ROGER: An On-Line Flight Efficiency Monitoring System using ADS-B Data

Shen Wang†, Aditya Grover†, Brian Mac Namee†, Philip Plantholt‡,
Javier López-Leónés‡, and Pablo Sánchez-Escalonilla§

†CeADAR Centre for Applied Data Analytics Research, University College Dublin, Dublin, Ireland
Email: shen.wang@ucd.ie, aditya.grover@ucd.ie, brian.macnamee@ucd.ie
‡Business Strategy and Development, Flightradar24 AB, Stockholm, Sweden. Email: philip@fr24.com
§Boeing Research and Technology Europe, Madrid, Spain. Email: javier.lopezleones@boeing.com
CRIDA, ATM R&D Reference Center, Madrid, Spain. Email: psescalonilla@e-crida.enaire.es

Abstract—Flight efficiency indicators reported monthly in the European area by the Performance Review Unit (PRU) help the air traffic management (ATM) community determine if excessive distances are being flown (compared with the ideal lengths of flight routes). Recent research, however, provides more indicators that comprehensively capture flight efficiencies in terms of other factors including fuel consumption, time adherence, and route charges. The efficacy of all of these indicators, however, is diminished as they are currently only available almost a month after flights take place. This is not sufficiently timely to use these indicators for the alleviation of unpredictable hotspots (i.e. sectors with congested air traffic), which often leads to unexpected ground delays. This paper proposes a methodology to calculate general flight efficiency indicators on-line in near real-time using nearest point search. A prototype system called ROGER (compRe-hensive On-line fliGht Efficiency monitoRing) is implemented using Apache Kafka and Spark. ROGER can digest large-scale heterogeneous datasets (i.e. mainly ADS-B data, the next generation aircraft surveillance technology) to compute indicators every 5 seconds. Our experiments on realistic datasets demonstrate that the proposed on-line indicator calculation method can achieve high accuracy compared with existing off-line approaches, and that ROGER can achieve desirable system performance in throughput and latency. A use case is also described showing how ROGER can assist in alleviating hotspots more effectively.

1. Introduction

Recent years have witnessed fast growing air traffic demand around the world. According to a 2017 European Commission report [1], global air traffic has been increasing, on average, by over 5% annually for more than 10 years—well above global GDP growth. Limited airspace and slowly increasing airport capacity, mean that it is becoming challenging for air traffic controllers and airlines to achieve optimal flight efficiencies.

To fully understand flight efficiency, in the European area, the EUROCONTROL Performance Review Unit (PRU) collects, analyses, and releases flight efficiency reports to ATM community actors (including airlines and air navigation service providers (ANSP)). Two flight efficiency indicators are frequently used: key performance indicator based on actual trajectory (KEA) and key performance indicator based on last filed flight plan (KEP). These measure the percentage of extra distance contained in the actual trajectory flown (KEA) or in the last flight plan trajectory (KEP) by a certain flight compared to the shortest possible trajectory between the origin and destination airports. To study more sophisticated measures of flight efficiency, Calvo et al [2] proposed new indicators that capture time-adherence and fuel consumption efficiency factors. [3] improved upon these and extended to new indicators that account for overall monetary cost. However, existing flight efficiency results only become available for analysis after a flight’s arrival. This significantly delays ATM actors in making adjustments to drive more efficient flights. For instance, short-term air traffic flow and capacity management (STAM) measures (e.g. re-routing, level-capping, ground delay) that are taken to avoid sectors of airspace being overloaded with flights could be more efficiently planned if efficiency indicators were available to airspace users in near real time. An on-line flight efficiency monitoring system is highly required.

One of the core technologies that could enable real-time ATM performance monitoring is automatic dependent surveillance-broadcast (ADS-B) [4]. ADS-B is a surveillance technology in which an aircraft regularly broadcasts its own position (calculated using on-board GPS) which, when received and aggregated by ground stations, can be used to track the aircraft. ADS-B can send richer information (i.e. location, weather, and flight) to more receivers (i.e. including ground stations and nearby aircraft) at a longer range, with higher frequency, and with lower deployment cost than on-ground radar systems. Companies, such as Flightradar24, collect ADS-B data through crowd-sourcing to provide a global real-time aircraft monitoring service. ADS-B is now widely equipped on civil aviation aircraft

1. See http://ansperformance.eu/about/
and will be mandatory in the coming years (it is compulsory for some aircraft in Europe from 2017 and will be so in the USA from 2020). Recent research work on ADS-B has mainly focused on the security aspects of the protocol [5], methodologies for fusing ADS-B data with existing radar surveillance technology [6], and future challenges [7] for both manned and unmanned aircraft. To our knowledge, there is limited proof-of-concept work using ADS-B data to improve ATM performance at the operational level. There are three contributions made in this paper: A methodology is proposed to calculate flight efficiency indicators on-line. Existing off-line efficiency indicators only have one value for each flight when the aircraft is landed at its destination airport. Using the proposed on-line method, flight efficiency indicators values can be calculated during flights approximately every 5 seconds. A prototype system called ROGER (comprehensive On-line Flight Efficiency monitoring) is designed and implemented. ROGER can digest large-scale heterogeneous datasets (e.g. ADS-B data, weather data, aircraft performance etc.) to compute on-line indicators using the proposed method at the same frequency as the input surveillance data stream. ROGER is implemented using the state-of-the-art big data streaming technologies from the open source community—Apache Kafka\(^3\) and Apache Spark\(^4\)—which ensures scalable and reliable system performance. A case study using realistic datasets is completed to verify ROGER. ROGER is evaluated using realistic datasets that cover almost all flights that departed and arrived in the European area on 20th and 24th February, 2017. The evaluation results show that the proposed on-line indicator calculation method in ROGER can achieve high accuracy when compared with existing off-line methods. Moreover, ROGER can cope with high throughput (i.e. up to 361 messages per second), respond with low latency (i.e. 33 seconds on average, 77 seconds for 99th percentile), and have no data loss, no duplicated data, and no out-of-sequence data. A use case is also described to show how ROGER can assist with the alleviation of hotspots.

2. System Overview

The ROGER system has been designed to meet following goals:

1) To process heterogeneous datasets that include both streaming (e.g. aircraft surveillance data) and non-streaming data (e.g. weather data, aircraft performance data, and flight plan data).

2) To calculate on-line flight efficiency indicators efficiently and accurately, given any updated flight trajectory point.

3) To achieve satisfiable system performance (i.e. throughput and latency) for all potential users.

This section describes the system designed to meet these goals.

2.1. Datasets

ROGER is underpinned by a number of key datasets that are described in this section.

Surveillance: We use ADS-B surveillance data that includes information such as the last updated time, location, registration number, and speed for each aircraft tracked. We source our ADS-B data from Flightradar24. The update frequency of this data is every 5 seconds.

Aircraft performance model: We use the base of aircraft data (BADA)\(^5\) service provided by EUROCONTROL. It contains descriptions of aircraft performance, including the procedure model of climb, cruise, and descent, along with engine parameters of major aircraft.

Flight plan: We use the demand data repository (DDR)\(^5\) service provided by EUROCONTROL as the source of flight plans. A flight plan describes the waypoints and schedule that each flight plans to adhere to. The last filed flight plan is determined 3 hours before the departure of a flight.

Operational context: We also retrieve operational context data from the EUROCONTROL DDR service. The operational context data includes the partitions of airspace sectors in the European area, as well as important attributes of each sector (e.g. capacity and route charge rate).

Weather: We use weather data provided by the Global Forecast System (GFS)\(^6\) that has information such as temperature, wind, and pressure for any four dimensional point describing date and time, latitude, longitude, and altitude.

2.2. System Architecture and Data Flow

Figure 1 illustrates the architecture of ROGER. The main components of ROGER are an input ADS-B surveillance data stream; a trajectory reconstruction service which generates trajectory states (e.g. instantaneous mass) given a sequence of actual surveillance points; a stream processor that calculates on-line efficiency indicators based on surveillance data; a trajectory generation service that computes the ideal trajectory for each flight in various cost types (e.g. distance, fuel) and updates these every 3 hours using the latest estimated initial aircraft mass from the reconstructed trajectory; and a data store to persist all calculated on-line efficiency indicator results.

---

5. http://www.eurocontrol.int/ddr
Key data flows in this architecture are labelled with digits 1 to 8 in Figure 1. These are described below:

1. The ADS-B surveillance data stream is ingested to a buffer to adapt the receiving rate to the subsequent processing rate.

2. The latest surveillance data in this buffer is then moved out and accumulated with previously received ADS-B data to formulate a full trajectory for each flight since its departure, which is required by its subsequent trajectory reconstruction service. We use the “call sign number” combined with “departure time” to uniquely identify a flight.

3. The trajectory reconstruction service is triggered every 5 seconds to derive extra states (e.g. instantaneous mass) for all updated actual trajectory points. To avoid a performance bottleneck, this service is called in a multi-threaded manner, considering each unique flight as the parallelism unit.

4. The reconstructed trajectories are sent on to an Apache Kafka buffer. This reliable buffer can ingest data with high throughput and low latency for more complicated processing tasks afterwards.

5. The Kafka stream producer application reads reconstructed trajectory streams from the buffer and sends them to the stream processor. This stream producer guarantees reliable message transmission with no duplication, no data loss, and no out-of-sequence messages.

6. The stream processor, which is implemented using Apache Spark Streaming, pulls the reconstructed trajectory streaming data every 15 seconds to aggregate a micro-batch and computes the actual cost values that correspond to all new reconstructed trajectory points, such as travelled distance, consumed fuel, and overall cost.

7. This stream processor also retrieves the ideal cost value using nearest point search from pre-loaded in-memory generated trajectory data, then calculates required flight efficiency indicators with the actual cost value computed by the stream processor.

8. The stream processor outputs the calculated on-line indicator results on to a PostGIS data store for subsequent complex spatial queries. For example, the air traffic network manager may check if in an airspace sector one indicator’s value is fairly distributed among airlines given a specific time range.

It is worth noting that the aircraft instantaneous mass (including initial mass) is estimated by Boeing’s trajectory reconstruction service as airlines do not share this confidential data. The more surveillance data received the higher accuracy this estimation is, which lead to a periodically updated trajectory generation service.

### 3. On-line Indicator Calculation

Almost all existing flight efficiency indicators [3] are computed as the percentage of extra cost of an actual flight trajectory (i.e. reconstructed trajectory $T_{rec}$) compared with its corresponding cost in an ideal trajectory (i.e. generated trajectory $T_{gen}$). We extend this idea to the on-line scenario over the course of a flight. Specifically, the indicator value at the current time step $t_i$ can be computed using:

$$IND_{t_i} = \left( \frac{C_{t_i}}{C_{t_i}^{*}} - 1 \right) \%$$

where $C_{t_i}$ defines the cost (i.e. distance $d$, fuel $f$, overall cost $c$) of $T_{rec}$ up to time $t_i$ (calculated on-line), $C_{t_i}^{*}$ represents the corresponding optimum cost value of $T_{gen}$ up to the point which is geographically nearest to the trajectory point in $T_{gen}$ at time $t_i$. $C_{t_i}^{*}$ can be retrieved directly from each type of generated trajectory (i.e. 4 in this study, details shown in Table 1) which are pre-computed and loaded in our system. We consider the geographically nearest point (rather than “the closest time point” or other potential metrics) in terms of the two dimensional (i.e. longitude and latitude) great circle distance, as it is reasonable to define 0% completion of a flight as when its current location is at an origin airport, and 100% completion of a flight as when the aircraft has geographically reached a destination airport. Table 2 lists ten on-line efficiency indicators calculated in this study, including their identifiers, the type of cost used, and the generated trajectory used.

#### TABLE 1. THE IDENTIFIERS, OPTIMISATION GOALS, AND CONSTRAINTS OF ALL GENERATED TRAJECTORIES $T_{gen}$

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Constraints</th>
<th>Optimisation Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEO_FP</td>
<td>flight plan and weather</td>
<td>distance</td>
</tr>
<tr>
<td>UP</td>
<td>flight plan and weather</td>
<td>–</td>
</tr>
<tr>
<td>FREE_CI</td>
<td>weather</td>
<td>time and fuel</td>
</tr>
<tr>
<td>OPT_CI</td>
<td>flight plan and weather</td>
<td>time, fuel, and route charges</td>
</tr>
</tbody>
</table>

#### TABLE 2. THE IDENTIFIER, THE TYPE OF COST, AND GENERATED TRAJECTORY USED FOR THE COMPUTATION OF ON-LINE FLIGHT EFFICIENCY INDICATORS

<table>
<thead>
<tr>
<th>Generated Trajectory $T_{gen}$</th>
<th>UP</th>
<th>GEO_FP</th>
<th>FREE_CI</th>
<th>OPT_CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost</td>
<td>distance</td>
<td>KEAP</td>
<td>KEA</td>
<td>KEAC1</td>
</tr>
<tr>
<td>fuel</td>
<td>KEAP</td>
<td>FEAP</td>
<td>FEAC1</td>
<td>FEAC2</td>
</tr>
<tr>
<td>cash</td>
<td>CEAP</td>
<td>CEAC1</td>
<td>CEAC2</td>
<td></td>
</tr>
</tbody>
</table>

The general steps to calculate indicators are as follows:

1. pre compute all cost values for $T_{gen}$ and load them fully on to our system; 2. as the surveillance data stream is ingested, all cost values of the latest updates in $T_{rec}$ are calculated on-the-fly; 3. the nearest points of the latest updates in $T_{rec}$ are then found in each type of the corresponding $T_{gen}$; 4. the indicator values can then be computed using the mappings shown in Table 2. For example, the value of KEA should be equal to the ratio of the excessive value of $d$ in $T_{rec}$ to $d$ in $T_{gen}$ - GEO_FP.

The calculations of all cost values for both $T_{gen}$ and $T_{rec}$ are given as:

**Travelled distance** $d_{t_i}$ : for a given flight that departs at time $t_0$, at a given time, $t_i$, the travelled distance is calculated as:

$$d_{t_i} = \begin{cases} 0, & \text{if } t_i = t_0 \\ dist(T_{t_i}, T_{t_{i-1}}) + d_{t_{i-1}}, & \text{otherwise} \end{cases}$$
where \( \text{dist}(T_t, T_{t-1}) \) defines the geographic distance between two consecutive trajectory points \( (T_t, \text{ and } T_{t-1}) \) calculated using Vincenty’s formulae [10].

**Consumed fuel** \( f_t \): Using the trajectory reconstruction service, the instantaneous mass \( m_t \) can be estimated for all trajectory points. Thus, the consumed fuel is computed as:

\[
f_t = m_{t0} - m_t
\]

**Overall cost** \( c_t \): The overall cost in euros (€) at \( t \) includes time, fuel, and route charges. Cost index, \( CI \), is a constant value that reflects airline’s perspective towards the proportion of time and fuel that should contribute to a flight’s total cost. This is used with the fuel price, \( P \), (in euros/kg) to merge and convert the cost of time and fuel into euros, then added to the route charge, \( rc_t \):

\[
c_t = P \times (f_t + CI \times (t_k - t_0)) + rc_t
\]

Figure 2. An example showing how route charge grows as new reconstructed trajectory points ingested.

The route cost \( rc_t \) is calculated as the summation of the route charge in the current zone \( \text{cur}_t \) and the route charge in all previously flown zones \( \text{prev}_t \) since a flight’s departure:

\[
rc_t = \text{cur}_t + \text{prev}_t
\]

We denote \( T_t \) as the current trajectory point, and \( T_{t-1} \) as the trajectory point at the previous time step. Given \( T_t \), we find its current airspace zone \( z_t \) (e.g. Germany or France, as shown in Figure 2) using a point-in-polygon query. We also denote \( M \) as maximum take-off weight (MTOW) which is the constant value for each aircraft, \( R(z_t) \) as the charge rate in zone \( z_t \), and \( T_{t\text{entry}} \) as the entry trajectory point in the current airspace zone \( z_t \). As shown in the off-line indicator calculation [11], the distance considered in route charges only depends on the entry and exit trajectory point in each flown airspace zone. We extend this calculation to the on-line scenario using the following equations:

\[
\text{cur}_t = \begin{cases} 
0, & \text{if } z_{t-1} \neq z_t \\
\text{dist}(T_{t\text{entry}}, T_t)/100 \times \sqrt{M/50} \times R(z_t), & \text{otherwise}
\end{cases}
\]

\[
\text{prev}_t = \begin{cases} 
\text{prev}_t + \text{cur}_t, & \text{if } z_{t-1} \neq z_t \\
\text{prev}_t, & \text{otherwise}
\end{cases}
\]

Specifically, as new trajectory points are coming in, we compare the existing charging zone of each consecutive trajectory point pair \( z_{t-1} \) and \( z_t \). If \( z_{t-1} = z_t \), then a flight has not finished its journey in its current zone and only \( \text{cur}_t \) needs to be updated. If \( z_{t-1} \neq z_t \), then the flight has just finished its journey in the previous zone \( z_{t-1} \), and has started another journey in a new zone \( z_t \) and so \( \text{prev}_t \) should be updated, and \( \text{cur}_t \) should be reset to 0. We also mark this point-pair as the exit \( T_{t-1} \) and entry point \( T_t \) pair of zones they belong to respectively.

We use the cost index from flight plan, maximum take-off weight (MTOW) from aircraft performance data, and the partition of airspace in Europe with its route charge rate from operational context dataset.

### 4. Evaluation

We perform evaluation experiments to measure the accuracy of the on-line indicator calculations implemented in ROGER and the efficiency with which they are calculated. This section describes the setup of these experiments and their results.

#### 4.1. Experiment Settings

We evaluate ROGER using an ADS-B dataset covering flights that depart and arrive in European airspace on 20th and 24th February 2017. This dataset contains about 30,000 flights and around 38 million cleaned ADS-B surveillance points evenly split across each day. We have access to GEO_FP and UP generated trajectories for all of these flights. As the trajectory generation service used in this study has not been optimized for FREE_CI and OPT_CI, we calculated these cost-based trajectories for a subset of flights (around 1,200 for each type), which depart between 12:00 and 14:00 on each day. It takes about 1 second for Boeing’s trajectory reconstruction service to reconstruct a single flight’s trajectory. Currently this service does not support fully parallel computing for streaming context and so we simulate the latency of trajectory reconstruction using a randomly selected positive value, from a normal distribution with mean value of 1.0 and standard deviation of 0.1. This simulated service reconstructs each trajectory in parallel. The parallelism capacity is set as 144 = 1 (the number of servers) * 24 (the number of CPU processors) * 6 (the number of cores per CPU processor). All tests are performed on a single machine equipped with Intel Core i7K 8700K (6-Core/12-Thread) 3.7 GHz, 32GB DDR4 RAM,2TB 7200RPM SATA 6Gb/s disk, and the Ubuntu 16.04LTS operating system.
4.2. Accuracy of on-line calculation

The accuracy of our proposed on-line indicator calculation method is measured in absolute error between offline (accurate) and online (approximate) indicator values. We obtain offline results from ATM experts in CRIDA as the ground truth data. These offline results contain 10 efficiency indicator values for over 1,000 verified flights in total for both 20th and 24th Feb, 2017. For each flight, we choose the last point of its on-line results to compare to the offline calculations.

As shown in Figure 3, most of the on-line indicator values are very close to the accurate off-line values, with the median absolute error less than 0.5 percentage points. The maximum absolute error occurs in indicator FEAC2 which is still less than 3.5 percentage points. The difference between on-line and off-line results could be because the nearest point of the last reconstructed trajectory point, might not be the last point of its corresponding generated trajectory. Moreover, the last reconstructed point might not always be exactly at a flight’s destination. Besides, the on-line results are calculated using several approximations (e.g. they do not distinguish between flight stages such as taking off, cruise, and landing, of a flight) for obtaining good system performance.

4.3. System performance

We measure the performance of the ROGER system in terms of throughput and latency. We define throughput as the number of data records processed by ROGER per 10 minutes from the ADS-B surveillance data source. The throughput in the whole system is shown in Figure 4. The peak traffic over two testing days is equivalent to about 361 records per second. Considering a lot of subsequent spatial computations (e.g. great circle distance calculation, nearest point search, point-in-polygon query), which are more computationally expensive than typical map and reduce operations in most big data applications, this throughput can still be considered heavy.

Under such throughput, the average latencies that ROGER achieves is shown in Figure 5 and Figure 6. This latency accounts for the time spent for each message, from when a surveillance trajectory point is received by ROGER, to the moment when a corresponding updated set of efficiency indicators is written in the sink data store. This includes the 5 seconds batch interval for the buffer between ADS-B data stream source and trajectory reconstruction service, the delay when reconstructing the trajectory, and the 15 seconds batch interval for Spark Streaming. A panel of expert airspace users, has recently set an update frequency (and so maximum allowed latency) target on efficiency indicator calculation of 5 minutes. Our results show performance much better than this target. The mean latency for messages processed by ROGER is about 33 seconds, while the 99th percentile latency observed for a message in our current dataset is less than 77 seconds. Besides, the longer latency generally occurs when generated trajectories are being updated, a process which occurs every 3 hours.

4.4. ROGER-assisted STAM

The main motivation of building ROGER for monitoring flight efficiency on-line is to enable better planning of STAM measures, which air traffic controllers (ATC) can use for re-routing, level-capping, or ground-delay to alleviate any detected hotspots (i.e. in a certain airspace sector the aircraft counts during a time interval is beyond its upper limit) in the tactical stage (i.e. day of operations), rather than pre-tactical (i.e. 1 to 6 days before operation) or strategic stages (i.e. longer than 7 days before flight operation). We show an example of using the output of ROGER to facilitate STAM decision-making. Assume that German airspace
Figure 6. The error bar (standard deviation) plot of latency every 30 minutes over the full day traffic scenario on 24th Feb (ICAO code: ED) is detected as a hotspot at 3pm. An ATC can perform spatio-temporal queries on the ROGER efficiency indicator data store to get up to date efficiency indicators for each flight within this region. The average efficiency indicators for each airline with flights in the hotspot can be calculated and hotspot alleviation measures can be applied to only flights from those airlines that are currently running efficiently so as to minimise the impact of hotspot alleviation measures. Figure 7, shows the median CEA-C2 value for each of the 5 airlines that have flights in this region calculated at 30-minute intervals in the time leading up to the hotspot arising. From this figure, a primary conclusion can be seen that if ATC intends to do STAM for some airlines, airline4 and airline1 (real airline names are hidden for confidentiality) between in 13:30-14:00 is definitely not a fair option as they have already sacrificed a lot in overall cost in this airspace.

5. Conclusions

To dynamically monitor flight efficiencies, we propose a methodology to calculate efficiency indicators on-line and implemented a prototype system, ROGER, using Apache Kafka and Spark Streaming. We showed that the calculated on-line indicator results using our proposed method have high accuracy when compared with off-line results. The experiments also demonstrate that ROGER can achieve desired performance in terms of throughput and latency. In the future we intend to fully integrate the trajectory generation and reconstruction services and evaluate use cases for taking advantage of real-time efficiency indicators. We will also explore ways in which the performance of the system could be improved.

Acknowledgments

This research was performed as part of the European Union’s Horizon 2020 AURORA project7 (Grant Number: 699340) and supported by Enterprise Ireland and IDA Ireland under the Technology Centres Programme (TC 20130013).

References
