OpenSky Report 2020: Analysing in-flight emergencies using big data

Xavier Olive¶‖, Axel Tanner**, Martin Strohmeier¶†§, Matthias Schäfer¶‡*, Metin Feridun§, Allan Tart¶, Ivan Martinovic¶†, Vincent Lenders¶§

¶OpenSky Network, Switzerland lastname@opensky-network.org ‖ONERA, Université de Toulouse, France xavier.olive@onera.fr **IBM Research, Switzerland axs@zurich.ibm.com

*TU Kaiserslautern, Germany schaefer@cs.uni-kl.de †University of Oxford, UK firstname.lastname@cs.ox.ac.uk ‡SeRo Systems, Germany schaefer@sero-systems.de §armasuisse, Switzerland firstname.lastname@armasuisse.ch

***Abstract*—Transponder codes are four octal digit numbers transmitted by an aircraft transponder in response to a secondary surveillance radar interrogation. These discrete transponder codes (also known as squawk codes) help with the clear labelling of an affected aircraft on radar screens. Three particular squawk codes are associated with specific situations: 7500 for hijacking, 7600 for radio failure and 7700 for general emergencies, often related to medical or technical issues.**

In this paper, we analyse more than 800 trajectories received by the OpenSky Network over a two-year period as they were broadcasting the 7700 emergency code. Background information about the reason of these emergencies is taken from social networks and other crowdsourced information sources on the Internet. We provide an overview of various reasons for in-flight emergencies, typical trajectory patterns and communication strategies by airlines. Based on our semi-labelled dataset of trajectories, we also train models able to suggest possible explanations for trajectories when no information is available.

I. Introduction

A transponder is a key electronic on-board device which helps identifying aircraft on air traffic control (ATC) radars. Traditionally, transponders produce a response on the 1090 MHz frequency after receiving a radio-frequency interrogation on 1030 MHz from Secondary Surveillance Radar (SSR) systems (both from the ground and other aircraft). In particular, they support the transponder Modes A, C and S, traffic alert and collision avoidance systems (TCAS) [1], and the novel Automatic Dependent Surveillance – Broadcast (ADS–B), which does not require interrogations.

Air traffic controllers use the term *squawk* to assign aircraft a transponder code which is used to identify an aircraft uniquely in a distinct flight information region (FIR). The most elementary information transmitted by transponders (Mode A/C) include information about pressure altitude (Downlink Format 4) and identification (or squawk code, Downlink Format 5).

Squawk codes are made of four octal digits; the traditional dials on a transponder read from zero to seven. In normal use, squawk codes are assigned to an aircraft by ATC and subsequently applied by the pilot. Beyond this procedure, there are conventions for squawk codes, which can be selected by aircraft if and when the situation requires or allows it, without permission from ATC. Three such emergency codes are applicable worldwide: 7500 is reserved for hijacking situations and may not be used during training as it triggers a very strict security protocol; 7600 reports a radio failure to the ATC, and 7700 is reserved for general emergencies.

Usually, emergencies are first reported to the ATC over the radio while the crew is assessing the situation or running through checklists (after following the old adage, “aviate, navigate, communicate”). Depending on the nature of the issue, on its gravity and on provided facilities for the airline in different airfields, pilots may choose to divert the aircraft or to continue to their final destination. When diverting, aircraft type and airline procedures may require specific manoeuvres and dumping fuel before landing. Squawking 7700 helps ATC taking the emergency into account in terms of separation, priority and logistics: they may have to get Aircraft Rescue and Fire Fighting (ARFF) ready on the runway if necessary. Finally, there are regular occurrences where transponder squawk codes have been set mistakenly and/or only for a short time. It is important to detect and filter such occurrences.

The wide variety of aircraft emergency situations include technical issues, e.g. landing gear failing to retract, pressurisation issues, cracked windshield; fuel issues, e.g. after holding or several failed attempts to land; and passenger issues, e.g. serious medical trouble or unruly passengers. Collecting and analysing such situations on a larger scale can provide two benefits: First, the insights from the use of transponder codes in real emergencies may assist in future safety analyses of emergency situations. Second, comparative knowledge about the differences between airspaces in their handling of 7700 codes could provide incentives to refine such procedures.

In this paper, we present the first large-scale study about emergency squawks, their causes and consequences. In order to present a comprehensive view, we collect open-source information about worldwide emergency situations over two years (2018 and 2019). Trajectory and squawk information are extracted from the OpenSky Network [2] database and background information related to the situations is found on external sources, including social networks such as Twitter and crowdsourced information related to incidents and accidents (such as The Aviation Herald [3]).

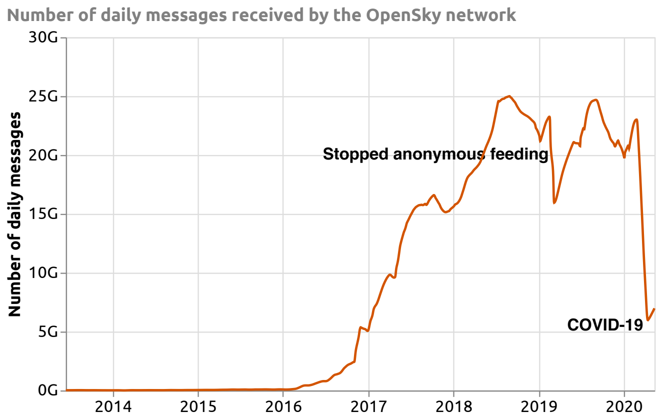

Fig. 1. The growth of OpenSky's dataset over time from 2013 to May 2020

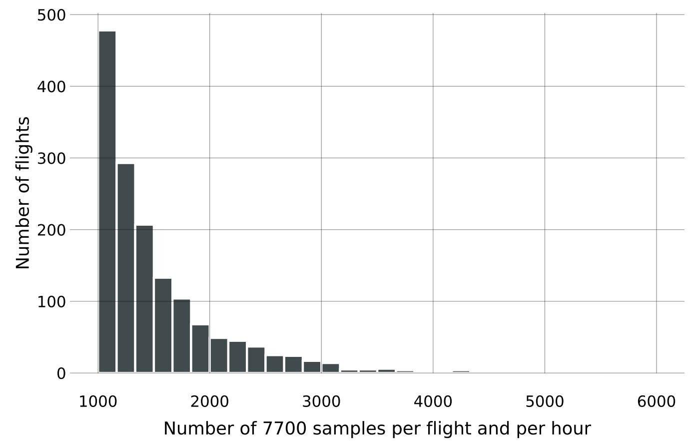

Fig. 2. Distribution of number of 7700 samples per flight per hour

The remainder of this work is organised as follows: Section II describes the overall data collection and preprocessing from various sources of information. Section III provides statistics and observations in the data. Section IV describes common trajectory patterns with real-world emergency situations. This is followed by Section V that with a short overview of findings of the other general emergency transponder codes in the data. Section VI discusses our results and some general points on emergency squawks while Section VII concludes this work.

II. Data Collection and Processing

A. The OpenSky Network

The OpenSky Network is a crowdsourced sensor network collecting air traffic control (ATC) data. Its objective is to make real-world ATC data accessible to the public and to support the development and improvement of ATC technologies and processes. Since 2013, it has continuously been collecting air traffic surveillance data. Unlike commercial flight tracking networks (e.g., Flightradar24 or FlightAware), the OpenSky Network keeps the raw Mode S replies as they are received by the sensors in a large historical database, which can be accessed by researchers and analysts from different areas.

The network started with eight sensors in Switzerland and Germany and has grown to more than 3000 registered receivers at locations all around the world. As of this writing, OpenSky's dataset contains seven years of ATC communication data. While the network initially focused on ADS-B only, it extended its data range to the full Mode S downlink channel in March 2017, which is also the base for this present work. The dataset currently contains more than 22 trillion Mode S replies and, before COVID 19, received more than 20 billion messages per day. Fig. 1 shows the growth and development over the past several years which saw the inclusion of the dump1090 and Radarcape feeding solutions and the integration of non-registered, anonymous receivers, which has been discontinued early 2019 to further ensure the quality of the feeder data. In March 2020, the number of daily messages dropped to about 30% of the previous level, reflecting the curtailment of air travel around the world due to the COVID 19 pandemic. Recovery has been slow so far and it may take years to return to the previous numbers. Besides the payload of each Mode S downlink transmission, OpenSky stores additional metadata. Depending on the receiver hardware, this metadata includes precise timestamps (suitable for multilateration), receiver location, and signal strength. For more information on OpenSky's history, architecture and use cases refer to [4], [5] or visit https://opensky-network.org.

B. Data collection

The OpenSky Network provides a SQL-like Impala-based API to the whole database of collected and decoded trajectory information. As 7700 emergency codes remain exceptional events, we first identified emergency trajectories based on the number of samples in a given flight that contained 7700 squawk codes. A first request to the database collected `icao24` addresses (a 6-character hexadecimal number, 24 bits, named after the International Civil Aviation Organization and uniquely associated to a transponder), callsigns (a 8 character alphanumerical identifier associated to a mission or a commercial route), first and last timestamps, a maximum value of altitude and a number of 7700 messages per hour.

The result of this request shows an exponential distribution (see Figure 2) of the number of 7700 squawk samples per flight. We find that there are many unreliable short alerts; either aircraft sending 7700 for only few seconds or a ground receiver wrongly decoding 7700 for an aircraft. One of the challenges of a large-scale distributed network of ADS-B receivers is to manage to collect enough data around the world: this is difficult to achieve without accepting data of a poorer quality. Before 2019 (see Figure 1), anonymous feeding of preprocessed data included false 7700 alerts from some receivers that we had to later manually remove from our dataset. For a longer discussion of data integrity issues in crowdsourced ADS-B networks, see our OpenSky Report 2018 [6].

Another main reason for false alarms in the data comes from transitions in squawk codes. While pilots are taught not to manually accidentally roll over the emergency codes, some transponder models seem to transition from a squawk code to the next one by not changing digits all at the same time. This also happens when aircraft transition from a squawk code to an emergency 7700 squawk code. For instance, in the dataset associated to this contribution, aircraft `3c9432` (`D-BEAR`) with callsign `AXG1801` started squawking 7700 on Sep. 17, 2019. After squawking 6603, the aircraft very shortly emitted 7703 before switching to 7700:

icao24	timestamp	squawk code
3c9432	2019-09-17 06:25:00Z	6203
3c9432	2019-09-17 06:41:11Z	7703
3c9432	2019-09-17 06:41:14Z	7700

TABLE I
TRANSITION FROM 6203 TO 7700 SQUAWK CODES

After analysing the non-processed flight data, we decided to limit our selection of trajectories to aircraft respecting the following constraints:

1) more than 1000 samples with a 7700 squawk code;
2) a valid `icao24` value (i.e., we remove `000000`, `000001` and other anomalous or wrongly decoded identifiers);
3) a callsign starting with three letters and one number (to remove most non-commercial flights);
4) a valid value of maximum altitude: most messages emitted from the ground do not send any altitude value and we focus here on flown trajectories.

After obtaining this preprocessed dataset of relevant emergency codes, we matched each selected `icao24`, start and end timestamps with samples the `flights_data4` table in the database, which contains one line of information per trajectory, including first and last recorded timestamps for each trajectories. Appropriate requests were written to efficiently download corresponding trajectories between January 1st 2018 and January 31th 2020.

C. Trajectory processing

All trajectories have been requested and preprocessed with the help of the `traffic` Python [7] library, which downloads OpenSky data, converts the data to structures wrapping `pandas` data frames and provides specialised semantics for aircraft trajectories (e.g., intersection, resampling or filtering). In particular, it iterates over trajectories based on contiguous timestamps of data reports from a given `icao24` identifier.

Data selected from the OpenSky database for our emergency dataset yielded 1506 trajectories. In spite of the efforts described in previous section, many trajectories in this first version of the dataset remained irrelevant:

1) no positional (latitude/longitude) information was included for a trajectory due to a lack of ADS-B capable transponder;
2) ground vehicles at airports were associated to a 7700 squawk code (Figure 3);
3) irrelevant legs from the same flight: when an aircraft is associated to a (mostly medical) 7700 emergency, diverts to the closest airport and takes off again after things have been cleared, we need to trim the trajectory down to the leg ending at the diversion airport.

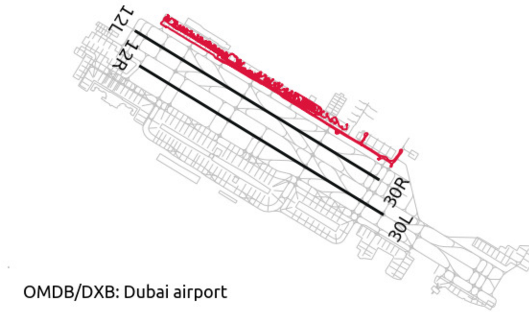

Fig. 3. Ground vehicle (identifier `896ed4`) continuously squawking 7700 on October 12th, 2019 at Dubai airport

Trajectories have been selected, filtered with the help of the `traffic` library grammar and enriched with information from OpenSky aircraft database, including aircraft registration and typecodes, as well as commercial International Air Transport Association (IATA) numbers and a planned route associated to a callsign. Ground vehicles were thereby filtered out with information from the aircraft database.

Then, a set of `traffic` methods has been chained to:

- remove non-positional samples;
- filter out invalid values of altitude;
- filter out invalid positions based on recomputed invalid ground speeds values (too fast often means invalid lat/lon position, too slow could mean that the transponder is still on after the aircraft reached its parking position);
- trigger a new iteration on all trajectories;
- remove flights with no 7700 samples;
- remove flights sending 7700 only in areas covered by a suspicious set of anonymous receivers;
- resample trajectories with one sample per second.

The resulting dataset contains 832 trajectories fully labelled with aircraft information (`icao24` identifier, registration number, typecode), flight information when available (flight commercial IATA number, planned origin and destination airports, detected take-off and landing airports and when applicable diversion airports)

The data presented in this paper is now provided as a generic import in the `traffic` library as `squawk7700`, triggering a download from a corresponding Zenodo repository [8] if the data is not in the cache directory of the user. ADS-B trajectory information are available in a first tabular file, a second one contains metadata information associated to each trajectory.

D. Additional data sources

For further enrichment and a complementary perspective on the cases identified in the OpenSky data, we looked at two publicly available data sources on the web.

- **Twitter:** We conducted searches through the Twitter web page, using callsign, registration, flight number as

	Africa	America	Asia	Europe	Mid. East	Russia
Africa				15		
America		361	6	44	3	1
Asia		1	11		3	
Europe	25	59	2	213	2	5
Mid. East	1	1	3	5	6	
Russia				4		12

TABLE II
NUMBER OF EMERGENCY FLIGHTS PER ORIGIN (LINE) AND DESTINATION (COLUMN) CONTINENT.

identifiers together with the `since` and `until` keywords as defined in its advanced search capabilities, in order to restrict the date ranges of the search, starting with the day of the first emergency squawk to 4 days later: this allows emerging explanations or details to be found.

- **The Aviation Herald:** Relevant articles were identified using date, flight numbers and registration. Data has been manually classified to evaluate the reasons of the emergency, the resulting action (divert, return, continue) and whether fuel was dumped or flights put on hold to burn fuel where possible from the sources.

Whereas Twitter provides fast, often unsourced, information within short text messages provided by of a mix of users, The Aviation Herald focuses on safety-relevant technical issues through detailed and thoroughly sourced reporting. Information provided by Twitter users can comprise first-hand accounts by companies, passengers or crew, but also bots and non-involved enthusiasts speculating wildly with unsubstantiated guesses.

III. Statistics

a) Global coverage: Data collected for this analysis heavily depend on the global coverage of the OpenSky Network of ADS-B receivers. Anonymous feeding has been discontinued in early 2019 in order to improve data quality: this had few impact on the dataset we constituted since anonymous feeders were a serious source of false positive.

Figure 4 compares the OpenSky Network global coverage (above) at the time of submission with the density of emergency trajectories (below). The distribution is consistent with the global coverage, although Northern France and the UK (zoomed in upon in Figure 5) seem to yield a rather high distribution of emergency squawks in the vicinity of Paris airports and along the path to transatlantic routes. This reveals different practices in dealing with emergencies according to countries.

Tables II and III list the most common airlines and aircraft types with a number of emergency situations to be compared with a reference number of flights in one month (January 2019) Bigger absolute values come from the bigger airlines and most short-haul aircraft: European airlines seem to yield a higher ratio, consistent with the higher distribution of 7700 situations in Europe (Figure 5), whereas the ratio for aircraft type seems more constant.

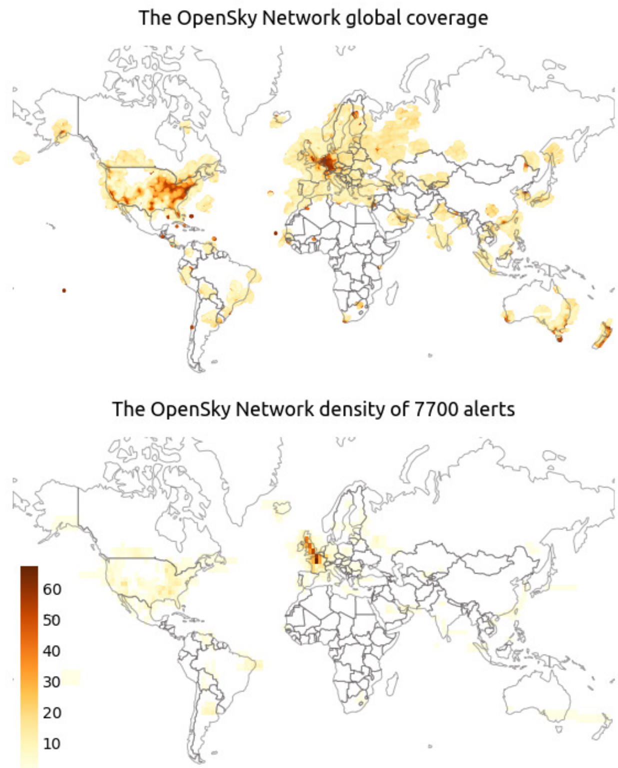

Fig. 4. Global coverage of the OpenSky Network (above) and density of 7700 alerts (below); the scale below refers to a number of unique aircraft. Note that this illustration is dependent on the density of OpenSky's coverage and air traffic in general.

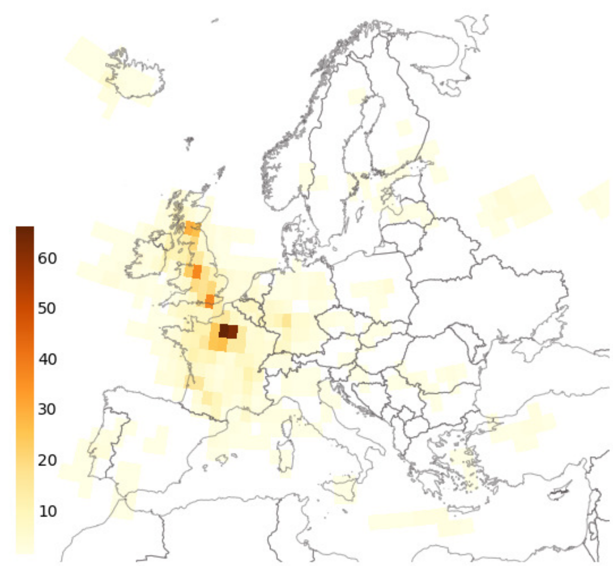

Fig. 5. Density of 7700 alerts in Europe. The clearly identifiable hotspots in the UK and around the Paris airspace are likely indicative of different approaches to handling emergency squawk procedures.

airline	7700	Jan 19	typecode	7700	Jan 19
AAL	47	72490	B738	122	173267
UAL	45	41965	A320	120	151494
EZY	40	577	A319	53	50296
DAL	38	53451	A321	42	59433
AFR	33	5118	B763	36	20156
SWA	30	63015	B737	35	51149
BAW	27	15213	B752	27	17372
RYR	19	1909	B739	27	30765
RPA	16	13585	E75L	27	33098
FDX	16	14535	B772	22	14460

TABLE III
TOP 10 BY AIRLINE AND AIRCRAFT TYPE

b) Causes of emergency: As specified above, additional data sources were used to evaluate known causes for an emergency, namely Twitter and reports from The Aviation Herald. In our methodology, we build and expand on earlier work in this area of data fusion [9].

In our results, we can broadly differentiate *technical*, *environmental* and *medical* causes. Table IV summarises the findings from these two sources including some subcategories in case a more detailed description was found. The subcategories are not always clear or unique, so should be considered with care.

We found *Twitter* messages for 419 cases and reports in *The Aviation Herald* for 90 cases of the total number of 832 cases identified in the OpenSky data.

Interesting to note is the high number of cases leaving traces in Twitter messages (419 out of 832 cases), which is, of course, to some extent, explained by some automated bots like the Twitter user `@OpenSkyAlerts` itself. That we do not find Twitter messages from these bots for *all* cases seems to be a limitation imposed by our way of accessing Twitter messages through their search mechanism. On the other hand, many other accounts add manually information related to emergency situations (for a more detailed discussion see Section III-0d).

The focus of The Aviation Herald is explicitly on safety-related issues, therefore in general medical emergencies are not covered. The two corresponding reports are indeed cases of medical problems of crew members and therefore included in its coverage.

c) Measures taken after emergency: Apart from extreme events (e.g. broken windshield, bird strike) requiring immediate action, when an incident occurs, pilots or crew first take time to assess the situation in the cockpit (technical issue) or in the cabin (smoke, passengers with an unruly behaviour or seeking immediate medical attention). Pilots may engage in holding patterns to run through checklists. The 7700 code is set upon requests from ATC, which reveals different practices according to countries. Subsequently, pilots coordinate with ATC and take a decision: follow their route or divert.

We detected 295 diverted flights in the dataset, 111 of them returned to their airport of origin. The choice of the airport of diversion usually depends on the nature of the emergency: when aircraft can divert to major hubs for the airline, it is easier to accommodate passengers on new flights, and to work through shift limits for the crew. In case of medical emergencies, aircraft usually resume their route after the passenger and their family have been disembarked and taken care of. When the aircraft is diverted to an inconvenient location, the airline may need to send a replacement aircraft with crew to the airport of diversion.

Category	Detailed Cause	Twitter	The Aviation Herald
Technical	Unspecified technical	34	-
	Engine	18	19
	Fire/Smoke/Smell	17	16
	Cabin Pressure	12	19
	Hydraulics	5	5
	Cracked window	5	3
	Landing gear	5	2
	Fuel related	4	6
	Instrumentation	-	3
	Flaps	-	1
	Heating	-	1
	Tires	1	1
	Hot air leak	1	1
	Maintenance	-	1
	Brakes	-	1
	Miscellaneous	-	6
	Slats	-	1
	Door	1	-
	Subtotal	103	86
Medical		72	2
External	Bird strike	4	1
	Weather related	3	1
	Bomb threat	2	-
	Operational issue	1	-
Unknown		234	-
Total		419	90

TABLE IV
MOST COMMON CAUSES FOR EMERGENCY AS OBSERVED IN TWITTER DATA AND THE AVIATION HERALD

`LFPG`	Paris CDG	France	27
`EGLL`	London Heathrow	United Kingdom	12
`LFPO`	Paris Orly	France	10
`EGPF`	Glasgow	United Kingdom	10
`LFBD`	Bordeaux	France	9
`EGPH`	Edinburgh	United Kingdom	7
`EDDF`	Frankfurt	Germany	7
`EIDW`	Dublin	Ireland	7
`LFLL`	Lyon	France	6
`EGKK`	London Gatwick	United Kingdom	5
`LFML`	Marseille	France	5
`LFRS`	Nantes	France	5
`EGCC`	Manchester	United Kingdom	5
`KPHL`	Philadelphia	United States	4

TABLE V
MOST COMMON 7700 DIVERSION AIRPORTS.

Table V plots the most common diversion airports in the data set. Most airports in the dataset are French or British airports: the first American airport is in 14th position in the list. This can be partly explained by different 7700 practices according to the countries.

d) Twitter analysis: For a large part of the emergency cases (419 out of 832) we find related messages in Twitter. Comparing OpenSky data and Twitter messages, we evaluate the delay between first observed 7700 code and the first related Twitter message for each case. For each account we therefore can generate a distribution of delays across the different cases mentioned by that account. Figure 6 highlights Twitter accounts that have sent messages related to at least 14 cases in our observations. Each account is placed according to the median of the distribution of the delays (x-axis) and the number of our observed cases, for which we see messages by that account (y-axis). The spread of the delays is further indicated by the horizontal error bar spreading from first to third quartile of the distribution. Major language for each account is indicated by a colour.

Overall the largest coverage of our observed emergency cases is by the `@OpenSkyAlerts` account, not very surprising as the observed emergency cases have been derived from the OpenSky data itself used for this analysis and this account is a bot maintained by the OpenSky Network. The delay is very consistently close to 5 minutes, as the Twitter bot is using this as the delay to avoid sending spurious alert messages. Looking at the respective messages by an account, we can clearly see a bot-like behaviour with a fixed message structure for some accounts (like `@OpenSkyAlerts`, `@SquawkAlert`). These have also nearly exactly one message per case, though the relative timing might vary. Other accounts, like `@planefinder`, `@FlightIntl`, `@airlivenet`, also have a bot-like behaviour looking at their messages, but provide in addition to the initial message also updates where additional information (place of diversion, reason for the emergency) is broadcast.

But there also seem to be many individuals interested in aviation (respectively aviation emergencies) that have a high coverage of cases with very short delay times and a high number of informative update messages, most specifically the account `@NikPhillips666`.

Nearly all accounts shown here sent messages covering cases through well above one year of our two-year observation period, most of the frequent messengers cover the full observation period, showing a consistent long-term interest of the users in the topic.

Still, although social media in many cases can provide quick and accurate additional information, it is still severely limited as the data is very noisy and the role of the messenger and therefore the reliability of the message is often unclear (e.g., an airline community manager for an airline, a direct observer sitting in the aircraft or an outsider stating a guess for the reason of the emergency).

IV. Analysis of Common Emergency Patterns

a) Medical issues: When a passenger or a crew member is unwell and their medical situation is life-threatening, the pilot will take the decision to divert to the closest airport with proper medical facilities. The aircraft will resume its route after the person has been taken care of. The dataset contains 72 situations labelled as medical emergencies; most information is provided by Twitter accounts.

Practices about the use of the 7700 emergency code differ according to countries, only one medical issue in the dataset

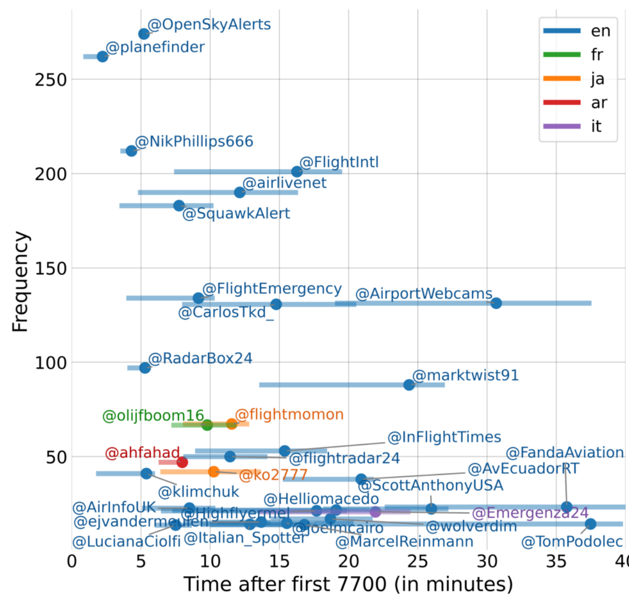

Fig. 6. Twitter accounts for which we have seen messages for at least 14 of our observed cases. The x-axis represents the median of the time delay between the first observed transponder code 7700 and the first observed tweet by that user for the cases. The horizontal error bar indicates the spread of delay times by spanning from first to third quartile of the distribution. The y-axis is the count of cases for which we have observed messages for the account. Colour highlights the main language of the tweets by that account.

diverted to an American airport (Dallas Fort-Worth `KDFW`). European ATC will often require aircraft to squawk 7700 for a priority landing, they will also get ARFF or ambulances ready at the airport. Twitter activities of popular accounts reveal many *non emergency* medical diversions to Gander `CYQX` or Saint-John `CYYT` airports on Newfoundland, Canada, before or after a transatlantic crossing. Figure 7 plots a situation when one of the pilots felt unwell. Passengers had to wait for several hours at the diversion airport before another pilot was able to reach Rennes `LFRN`.

b) Cabin pressure issues: During a flight, commercial aircraft are usually pressurised at an equivalent altitude of 10,000 ft (slightly more than 3,000 m). Issues with the pressurisation system or cracked windshields in the cockpit will most often lead to a steep descent to 10,000 ft (at up to $-6,000$ ft/min), which is illustrated in Figure 8. Then, flights are not necessarily diverted as it may become safe to fly at this altitude with a defunct pressurisation system. The dataset contains 31 situations labelled as cabin pressure or cracked windscreen issues, 27 of which leading to a diversion.

Regarding the situation shown in Figure 8, it is possible that most passengers did not notice any emergency until landing. The Aviation Herald mentioned the issue but only bots reacted to this event on Twitter.

c) Fuel dumping: Long-haul flights embark large amounts of fuel before take-off. Emergency landings may be performed at any time but a heavy landing provides a risk of significant damage to the structure of the aircraft. In order to avoid immobilising the aircraft for a mandatory maintenance

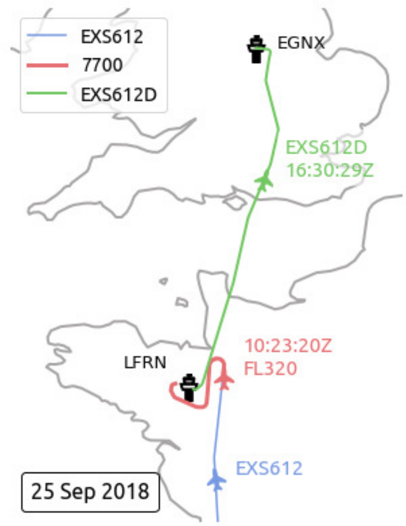

Fig. 7. On September 25th, 2018, Jet2.com flight `EXS612/LS612` from Girona `LEGE` to East Midlands `EGNX` diverted to Rennes `LFRN` after one of the pilots felt unwell. Another pilot was flown to Rennes and the same aircraft took off again five hours later with callsign `EXS612D`.

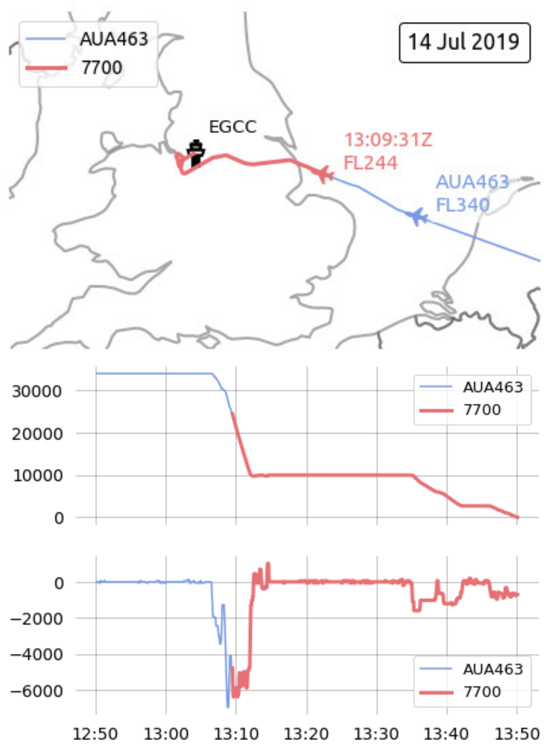

Fig. 8. On July 14th, 2019, Austrian Airlines flight `AUA463/OZ463` from Vienna `LOWW` to Manchester `EGCC` started an early abrupt descent (down to -6,000ft/min) before starting to squawk the emergency code. Once at 10,000ft, the aircraft was considered safe to fly to its final destination without the need to divert. The same aircraft flew back to Vienna two days later, after a probable maintenance in Manchester.

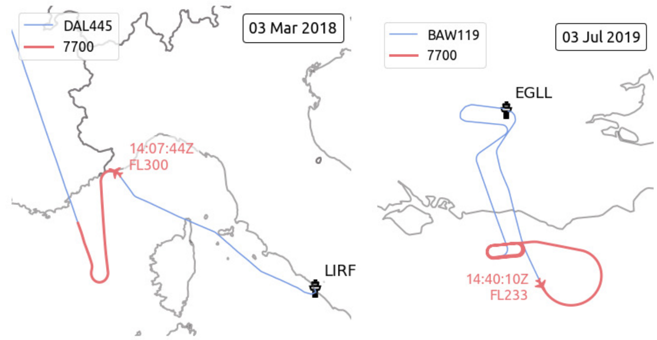

Fig. 9. On March 3rd, 2018, Delta Airlines flight `DAL445/DL445` from Rome `LIRF` to New York `KJFK` detected an instrument issue when reaching the vicinity of Nice. They turned over the Mediterranean sea to jettison fuel before diverting to Charles-de-Gaulle `LFPG`. On July 3rd, 2019, British Airways flight `BAW119/BA119` bound for Bangalore `VOBL` interrupted their climb before FL240 and entered into a holding pattern over the Channel to jettison fuel before requesting a priority landing at Heathrow `EGLL`.

after a heavy landing, it is common practice to jettison fuel under close monitoring with the ATC until the aircraft is light enough to land safely. Fuel dumped above 6,000 ft will dissipate before reaching the ground.

Figure 9 shows two common patterns for fuel dumping: hippodrome-shaped holding patterns or a long detour above uninhabited areas. In the dataset, 32 aircraft were found to dump fuel according to external sources, among them 23 aircraft engaged in several holding patterns.

d) Diversions due to weather: Poor weather conditions are a common source of disruption in airline operations. Even if the airport is still open to landing in degraded mode, procedures may require a diversion after two failed attempts at landing at the same field. As aircraft divert to neighbouring airports, some of them could become short of fuel and ask for priority landing. Squawking 7700 can assist ATC in scheduling landing operations with this factor in mind.

Figure 10 plots two such situations. On January 14th 2020, Storm Brendan forced airlines bound to London to divert: crosswind gales of up to 80mph caused major disruption. Flight `BAW79J/BA2777` entered a holding pattern before diverting to Birmingham `EGBB`. The 7700 emergency code was triggered after the decision to divert. On January 20th 2019, severe weather conditions (wind, snow and ice) at Boston Logan airport `KBOS` cancelled more than 10% of the scheduled flights. After `UAL342/UA342` entered a holding pattern, they took the decision to divert before any landing attempt. They triggered the 7700 emergency code later: the hypothesis of low fuel is reasonable and is labelled accordingly based on analysis expressed on Twitter.

e) Major technical issues: Serious technical issues may lead to a diversion to the closest airport, not necessarily a major hub for the airline. When the aircraft is not fit for taking off again, solutions to accommodate passengers must be taken: they may be rebooked on different flights, driven by bus to a major airport or a replacement aircraft may be reserved or dispatched to fly people forward to their intended destination.

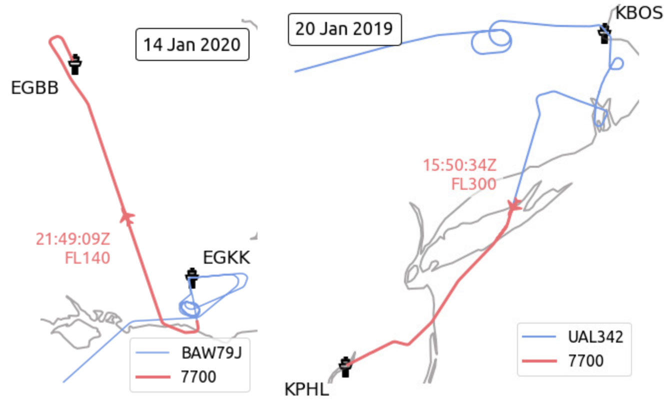

Fig. 10. British Airways BAW79J/BA2777 from Jersey EGJJ to London Gatwick EGKK diverted to Birmingham EGBB after two failed attempts at landing in gusty crosswinds of Storm Brendan in early January 2020. On January 20th, 2019, many flights to and from Boston Logan airport KBOS were cancelled because of severe wind, snow and ice conditions. United Airlines UAL342/UA342 from Chicago KORD was diverted to Philadelphia KPHL without any landing attempt.

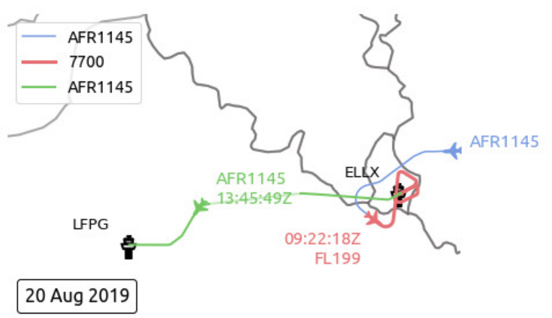

Fig. 11. On August 20th, 2019, Air France AFR1145/AF1145 from Moscow Sheremetyevo UUEE to Paris CDG/LFPG diverted to Luxembourg ELLX after an issue with the breaking system. The aircraft did a first low approach, probably asking the tower for a visual inspection, before landing. The same aircraft resumed its flight to Paris few hours later.

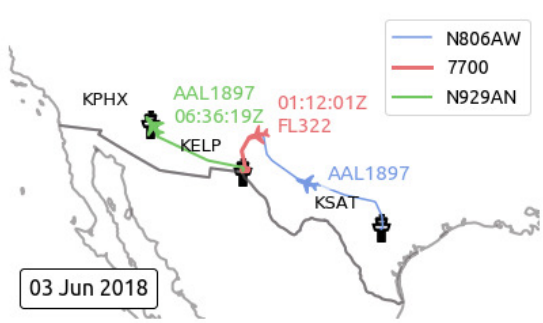

Fig. 12. On June 3rd, 2018, American Airlines AAL1887/AL1887 from San Antonio KSAT to Phoenix KPHX diverted to El Paso KELP after experiencing turbulence and major damages to the aircraft. Few hours later, a replacement aircraft flew passengers to their final destination.

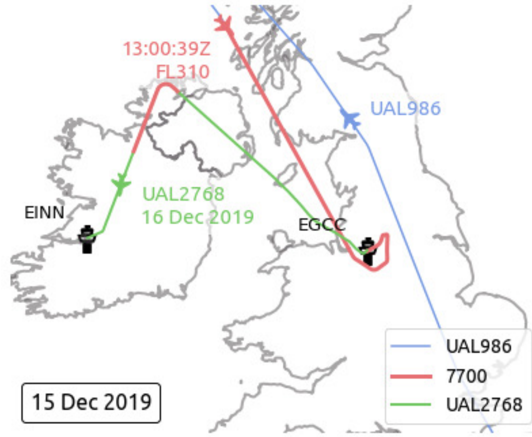

Fig. 13. On December 15th, 2019, United Airlines UAL986/UA986 from Paris CDG/LFPG to Chicago KORD diverted to Manchester EGCC after a suspicion of fuel leak. On the next day, as the aircraft was about the be ferried back to Chicago, it diverted to Shannon EINN with the same issue. It was finally sent back to Washington KIAD on December 19th.

Illustrations include an emergency landing in Luxembourg for brake issues: the same aircraft took off later and returned to Paris with passengers on board (Figure 11); an emergency landing in El Paso after an aircraft experienced severe turbulence with serious damages on the windshield and nose: a replacement aircraft was used to take passengers to their final destination (Figure 12). Figure 13 displays an aircraft diverting to Manchester upon suspicion of a fuel leak: passengers were taken by bus to Heathrow. The next day, the same aircraft was ferried back to Chicago without passengers but had to divert again to Shannon airport.

f) Atypical approaches: Many emergency situations, technical or medical, do not lead to a diversion. The 7700 squawk code triggers a particular way of displaying the aircraft on all radar screens, even for controllers in charge of different sectors in the same ACC. Displaying 7700 may help ATC assisting the aircraft in providing ARFF on the runway or making them fly unbeaten tracks for a swift priority landing.

Figure 14 displays the trajectory of AFR48FC between Toulouse and Paris–Orly. All trajectories flying the same callsign in the same month are plotted in the background for reference: they mostly match the STAR procedure through ODILO (IAF). AFR48FC had a technical issue and squawked 7700 as they bypassed the usual procedure, descended with a very steep profile (several segments at less than -3,000 ft/min) for an emergency landing. Only bots reacted to the event on Twitter with no relevant information and it is reasonable to think most passengers did not notice anything unusual apart from the discomfort of a steep descent.

g) Non-commercial flights: The Airbus A300-600ST (Beluga) transports oversized aircraft components (like wings) between Airbus factories in Europe, including Toulouse, Saint-Nazaire (France), Hamburg (Germany) or Seville (Spain). On March 5th, 2018, one such aircraft bound for Saint-Nazaire LFRZ turned back to Toulouse during climb and aligned on

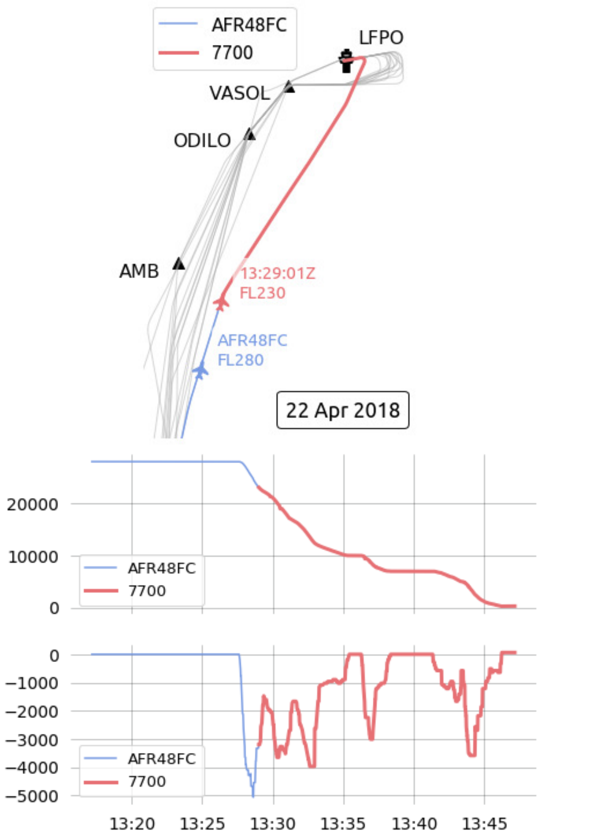

Fig. 14. On April 22nd 2018, Air France flight `AFR48FC/AF6127` from Toulouse `LFBO` to Paris Orly `LFPO` squawked 7700 as they left the usual STAR procedure route (`AMB ODILO VASOL`) for a steep descent and a circle to land manoeuvre to landing.

runway 14R before heading aside, probably investigating a possible issue. Only after getting out of the two holding patterns did the Beluga start squawking 7700, probably to remind tower controllers they might need a particular attention, shall they need extra time on the runway or in case of a missed approach.

On July 15th 2019, private aircraft `F-GJFE` covered the Tour de France to relay audio and video signals from helicopters flying below to cover the 10th stage of the Grande Boucle between Saint-Flour and Albi. Few minutes after Wout Van Aert aced the sprint to stage victory in Albi, the aircraft headed west with an emergency squawk code before turning back to park in Castres `LFCK`. The nature of the issue remains unclear. There would have been time for maintenance on July 16th if needed, as it was a scheduled rest day. The aircraft was back in service on the Tour on July 17th between Albi and Toulouse.

Non-commercial flights are not very well covered in social networks as the general public is not much impacted. Naturally, emergency situations occur, too, but it is more difficult to get relevant contextual information from public sources.

V. Other Transponder Codes

While this work focuses on the 7700 emergency squawks, we briefly discuss the other two universal transponder codes,

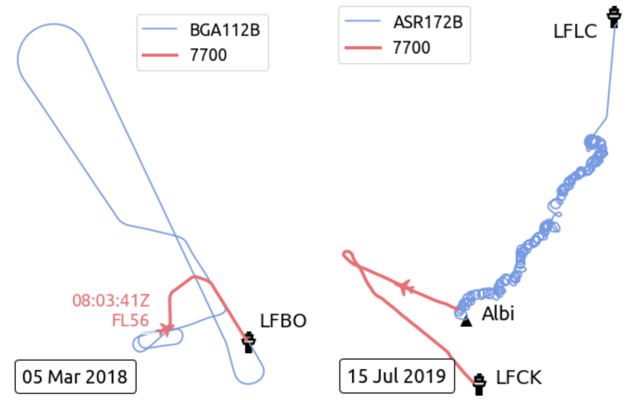

Fig. 15. Airbus Beluga flight `BGA112B` returned to Toulouse with a possible technical issue on March 5th, 2018. On July 15th 2019, private aircraft `F-GJFE` with callsign `ASR172B` covered the Tour de France between Saint-Flour and Albi to relay audio and video signals for television.

7500 and 7600 as they were seen, pre-filtered and stored by OpenSky's alert service. [10] We analysed a four-month period between 17 January 2020 and 17 May 2020.

A. 7500

Over the four months, four hijack codes of significant length were reported, none of which could be suspected as a relevant case. All were still relatively short lived (the median time was around 5 minutes); indicating a mistake by the pilot(s). For example, in one of the cases the aircraft eventually switched to 7600 on approach.

Taking a closer look at the raw flight data for the same period, over 8000 distinct aircraft reporting squawk 7500 were filtered out, as their squawks were only sent a single time or for a few seconds. In some cases, the short sequences seem to be glitches in systems or in transmission, e.g., a stream of code 1000 followed by a short sequence (2-3 seconds) of code 7500, followed again by code 1000. Some of these sequences are observed from state-of-the-art, new aircraft.

In further examining squawk 7500 incident tweets informally sampled using Google search, we found 29 tweets reporting hijack codes between 2017 and 2020. Out of these, 17 were issued from USAF aircraft, and none were related to an actual incident.

It is crucial to note that such errors with regards to 7500 can cause major disruption and are not to be taken lightly. For example, Amsterdam Schiphol airport made the news on November 6th 2019 when an Air Europa Airbus A300 `EC-LQP` accidentally started squawking the hijack code 7500. The aircraft was parked at the gate (outside the ground coverage of the OpenSky Network). Security procedures seriously disrupted the airport operations and parts of the terminal were closed to the public for several hours.

B. 7600

Over the four-month period, 70 radio failures were reported. While we cannot assume this period to be representative due to the effect of the COVID-19 pandemic on aviation, in particular commercial and scheduled flights, we identified and categorised these occurrences. 5 (7%) were military planes, 10 (14%) commercial airliners, and the remaining 55 (79%) smaller private aircraft, mostly Cessnas or Pipers. Just over half (37 of 70 or 53%) occurred during approach with the 7600 code set during landing. The mean (median) time that the transponder code was set was 22 (11) minutes, respectively.

VI. Discussion

As known within the aviation community and confirmed by our data analysis, a large majority of 7700 do not point towards safety-relevant events but are short-lived errors, which are possible at several levels (transponder setting, message decoding). Of those that can be assumed to be correctly set and decoded, many will still not produce a safety-related incident report or a noticeable diversion from planned flight plan or operating procedures. In case of a medical or unruly passenger issue or similar cases where the aircraft works normally, information provided by airline or third parties may shed some light on the reason as shown in our analysis in this paper.

While squawking 7700 in case of an emergency is not a requirement, it can be helpful for pilots in particular when in non-domestic airspaces. It is a proven way to get the undivided attention and priority by all ATC facilities in the area, even if a mayday – pan-pan call via radio is not an option [11]. As such, the 7700 squawk remains an important tool for ATC as any such aircraft is displayed prominently on all screens of the concerned Area Control Centers (ACC). As an alternative and complement to radio calls, it helps to clearly visualise that there is an important issue [12].

As using 7700 is not a requirement or always feasible, it should also be noted that there are many true emergencies where the transponder code is not set. Analysing this number was out of the scope of this work but it is estimated to be significant.

With regards to comparative analysis, it is important to consider the differences between airspaces and countries. For example, there are ATC facilities, which assign 7700 squawks to aircraft with any type of anomaly, e.g. routing problems, paperwork issues, or minor technical problems.

Besides the reason for applying emergency squawks, the timings can also differ. For example, we observe aircraft squawking 7700 when doing specific patterns related to the emergency and switch it off after the decision is taken (diversion for instance); some others investigate first and then switch the squawk to 7700 until landing.

With the high-profile nature of aviation accidents, there are regular cases of users and news reports jumping prematurely on reports of squawk codes instantly available through ADS-B networks, despite the low information value provided by 7700 squawk occurrences, as outlined in this paper. Taking all discussed issues together, we hope that our work provides some accessible background information that can counteract the fast-paced internet news and reporting cycle on aviation accidents and prevent at least some premature speculation leading to potential anxiety and loss of trust.

VII. Conclusion

In this work, we used the data collected by the OpenSky Network over several years to analyse the occurrence and impact of emergency transponder codes, the well-known squawk 7700. While our results indicate that analysing the three universal squawks still requires significant manual work and some of the methods and gained insights are potentially airspace-specific, we showed that many relevant case studies can be obtained with our approach. By filtering and enriching the data with external data sources, we created a dataset that offers ground truth and thus opportunities for teaching, for ATC and pilot training using realistic emergency situation references and for further research. Eventually, it may also provide insights into the possible harmonisation of the handling of transponder codes both within and across airspaces.

Acknowledgements

The authors would like to give special thanks to Simon Hradecky, the driving force behind The Aviation Herald, for his helpful comments and insights.

References

[1] M. Schäfer, X. Olive, M. Strohmeier, M. Smith, I. Martinovic, and V. Lenders, "OpenSky Report 2019: Analysing TCAS in the Real World using Big Data," in *2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC)*. IEEE, 2019, pp. 1–10.

[2] M. Schäfer, M. Strohmeier, M. Smith, M. Fuchs, R. Pinheiro, V. Lenders, and I. Martinovic, "OpenSky Report 2016: Facts and figures on SSR mode S and ADS-B usage," in *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. IEEE, 2016, pp. 1–9.

[3] "The Aviation Herald," accessed 2020-05-20. [Online]. Available: https://www.avherald.com/

[4] 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 IEEE/ACM International Symposium on Information Processing in Sensor Networks (IPSN)*, Apr. 2014.

[5] M. Strohmeier, M. Schäfer, M. Fuchs, V. Lenders, and I. Martinovic, "OpenSky: A Swiss Army Knife for Air Traffic Security Research," in *Proceedings of the 34th IEEE/AIAA Digital Avionics Systems Conference (DASC)*, Sep. 2015.

[6] M. Schäfer, M. Strohmeier, M. Smith, M. Fuchs, V. Lenders, and I. Martinovic, "OpenSky report 2018: assessing the integrity of crowdsourced mode S and ADS-B data," in *2018 IEEE/AIAA 37th Digital Avionics Systems Conference (DASC)*. IEEE, 2018, pp. 1–9.

[7] X. Olive, "traffic, a toolbox for processing and analysing air traffic data," *Journal of Open Source Software*, vol. 4, no. 39, p. 1518, Jul. 2019. [Online]. Available: http://joss.theoj.org/papers/10.21105/joss.01518

[8] X. Olive, A. Tanner, M. Strohmeier, M. Schäfer, M. Feridun, A. Tart, I. Martinovic, and V. Lenders, "Reference datasets for in-flight emergency situations," Jul 2020. [Online]. Available: https://zenodo.org/record/3937483

[9] A. Tanner, A. Blarer, and V. Lenders, "Fusing flight data with social media data," in *NATO Science and Technology Organization: Big Data and Artificial Intelligence for Military Decision Making*, may 2018.

[10] The OpenSky Network, Online, accessed 2020-05-20. [Online]. Available: https://opensky-network.org/network/alerts

[11] I. Petchenik, "Squawking 7700—In-flight Emergencies from a Pilot's Perspective," Online, Oct. 2015, accessed 2020-05-20. [Online]. Available: https://www.flightradar24.com/blog/squawking-7700-in-flight-emergencies-from-a-pilots-perspective/

[12] ——, "Squawking 7700 — An Air Traffic Controller's Perspective on In-flight Emergencies," Online, Oct. 2015, accessed 2020-05-20. [Online]. Available: https://www.flightradar24.com/blog/squawking-7700-an-air-traffic-controllers-perspective-on-in-flight-emergencies/