A Flight Delay Reporting and Analysis Platform Through Secure Information Sharing

Guney Guner*, Baris Baspinar*, Emre Koyuncu*, Gokhan Inalhan*, Massimiliano Zanin‡, Vaishali Mirchandani†, Alberto Enrich Gonzalez†, Julio César Triana Castillo†, Emilio Álvarez Pereira† and Cengiz Paşaoğlu§
*Istanbul Technical University, Istanbul, Turkey
‡Innaxis Foundation & Research Institute, Madrid, Spain
†Team & Cloud, Madrid, Spain
§Devlet Hava Meydanlari İşletmesi, Istanbul, Turkey

***Abstract*—Considering the future needs of the seamless flow of information between the stakeholders, applying secure information sharing and calculation that allows untrusted parties to perform computation over a data set will be a delicate issue for air transportation implementations. In this work, we have developed a secure system that is specific to delay report gathering from different stakeholders and analysing based on secure multi-party computation. Considering the needs of a secure reporting system, we have developed a web-based portal enabling participants to manage their contributions. To demonstrate the feasibility of such reporting and secure analysis, we have utilized real-world examples through the historical data analysis. Moreover, we have also performed the cost-benefit analysis through the computational effort assessing of such SMC-based delay analysis solution for the realistic operational environment, provided results of these analyses, and given detailed discussed on them.**

I. Introduction

There is a global interest in US and Europe to transform the current information handling in air transport system into an highly efficient and secure system through SESAR in Europe, NextGen in USA, OneSky in Australia, SIRIUS in Brazil, or CARATS in Japan. The common idea behind of these programs is: efficiency can be improved only by ensuring a continuous flow of information between the agents and stakeholders involved in the operation. While some examples involving tactical information sharing such as flight intent exchange or price negotiations for slot exchange by airlines; some needs for strategical improvements, e.g. analysis of past incidents, thus of historical operational data.

Achieving such seamless flow of information entails two important and contradictory challenges. First, most ATM data are considered confidential and sensitive and, hence, private - both for their commercial value, and for the political or social consequences some of the analyses may cause; any solution should thus guarantee an adequate level of confidentiality. Second, at the same time, data should be stored and processed in a safe and efficient way, which usually implies the use of a cloud-based infrastructure. This may generate security problems, as the exact location of data in the cloud is generally not known. SESAR's System Wide Information Management (SWIM) [12], only partially tackle these two problems. Specifically, SWIM is based on a public-key infrastructure, allowing users to only access those sets of data included in their authorization class. Data are released to the parties requiring them, hence the security of the system is as good as the security of the worst procedure implemented by the entities. As a result, the idea behind this paradigm is trust between the users and system managers.

TABLE I: Standard computation vs. SMC-based computation

Standard Computation	SMC
Parties want to collaborate to perform a computation, and trust each other. Inputs and outputs are publicly shared.	Parties do not trust each other, and one partys inputs should not be known by the others.
The computation is performed in a single location.	The information is divided in shares, and the computation is performed at different locations.
As inputs and outputs are public, any party can check the correctness of the computation.	Inputs are private, and thus a party may try to distort the result of the computation. Methods should be put in place to prevent this possibility.
The cost of the computation is only due to the operations performed.	Additional computational costs can appear, due to share manipulation and communication.
Any computation can be performed (in the Turing sense).	SMC is currently limited to linear arithmetic and Boolean circuits [7].

Secure Multi-party Computation (SMC) is a set of techniques and algorithms that allows two or more untrusted parties to perform some kind of computation over a data set. The basic principle behind this is that the input information is divided into a number of shares, which are transmitted to different computation server. While each share is not enough to recover the initial information, protocols can be designed to perform operations over them; at the end, the result is collectively calculated by the computation servers, while no one of them has enough information on its own to recover any input. This allows, once the computation is over, to recover the output of that computation, without any additional knowledge on the information provided by the parties. In other words, instead of providing any party with the full data set (and thus creating a security issue to be managed) or denying the access to it (effectively blocking any possibility of using the data), the data owners could allow third parties to run computations on encrypted information, without real access

to the full dataset. Table I provides a summary of differences between a standard and an SMC-based computation. Secure computation has hitherto been used to solve several real-world problems, from secure sealed-bid auction [5], elections with an electronic voting scheme [16], benchmarking [3], up to defense applications in military operations [13].

In this work, we have developed a secure system for delay report gathering from different stakeholders and analysing utilizing Secure Multi-party Computation (SMC). Such reporting system enabling to collect reports from different stakeholders allows further investigating the causes of delay that sometimes it is not easy to judge by looking in which phase delay occurs. Considering the needs of a secure reporting system, we have also developed a web-based delay analysis portal enabling participants to see the open questionnaires and put their inputs. In order to conceptually demonstrate the feasibility of such reporting and secure analysis, we have selected many real-world examples through the historical data analysis, exemplifying the kind of use one can make of the delay analysis portal. It is envisioned that in addition to providing secure calculations through the SMC tools, providing frequent delay inquiries could potentially improve assessing the causes. Therefore, to address the cost-benefit, we have also performed analyses by the means of computational effort assessing of SMC-based delay analysis solution for the realistic operational environment. The results of these analyses are provided and discussed in details.

The paper is organized as follows: first, a brief explanation for Secure Multi-Party Computation is given in Section II. Delay Report Analysis Portal including data frames for different stakeholders are introduced in Section III. Section IV explains real-world data selection and simulation results. Finally, Section V provides conclusions and remarks.

II. Brief Secure Multi-Party Computation

In the last decade, the increasingly use of cooperative computation, as well as the new ways of decentralized and distributed computing, *i.e.* peer to peer networks and cloud computing, has fostered the need of such technology, in order to solve problems in which many parties need to provide inputs for a computation, however, no mutual trust can be ensured. Some examples include secure decentralized elections [14], [4], secure auctions [1], secure benchmarking or retrieval of private information, *e.g.* biomedical records [10].

Different approaches, or *primitives*, have been used to implement SMC protocols for different applications. Independently on the problem to be solved, *e.g.* ranking, auction or set intersection problems, the protocol has to be constructed by means of a combination of these primitives, being therefore the building blocks of any SMC solution. The four combinations that have by and large been used in real-world applications are *Secret Sharing* [15], [2], *Oblivious Transfers* [14], *Garbled Circuits* and *Homomorphic Encryption* [9].

As its name suggests, Secret Sharing is a set of techniques aimed at distributing a secret, *i.e.* private information that should be concealed, among a group of participants, each one of them receiving just one piece of the secret. The secret can then be reconstructed only when a sufficient number of participants work collaboratively, as individual shares are of no use on their own. For instance, suppose that one is to encode the secret, in this case a binary number s, among different parties. To all (except one) parties, the user would send a random number p_i, while the last would receive the result of $s \oplus p_1 \oplus p_2 \oplus \ldots \oplus p_{n-1}$, $\oplus$ being the bitwise exclusive OR (XOR) operation. To recover the secret, all parties should collaborate, and calculate the bitwise XOR of all parties numbers ps. Suppose next that all parties want to perform a Boolean operation on private numbers they own. Following the previous example, each one of them can divide and share its number through a set of shares ps; afterwards, all parties execute the Boolean operation on the shares they have, and finally they collaboratively retrieve the final results.

Following section explains the delay report portal and reference data frame for different level of users.

III. Analysis of Delay Report Portal

Analyzing of Delay Reports through the secure multi-party computation (SMC) aims to build a system for delay reporting using cleared information coming from different participants, securely merged in order to achieve additional knowledge about causes of delays. Here, cleared information refers to delay information whose causes and amounts have already been processed by the stakeholders. All stakeholders are considered semi-honest parties (honest but curious): the stakeholders will always follow the protocol correctly and will always send well-formed messages, thus not affecting the outcome of the computation. However, if the possibility is offered, they will try to learn any private information by examining all the data they can get. Figure 1 depicts the delay report analysis structure with SMC computation.

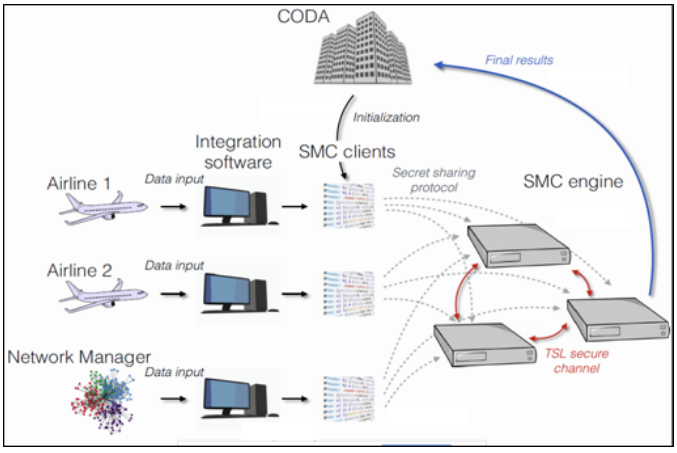


Fig. 1: Delay Report Analysing with secure computation

Through this reporting structure, several stakeholders collaborate by introducing delay information inside the system: pilots, airline representatives, controllers, network manager, airport representatives, and handling organizations. In validation study, participants has been categorized in two groups; the

first type represents participants that can provide information about one cause of delays for all flights considered in a given scenario, such as the Network Managers (i.e. NM or CFMU officer), since they have access to delay data related to network regulations for all flights across Europe. The second type of participants have information about a limited set of flights: such as airlines, which have visibility only on their operations. The following four statistical calculation library based on SMC are built to perform delay analysis:

- Statistical indicators related to the delay minutes of all flights in a route during a specific time window (mean, median, standard deviation).
- Airlines ranking by means of total delay minutes in a route during a specific time window.
- Statistical indicators related to the delay minutes of each cause in a route during a specific time window (mean, median, standard deviation, maximum delay time per cause).
- Statistical indicators related to the delay minutes of all the flights in a route during a specific time window (mean, median, standard deviation), excluding each airlines flight with the highest delay and thus excluding extreme values.

Following subsection gives the real-word data selection from different delay causes for the simulation purposes.

A. Data Provision for Delay Analysis and Simulations

In order to provide the validation of the delay analysis portal through the realistic events, we have utilized following datasets:

- ALLFT+: historical flight data from the DDR2 repository (EUROCONTROL);
- Capacity reports: reported airport capacities from the DDR2 repository (EUROCONTROL);
- NOP Network Operation Portal: Air traffic network headline news from Europe (provides some news for the events causing delay);
- Other online weather reporting portals providing METAR Data.

By considering data availability, only a limited number of delay causes could be observed (or estimated). The current practice in the airline industry is to adopt the International Air Transport Association (IATA) delay coding system, a standard recommended in the Airport Handling Manual (AHM 730) published by IATA [8]. Therefore, the following subspace (with corresponding IATA codes) of the Standard IATA Delay Codes is considered in the simulations by considering available datasets given above.

Table II includes 13 delay causes that are measurable for Network Managers (NM) and to Airlines (AOC). The stakeholders might have also information about the other parties (indirectly related), but we suppose that they are most likely not measurable and the information is only at speculation level.

TABLE II: Observable delay causes through the available data library

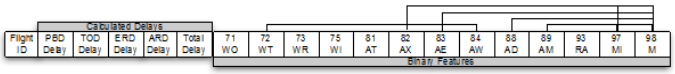
Weather
71(WO) – DEPARTURE STATION – [NM + AOC]
72(WT) – DESTINATION STATION – [NM + AOC]
73(WR) – EN ROUTE OR ALTERNATE – [NM + AOC]
75(WI) – DE-ICING OF AIRCRAFT – [AOC]
removal of ice and/or snow, frost prevention excluding unserviceability of equipment
Air Traffic Flow Management Restrictions
81(AT) – ATC EN-ROUTE DEMAND/CAPACITY – [NM]
standard demand/capacity problems
82(AX) – ATC STAFF/EQUIPMENT EN-ROUTE – [NM]
reduced capacity caused by industrial action or staff shortage, equipment failure, military exercise or extraordinary demand due to capacity reduction in neighbouring area
83(AE) – RESTRICTION AT DESTINATION – [NM]
airport and/or runway closed due to obstruction, industrial action, staff shortage, political unrest, noise abatement, night curfew, special flights
84(AW) – WEATHER AT DESTINATION – [NM]
Airport And Governmental Authorities
88(AD) – RESTRICTIONS AT DESTINATION – [NM + AOC]
airport and/or runway closed due to obstruction, industrial action, staff shortage, political unrest, noise abatement, night curfew, special flights
89(AM) – RESTRICTIONS AT DEPARTURE – [NM + AOC]
including Air Traffic Services, start-up and pushback, airport and/or runway closed due to obstruction or weather, industrial action, staff shortage, political unrest, noise abatement, night curfew, special flights
Reactionary
93(RA) – AIRCRAFT ROTATION – [AOC]
late arrival of aircraft from another flight or previous sector
Miscellaneous
97(MI) – INDUSTRIAL ACTION, OWN AIRLINE – [NM + AOC]
98(M) – INDUSTRIAL ACTION OUTSIDE OWN AIRLINE – [NM]

To give the coverage of those 13 delay causes, in Europe, NM Monthly Network Operation Report for May 2015 [6] states that En-Route Capacity (i.e. 81 in IATA code) with 27%, Airport Capacity (81(AT)) with 22.2% and Airport Weather (i.e. 71,72,73,75 and 84) with 16.6% were the main causes of delays in May 2015. The delays by En-Route Staffing and En-Route Events (belongs to 82(AX)) are accounted for 10.4% of all delays. Airport Staffing and Airport Events (belongs to code 83(AE), 88(AD) and 89(AM)) are accounted for 4.6% of all delays

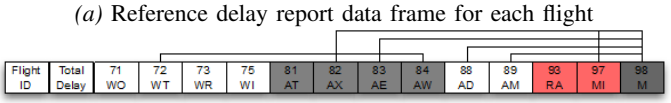
Figure 2a demonstrates a reference delay report data frame for a specific flight. This frame includes calculated delays (in minutes) for every phase of the flight and binary features (0 or 1) for the delay causes. Coupled effect on the delay causes will not be considered, since the observation and distinguishing is not possible. For guidance, relevant causes are depicted in the data frame through the connected boxes that might be set together if it is available to the participants.

The airline delay report frame (Figure 2b) includes how much delay is generated by the corresponding delay cause. The dark colored causes are locked, in that the airline representative cannot select these causes. Please note that, the red colored delay causes are directly in the field of the airline's responsibility. The airport representatives might not be fair in reporting delays associated with these causes.

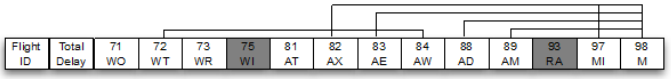
Similar to the airline delay report frame, the NM delay report frame (seen in Figure 2c) includes how much delay is generated by which delay cause. The dark colored causes are locked, i.e. the NM representative cannot select 75(WI) and 93(RA).



(a) Reference delay report data frame for each flight



(b) Delay report input data format for the airline



(c) Delay report input data format for the NM

Fig. 2: Delay Report Formats

In the simulation, the human evaluators are used as Airline and NMs. The human subjects (they will be mostly aware of the basics of ATM) are presented with the reference data frame for each flight, and they are asked to create their data models for the AOCs and NMs. as seen in Figure 2b and Figure 2c respectively. Following subsections gives the web-based portal to collect the reports and next section explains the dataset generation and simulation results.

B. Web-based Portal for Analysis of Delay Reports

To create an efficient report collection system for the simulations, we have developed a *Delay Analysis Portal* using SMC statistical libraries. This portal provides a web-based server for different level of users such as referee and participants (i.e. AOCs or NMs). Airline Operation Center (AOC) and Network Manager (NM) have certain pages specific to "delay analysis". The "delay per flight" and "delay per cause" analysis creation pages of the *Delay Analysis Portal* for the referee are depicted in Figure 3.

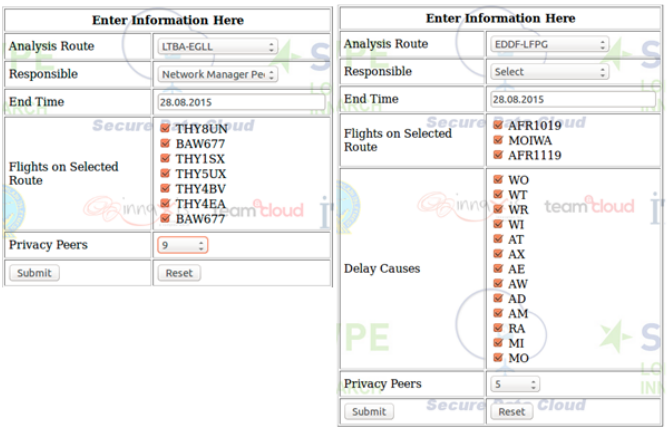


Fig. 3: Delay per flight and delay per cause analysis creation pages of referee through the simulation web portal

Following section explains input data selection from real-world events in order to demonstrate feasibility of Delay Report Analysis portal basedn on SMC.

IV. Delay Report Analysis Simulations

The section explains input data selection from real events such as weather restrictions or capacity shortage, and provides details about delay report data frames for the airlines and network managers.

A. Simulation Data Selection

We have selected 4 delay causes relevant to weather; demand capacity; special events; and aircraft rotation, and specific flight examples associated with these causes from the available dataset:

1) Weather related delays: Weather information is usually available to both Airline Operation Centers (AOC) and Network Managers (NMs). The weather delay causes are codified with 71(WO), 72(WT), 73(WR), 75(WI) and 84(AW) IATA delay codes in Table II. For the simulation prototype, the relation with delay caused by weather at both origin and destination is done by selecting the 5 problematic days, which are given in Table III. The example delay cause report associated with weather issues are given in IV. Note that, in order to not publish calculated delays for the flights, we have hide the names of the airlines through assigning three letters, e.g. AAA, BBB, CCC etc.

TABLE III: Selected weather related issues

Airport	Date	Issue
EGLL	January 23, 2013	haze: from 2km to 5km visibility
LTBA	February 01, 2015	wind: 20-23 knot
LTBA	January 31, 2015	wind: 23-31 knot
EHAM	November 24, 2012	fog: visibility is less than 1 km
LOWW	May 20, 2015	tropical storm
LTBA	May 20, 2015	heavy rain

2) Demand Capacity Problems: This information is measurable by and more visible to NMs, and we therefore assume that such information can be reported by them. The delay causes by Demand Capacity Problems are encoded by 81(AT) IATA delay code in Table II. We have chosen 5 problematic days (Table V) associated with demand-capacity imbalances.

TABLE V: Demand-Capacity Problems

Airport	Date	Issue
LIRF	June 11, 2015	20% capacity reduction due to terminal unavailability
LEMD	May 13, 2015	Reduced landing rate due to work in progress
LEMG	April 23, 2015	Reduced capacity due to ATC equipment failure
LPPT	June 12, 2015	Reduced capacity due to runway problem
EGLC	May 19, 2015	Reduced capacity due to runway problem

3) Restrictions, Industrial Actions and Special Events: Some restrictions and events are visible to both AOCs and NMs, some not. These types of events are reported before or during the occurrence by the NMs specific to their class. The delay causes by restrictions, industrial actions and other special

TABLE IV: Example reference delay reports for weather related issues (airline names are intentionally masked)

flight ID	date	org	dest	TtOff	Tland	dep dly	arr dly	13 delay causes											
AAA914	20130123	EDDF	EGLL	15:18	16:20	12	21	0	1	0	0	0	0	0	1	0	0	0	0	0
AAA6JA	20130123	EDDF	EGLL	17:07	18:09	5	17	0	1	0	0	0	0	0	1	0	0	0	0	0
CCC1890	20150201	LFPG	LTBA	11:48	14:50	16	45	0	1	0	0	0	0	0	1	0	0	0	0	0
BBB9LA	20150201	EDDM	LTBA	14:18	16:29	24	79	0	1	0	0	0	0	0	1	0	0	0	0	0
AAA1298	20150131	EDDF	LTBA	8:26	10:55	3	18	0	1	0	0	0	0	0	1	0	0	0	0	0
BBB1SX	20150131	LTBA	EGLL	6:35	9:58	12	19	1	0	0	0	0	0	0	0	0	0	0	0	0
DDD1141	20121124	EHAM	ENGM	6:03	7:34	11	0	1	0	0	0	0	0	0	0	0	0	0	0	0
BBB3KX	20150520	EGLL	LTBA	16:10	19:27	29	29	0	1	0	0	0	0	0	1	0	0	0	0	0
EEE680	20150520	EGLL	LTBA	17:10	20:36	90	93	0	1	0	0	0	0	0	1	0	0	0	0	0
BBB9JS	20150520	LTBA	LTAC	15:20	15:58	21	27	1	0	0	0	0	0	0	0	0	0	0	0	0
BBB59P	20150520	LTBA	LTAC	19:19	19:59	14	17	1	0	0	0	0	0	0	0	0	0	0	0	0
BBB5A	20150520	LTBA	LTAC	20:19	20:59	23	26	1	0	0	0	0	0	0	0	0	0	0	0	0

events (e.g. military exercise, runway closed) are seen as 82(AX), 83(AE), 88(AD), 89(AM), 97(MI) and 98(M) IATA delay codes in Table II. For simulation purposes, 5 problematic days have been selected from the NOP portal seen in Table VI.

TABLE VI: Selected special events, industrial actions, restriction for different airports

Airport	Date	Issue
LIRF	October 18, 2013	Industrial Action
LFPG	January 30, 2014	Industrial Action
LFPO	January 30, 2014	Industrial Action
LTBA	May 16, 2015	Military Exercise
EBBR	May 27, 2015	Specific Event

4) Aircraft Rotation related delays: The considered data set (ALLFT+) remarkably includes tail numbers for most of the flight. Sorting the aircraft according to their tail numbers enables to evaluate the late arrivals coming from previous leg (which is seen as a (93) RA IATA delay code). We have selected specific examples, which are given in Table VII, with late arrival through the planned and actual flight data logs. Note that, we have veil the airline companies.

TABLE VII: Example flight rotations for different aircraft (airline names are intentionally masked)

Flight ID	Date	Dep	Arr	planned takeOff	planned landing	actual takeOff	actual landing
AAA96B	20130318	LFOB	LIRA	09:10	10:49	09:29	11:17
AAA6072	20130318	LIRA	LICT	11:10	11:50	11:31	12:18
AAA9094	20130318	LICT	LIME	19:40	21:06	19:38	20:59

Flight ID	Date	Dep	Arr	planned takeOff	planned landing	actual takeOff	actual landing
BBB709	20130319	ENBO	ENST	18:25	19:01	18:40	19:11
BBB709	20130319	ENST	ENVA	19:10	20:07	19:28	20:15
BBB722	20130319	ENVA	ENBN	20:45	21:26	20:49	21:24

Flight ID	Date	Dep	Arr	planned takeOff	planned landing	actual takeOff	actual landing
CCC932	20130320	ENHF	ENMH	10:48	11:20	11:08	11:39
CCC932	20130320	ENMH	ENBV	11:37	11:46	11:57	12:07
CCC932	20130320	ENBV	ENBS	12:09	12:18	12:25	12:39
CCC932	20130320	ENBS	ENVD	12:42	12:55	13:04	13:19
CCC985	20130320	ENVD	ENTC	13:32	14:40	13:54	14:54
CCC908	20130320	ENTC	ENHF	15:45	16:31	15:36	16:11

The delay cause associated with flight rotation refers to 93(RA) in the IATA code table, thus the binary indicator associated with 93(RA) is set 1 in reference delay reports.

B. Delay Reporting Simulation Results

For each selected flight associated with the causes listed in Table II, the corresponding information was structured in the reference delay report format. Delay reports are generated for both airline operation center (AOC) and network manager (NM), considering the observability of the causes to both participant. The following subsections explain the simulation process for three SMC analysis scenarios: Delay per Flight Analysis; Ranker; and Delay Per Causes Analysis.

1) Analysis by routes: This analysis aims at aggregating all the information about flight delays, and at generating a set of statistical descriptors (including average, median and standard deviation) about those flight delays. The secure computation is performed over the sum of all partial medians, the sum of all delays by airlines, and the product of the latter by the number of flights. The remaining steps, which require a more complex mathematical framework, are performed in a traditional way: no confidentiality is lost.

For the simulation of delay per flight analysis, the delay analysis portal is used for gathering data from participants. First, the referee creates a route-specific analysis action. The delay report portal then gathers the reports from the participants and then sends the reports to the SMC server. SMC server generates the results and provides them to the referee.

The Table VIII provides some examples to demonstrate the outputs of the delay per flight analysis. For instance, the first result indicates that the average delay in EDDF-EGLL is 38 minutes, with 2.55 minutes standard deviation and 18 minutes median delay.

TABLE VIII: Result Sample of Analysis by routes

ID	Route	Number of Flights	Results
21	EDDF-EGLL	2	[38.0, 18.0, 2.55]
22	EDDF-LTBA	11	[74.0, 68.0, 35.58]
23	EDDF-EHAM	2	[23.0, 7.5, 8.51]
24	EDDF-LIRF	5	[18.0, 13.0, 5.81]

The analysis results are obtained from a batch simulation process through the delay analysis portal, where experts are used to mimic AOCs and NMs.

2) Ranker: This analysis yields a ranking of the participants, comparing the total delay their flights have suffered. All

total delays are compared one by one and sorted in a secure way.

In order to perform the ranking simulation, the associated input sets for specific routes are used and results are gathered through the delay analysis portal.

The following output (Table IX) is given as an example, for a ranker that sorts the most delayed airlines for specific routes, e.g. most delayed airlines is BBB in EDDM-LTBA route.

TABLE IX: Result Sample of Ranker

ID	Route	Results
9	EDDF-LTBA	[AAA, BBB, CCC, DDD]
13	EDDF-LFPG	[EEE, FFF]
14	EDDM-EGLL	[GGG, DDD]
15	EDDM-LTBA	[BBB, DDD]
16	EDDM-LEMD	[HHH, DDD, JJJ]

The analysis results are obtained through the delay analysis portal, where the experts/students are used to mimic AOCs and NMs.

3) Analysis by causes of delays: This analysis evaluates the delays as a function of the cause of the delay. The objective is to know the total delay introduced by each cause, the average per flight, and other statistical metrics. For the simulation of delay per cause analysis, the delay analysis portal is used for gathering data from participants. First, referee creates an analysis, gathers the reports, and then sends the results to the SMC server.

Note that input data are in following data format:

peer02: 1, 13, 0, 45, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0

where *peer02* is the name of the participant and the first two slots indicate the number of flight, e.g. N = 1, and the number of causes, e.g. M = 13. The remaining 13 fields refer to the delay-minutes for each cause, e.g. this flight takes 45 minutes delay due to cause 2.

Following example results of analysis by delay causes is given below. First example is selected from Istanbul Airport on 1 February 2015, where the wind speed was 20-23 knot. Second example is selected from Vienna and Istanbul Airports on 20 May 2015, where there was a storm and heavy rain. Third example (capacity reduction) example is selected from Rome, where there was passenger congestion in 11 June 2015.

In first example, the result of the delay analysis indicates that the major delay cause in LFPG-LTBA route is the 72(WT) with average 31.5 minutes delay, on the other words, weather phenomena in destination airport. In second example, the result of the delay analysis indicates that the major delay cause in LTBA-LOWW route is the 84(AW) [weather at destination] with an average of 25 minutes delay. In third example, the result of the delay analysis indicates that the major delay cause in LIRF-LEMG route is the 81(AT) [demand/capacity restriction] with average 6.5 minutes delay.

The analysis results were obtained from batch simulation process through the delay analysis portal, where expert/student subjects are used to mimic AOCs and NMs.

TABLE X: Analysis results for causes of delays

Route	LFPG-LTBA
Date	20150201
Inputs	peer02:1,13,0,45,0,0,0,0,0,0,0,0,0,0,0 NMpeer02:1,13,0,18,0,0,0,0,0,0,26,0,0,0,0,0
Total delays	[0.0, 63.0, 0.0, 0.0, 0.0, 0.0, 0.0, 26.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Non-zero causes	[0.0, 2.0, 0.0, 0.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Median of each cause	[0.0, 31.5, 0.0, 0.0, 0.0, 0.0, 0.0, 13.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Std of each cause	[0.0, 182.25, 0.0, 0.0, 0.0, 0.0, 0.0, 169.0, 0.0, 0.0, 0.0, 0.0, 0.0]

Route	LTBA-LOWW
Date	20150520
Inputs	peer04:1,13,25,28,0,0,0,0,0,0,0,0,0,0,0 NMpeer04:1,13,1,2,0,0,0,0,0,50,0,0,0,0,0
Total delays	[26.0, 30.0, 0.0, 0.0, 0.0, 0.0, 0.0, 50.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Non-zero causes	[2.0, 2.0, 0.0, 0.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Median of each cause	[13.0, 15.0, 0.0, 0.0, 0.0, 0.0, 0.0, 25.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Std of each cause	[144.0, 169.0, 0.0, 0.0, 0.0, 0.0, 0.0, 625.0, 0.0, 0.0, 0.0, 0.0, 0.0]

Route	LIRF-LEMG
Date	20150423
Inputs	peer06:1,13,0,0,0,0,0,0,0,0,0,0,0,0,0 NMpeer06:1,13,0,0,0,0,13,0,0,0,0,0,0,0,0
Total delays	[0.0, 0.0, 0.0, 0.0, 13.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Non-zero causes	[0.0, 0.0, 0.0, 0.0, 1.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Median of each cause	[0.0, 0.0, 0.0, 0.0, 6.5, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]
Std of each cause	[0.0, 0.0, 0.0, 0.0, 42.25, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0]

C. Computational Cost

One of the mains challenges limiting the applicability of SMC to real-world problems is the large computation cost required to perform even simple analyses. Comparing two numbers using SMC requires multiple computational steps, from dividing the initial data in shares to manipulating them in separate servers. For instance, the computational cost of a protocol based on secret sharing scheme of n players usually implies the creation of n^2 shares, representing a cost by operation of $O(n^2)$. The situation is even more complicated when non-linear operations are included in the mix, like comparisons and multiplications, which greatly increase the computational complexity and the evaluation cost. In order to assess the feasibility of a SMC paradigm, a set of simulations have been run, using the data models. The charts report the results of a set of velocity tests performed on the functional secure servers, as a function of the data input the number of clients (i.e. of participants). Three distinctive metrics have been defined, as part of the total execution time of each analysis:

- *Computation cost (blue bars)*: Time required to create and manipulate the shares.
- *Communication cost (green bars)*: Time spent by the SMC servers to transmit information among themselves, as required to perform the secure computation.
- *Communication overhead (yellow bars)*: Any other time cost, including the initial setup of the system, authentication of the clients, network discovery, synchronization between servers, etc.

In the following charts, an average of the results for an increasing number of flights per airline (left) and an increasing number of participating airlines (right) can be seen for analysis by route.

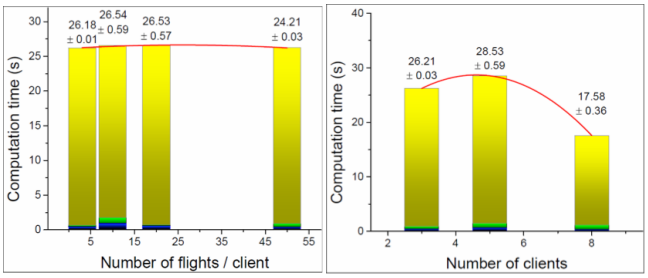


Fig. 4: Computational time as a function of the number of flights per client(left) and the number of clients (right)

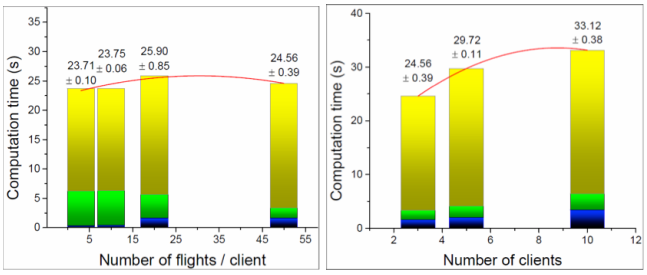


Fig. 5: Computational time as a function of the number of flights per client(left) and the number of clients (right)

Similarly, regarding the analysis by cause of delays, an average of the results for an increasing number of flights per airline (left) and an increasing number of participating airlines (right) is given in the Figure 5.

The ranker SMC Library was given, as its algorithmic complexity is obviously lesser than the delay per route and delay per cause SMC analysis library. Considering the results of computational time analysis, one can come up with following results:

- Using more computation servers increases the computation time of the SMC Libraries. This rise is not so large as to be a time related problem, whereas it presents a clear security advantage: in order to decrypt the Secret Sharing protocol, a harmful party needs to access all servers at one time.
- In a complete secure system it is important to have complete control over the execution times.

Overall, all secure computations can be executed in acceptable times, even when the number of participants increases beyond what initially estimated. Thus, the obtained results confirm the feasibility of SMC solutions in air transport environment as the complete execution times are, in average, under the one-minute bar.

V. Conclusions and Remarks

In this contribution, we have presented a secure system for delay report gathering from different stakeholders and analysing through Secure Multi-party Computation (SMC) library. Considering the needs of a reporting system, web-based delay analysis portal have been developed enabling participants to see the open questionnaire and introduce their inputs. In the simulation phase, we have selected many real-world examples through the historical ALLFT+ data analysis, exemplifying the kind of use one can make of the delay analysis portal. Experts and students have been used as reporters for both airlines and network managers. Hence, we have conceptually demonstrated the feasibility of such web-based reporting and secure analysis. It is envisioned that in addition to providing secure calculations through the SMC tools, providing frequent delay inquiries could potentially improve assessing the causes. However, in our experiments, we have observed that manipulating the results, of course, is easy if one intentionally acts in a biased manner. For example, when we have asked airline representatives to "care their businesses", they have assigned relatively small values to 93(RA) [delay due to aircraft rotation] and 97(MI) [delay due to industrial action in own airline], thus, this lead to biased outputs. Therefore, it can be said that, instead of asking report for delay through the secure information sharing, essential information sharing could be more effective in the analysis of delay.

Beyond the specific results here discussed, this contribution aimed at highlighting the necessity and feasibility of applying SMC techniques in AT and ATM. Any practitioner in air transport could easily identify a plethora of different scenarios in which private information cannot be shared, and yet some computation should be collaborative performed on it. Among others, some of them can be: ranking of airlines, both considering business elements (flight efficiency, occupancy rates, etc.) and the behavior of their own pilots (e.g. number of safety events encountered in specific routes); the study of safety data, e.g. the detection of abnormal days or airspace regions from a safety point of view; or contingency planning, involving the optimization of resources of both airports and airlines during abnormal operations. On the other hand, the Literature provides a large set of algorithms that, even if not developed with specific AT problems in mind, can easily be adapted. One of the mains challenges limiting the applicability of SMC to real-world problems is the large computation cost required to perform even simple analyses. In this contribution, we demonstrate that this limitation can be avoided by combining good programming techniques (in terms of data preparation and handling) with a cloud-based architecture. Even with high numbers of participants, all analyses here described can be performed in less than one minute, well below the time constraints set by, for instance, a slot trading problem.

While the use of SMC in air transport seems promising, there are still some issues and open problems that have to be tackled in future research activities. Two of them are of special relevance in the context here described. The first one is the kind of attackers the system is able to handle. The work described in this contribution is based on the assumption that parties are honest but curious: they will honestly collaborate in the computation, sending real data, but will try to deduce other parties inputs if the possibility arises. A different scenario may involve the presence of malicious parties, i.e. parties that

actively try to break the system by any mean at their disposal. While algorithms and protocols are available to handle such situations, their computational cost is usually extremely high [11], [17]. The second problem is the integration of such computation paradigm into existing air traffic concepts, the most notable being SWIM [12]. Future research work should be devoted to understand how both concepts could be integrated, thus effectively transforming SWIM into a platform to perform secure computations.

ACKNOWLEDGMENTS

This work is co-financed by EUROCONTROL acting on behalf of the SESAR Joint Undertaking (the SJU) and the EUROPEAN UNION as part of Work Package E in the SESAR Programme. Opinions expressed in this work reflect the authors' views only and EUROCONTROL and/or the SJU shall not be considered liable for them or for any use that may be made of the information contained herein.

REFERENCES

[1] A. Ben-David, N. Nisan, and B. Pinkas. Fairplaymp: a system for secure multi-party computation. In *Proceedings of the 15th ACM conference on Computer and communications security*, pages 257–266. ACM, 2008.

[2] G. R. Blakley et al. Safeguarding cryptographic keys. In *Proceedings of the national computer conference*, volume 48, pages 313–317, 1979.

[3] D. Bogdanov, R. Talviste, and J. Willemson. Deploying secure multi-party computation for financial data analysis. In *Financial Cryptography and Data Security*, pages 57–64. Springer, 2012.

[4] R. Cramer, R. Gennaro, and B. Schoenmakers. A secure and optimally efficient multi-authority election scheme. *European transactions on Telecommunications*, 8(5):481–490, 1997.

[5] I. Damgrd, M. Geisler, and M. Krigaard. Efficient and secure comparison for on-line auctions. In J. Pieprzyk, H. Ghodosi, and E. Dawson, editors, *Information Security and Privacy*, volume 4586 of *Lecture Notes in Computer Science*, pages 416–430. Springer Berlin Heidelberg, 2007.

[6] EUROCONTROL. Monthly network operation report – analysis. Technical report, May 2015.

[7] M. Harper. Fully homomorphic encryption. Technical report, Mathematics Department of Washington U., 2014.

[8] IATA. Airport handling manual. Technical report, 2015.

[9] V. Kolesnikov, A.-R. Sadeghi, and T. Schneider. How to combine homomorphic encryption and garbled circuits. *Signal Processing in the Encrypted Domain*, 100:2009, 2009.

[10] B. A. Malin, K. El Emam, and C. M. O'Keefe. Biomedical data privacy: problems, perspectives, and recent advances. *Journal of the American medical informatics association*, 20(1):2–6, 2013.

[11] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Bacşar, and J.-P. Hubaux. Game theory meets network security and privacy. *ACM Computing Surveys (CSUR)*, 45(3):25, 2013.

[12] J. S. Meserole and J. W. Moore. What is system wide information management (swim)? *Aerospace and Electronic Systems Magazine, IEEE*, 22(5):13–19, 2007.

[13] R. Pathak and S. Joshi. Secure multi-party computation protocol for defense applications in military operations using virtual cryptography. In *Contemporary Computing*, pages 389–399. Springer, 2009.

[14] K. Sako and J. Kilian. Secure voting using partially compatible homomorphisms. In *Advances in CryptologyCRYPTO94*, pages 411–424. Springer, 1994.

[15] A. Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979.

[16] S. Vaya. Realizing secure multiparty computation on incomplete networks. In *Security and Cryptography (SECRYPT), Proceedings of the 2010 International Conference on*, pages 1–8, July 2010.

[17] G. Zhang, Y. Yang, X. Zhang, C. Liu, and J. Chen. Key research issues for privacy protection and preservation in cloud computing. In *Cloud and Green Computing (CGC), 2012 Second International Conference on*, pages 47–54. IEEE, 2012.