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 Corrigendum 1*, RTCA Inc., Dec. 2011.

[6] M. Strohmeier, V. Lenders, and I. Martinovic, "On the security of the automatic dependent surveillance-broadcast protocol," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1066–1087, 2015.

[7] M. Leonardi, E. Piracci, and G. Galati, "Ads-b vulnerability to low cost jammers: Risk assessment and possible solutions," *2014 Tyrrhenian International Workshop on Digital Communications - Enhanced Surveillance of Aircraft and Vehicles (TIWDC/ESAV)*, pp. pp. 41–46, 2014.

[8] J. Butts, D. McCallie, and R. Mills, "Security analysis of the ads-b implementation in the next generation air transportation system," *International Journal of Critical Infrastructure Protection*, vol. 4, no. 2, pp. 78–87, 08 2011.

[9] M. Leonardi, E. Piracci, and G. Galati, "Ads-b jamming mitigation: a solution based on a multi-channel receiver,," *IEEE Aerospace and Electronic Systems Magazine*, vol. 32, no. 11, pp. 44–51, 11 2017.

[10] *Annex 10 to the Convention on International Civil Aviation Aeronautical Telecommunication*, ICAO, 1998.

[11] A. Goldsmith, "Wireless communication," *Cambridge University Press*, 2005.

[12] G. Galati, M. Leonardi, E. Piracci, N. Petrochilos, and S. Samanta, "The transponder data recorder: Implementation and first results," *IEEE Aerospace and Electronic Systems Magazine*, vol. 29, no. 2, pp. 6–13, 02 2014.

[13] M. Leonardi, L. D. Gregorio, and D. D. Fausto, "Air traffic security: Aircraft classification using ads-b messages phase-pattern," *Aerospace*, 10 2017.

[14] M. Strohmeier, "Security in next generation air traffic communication networks," Ph.D. dissertation, University of Oxford, 2016.

[15] 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," *ACM Conference on Mobile Computing and Networking*, 2016.

[16] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: a survey," *ACM Computing Surveys*, 2009.

[17] M. Moller, "A scaled conjugate gradient algorithm for fasy supervised learning," *Neural Network*, vol. 6, pp. 525–533, 1993.

[18] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed., W. . Sons, Ed. Wiley-Interscience, 2001.

[19] J. Shlens, "A tutorial on principal component analysis," 2014.

[20] 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*, A. N. York, Ed., Denver, Colorado,USA, 10 2015.