A model for enhancing robustness of aircraft and passenger connections

Mohamed Ali Aloulou a,b, Mohamed Haouari c,⇑, Farah Zeghal Mansour d

a PSL, Université Paris Dauphine, 75775 Paris Cedex 16, France
b CNRS, LAMSADE UMR 7243, France
c Department of Mechanical and Industrial Engineering, College of Engineering, Qatar University, P.O. Box 2713, Doha, Qatar
d Department of Industrial Engineering, Ecole Nationale d’Ingénieurs de Tunis Université de Tunis El Manar, BP 37, le Belvédère, 1002 Tunis, Tunisia

A R T I C L E   I N F O

Article history:
Received 4 December 2012
Received in revised form 23 March 2013
Accepted 25 March 2013

Keywords:
Robust airline planning
Flight delays
Aircraft routing
Simulation

A B S T R A C T

Airline schedules are often subject to various uncontrollable factors that cause disruptions and delays. These delays not only constitute, for many passengers, the painful part of air travel, but also represent for airlines a significant financial burden. In this paper, we present a model for building robust aircraft routes that are less vulnerable to disruptions through judiciously distributing slacks to connections where they are most needed operationally. Toward this end, the model assigns legs to aircraft and determines the flights departure times, while maintaining the designated time-slot assignments at airports, and satisfying operational constraints. The considered objective function is a newly proposed surrogate measure of robustness that implicitly captures the robustness both pertaining to aircraft and passenger connections.

Computational experiments carried out on real-world-based instances, with up to 1278 flights and 251 aircraft, show that the model yields solutions that are remarkably robust. In particular, a simulation study reveals that the total delays, number of delayed flights, and number of missed connections can be significantly reduced.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The application of optimization techniques to the airline industry is considered as one of the most impressive success stories of operations research. Indeed, since the early 1960s, literally thousands of papers have been published so far dealing with numerous optimization problems arising in the air transport industry. These intense research efforts address highly complex and large-scale problems including, but not restricted to, schedule planning, fleet assignment, aircraft routing, maintenance planning, crew scheduling, and revenue management. Needless to emphasize the overwhelming impact the information and optimization technologies had on increasing the profitability of the airline industry and enabling commercial aviation operations to be supported by what is possibly the most complex man-made system in the world (Ball et al., 2007).

Notwithstanding these impressive achievements, a seemingly paradoxical fact is that the actual relevance of optimization models to the airline industry is legitimately questionable. In this context, Barnhart and Cohn (2004) raised the question “Are optimal plans optimal in practice?”. Put in simple words, this intriguing question aims at emphasizing the significant gap that often exists between the planned optimized schedules from those really implemented. Actually, optimized schedules are scarcely, if ever, implemented as planned. The main reason is that planned schedules are often subject to many sources of uncertainties and uncontrollable factors that perturb the aircraft schedules, crew schedules, and passenger itineraries. These perturbations cause delays, flight cancellations, and missed connections. Hence, a schedule that is optimal under

⇑ Corresponding author. Tel.: +974 70131236.
E-mail address: mohamed.haouari@qu.edu.qa (M. Haouari).

0968-090X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trc.2013.03.008
the assumption of complete information and deterministic environment may largely be affected by unpredictable and stochastic disruptions.

As a consequence, aircraft arriving late, missed connections, and canceled and diverted flights continue to plague carriers and have become an inevitable painful part of air travel. Statistics publicly available from the US Department of Transportation (DOT) reveal the extent of these disruptions.\(^1\) In 2011, US airlines recorded just 79.62% of their flights arrived to their destinations with a delay not exceeding 15 min. Flight delays were even higher for some regional airlines which recorded more than one of every four flights was delayed in June 2011. In Europe, the magnitude of this flight delay problem is somewhat comparable. Indeed, in 2005 over 20% of European airlines flights departed more than 15 min later than their scheduled departure time (see Santos and Robin (2010)).

Major sources of delays include:

- Aircraft shortages: Aircraft breakdown, late-arriving aircraft (that is, a previous flight with same aircraft arrived late, causing the present flight to depart late), longer than expected ground operations (including maintenance checks, cleaning, catering, baggage unloading/loading, fueling, deicing, etc.).
- Crew shortages: Delayed connecting cockpit crews or flight attendants, crew sickness.
- Passengers shortages: Longer than expected embarking or disembarking times, longer than expected time at security control areas.
- Airspace capacity shortages: Congested facilities (including boarding gates, ramp facilities, aircraft parking areas, runaways, etc.), heavy traffic volume and air traffic control.
- Inclement weather: Refers both to non-extreme weather conditions that might slow down operations, as well as extreme meteorological conditions (such as sandstorm) that temporarily prevent aircraft from taking off or landing.

Not surprisingly, these disruptions not only damage the airlines image and reputation, but also prove extremely harmful to the air transportation sector. Barnhart et al. (2012) quote that according to the US Congress Joint Economic Committee, the cost associated with domestic flight delays in the United States during 2007 was estimated at $25.7 billion.

Because of soaring oil prices, narrow profit margins, and steady increase of airspace and airports congestion, airlines are seeking opportunities to effectively manage disruptions to reduce costs and enhance revenues. As a result, airline companies and operations research analysts have developed two broad strategies: Reactive strategies and proactive strategies. The former refer to a set of recovery approaches that aim at using optimization to plan for a cost-effective return to a feasible schedule after the occurrence of a disruption. Potential decisions include re-routing aircraft, delaying or canceling flights, reassigning crews, and accommodating disrupted passengers on alternative itineraries. Kohl et al. (2007) provide a description of the disruption management process, and deliver a detailed overview of the numerous aspects of airline disruption management. For comprehensive analyses of recovery strategies, models, and associated solution algorithms, we refer to Ball et al. (2007), Barnhart (2009) and Clausen et al. (2010).

In contrast with these recovery decisions, which are meant to be implemented in the aftermath of the occurrence of a disruption, proactive strategies aim at producing at the planning stage robust plans that are expected to be less vulnerable to disruptions or leading to cheaper (or easier) recovery processes.

In this paper, we are concerned with the development of a proactive approach to generate robust aircraft routes that are minimally affected by late-arriving aircraft. At this point, it is worth mentioning that the late-arriving-aircraft “syndrome” is largely ranked as the primary cause of delays. Indeed, it represents 40.8% of total delay minutes recorded in 2011 by US carriers. Fig. 1 shows that the importance of this delay cause has been steadily increasing since 2003.\(^2\)

In this paper, we make the following contributions:

1. Given a flight schedule together with the associated fleet assignment and passenger itineraries, we propose a Mixed Integer Programming (MIP) model that aims at revising the flight departure times and assigning aircraft to legs in a fashion that leads to judiciously inserting buffers across the connecting flights to provide protection against possible delays. Toward this end, we propose an objective function that implicitly captures the robustness pertaining to both aircraft and passenger connections.
2. We present the results of a Monte Carlo simulation study that demonstrates that the proposed plans exhibit a remarkable robustness and are less vulnerable to disruptions caused by late-arriving aircraft.

The remainder paper is organized as follows. In Section 2, we present a review of the literature pertaining to robust airline operations planning models. Sections 3 is devoted to the description of a surrogate robustness measure. In Section 4, we present the MIP model. Section 5 is devoted to the analysis of the computational performance of the proposed model and the results of a simulation study. Finally, Section 5 provides some concluding remarks and recommendations for future research.

\(^{1}\) http://www.bts.gov/xml/ontimesummarystatistics/src/index.xml.

2. Literature review

A practical way to deal with uncertainty in airline operations planning is to design plans that incorporate some specific robustness-inducing features, which attempt to mitigate the impact of unpredicted disruptions. This goal can be achieved by incorporating within the planned schedules any of the following two properties: stability and flexibility (Düeck et al., 2012). The term stability refers to the insensitivity of the schedule to minor changes in the operating environment. On the other hand, the concept of flexibility refers to the ability of a schedule to be easily repaired.

2.1. Approaches used for generating stable plans

Building a stable plan entails the development of a schedule that will have a lesser likelihood of being damaged by unpredicted disruptions. A popular way to generate stable plans involves inserting slacks into the schedule to enable the absorption of small-scale delays and to prevent their propagation in the downstream network. In this regard, Lan et al. (2006) investigated approaches for building stable aircraft routes that are obtained through an optimal allocation of slack in order to absorb the delay propagation. They proposed two models, the first of which builds robust aircraft routes while minimizing the expected total propagated delay of selected aircraft routes, and the second selects flight departure times to minimize the expected number of disrupted passengers, while maintaining the current fleeting, aircraft routings, and passenger itineraries. Recently, an enhanced variant of the first model of Lan et al. (2006) has been investigated by Dunbar et al. (2012) who proposed an approach to simultaneously produce robust aircraft routes and crew schedules (but without considering the flight retiming option). To that aim, both problems were formulated as set partitioning problems with objective functions seeking to minimize the total expected propagated delay costs. Furthermore, the issue of improving robustness through flight retiming has also been investigated by Lee et al. (2007). These authors developed a multi-objective model that seeks to revise the departure times of a flight schedule without altering the fleet assignment, the aircraft routings, and the crew pairings. The problem was solved using a multi-objective genetic algorithm. A further contribution dealing with using the flight retiming strategy to improve the solution’s robustness has been proposed by Ahmadbeygi et al. (2010) who developed a linear programming model to improve robustness of aircraft schedules through flight retiming and subject to a fixed aircraft routing. As an objective function, they considered the sum of the impacts of each individual primary delay (selected out of a discrete set of possible delay values) as it propagates downstream, weighting these delays by their relative probabilities.

2.2. Approaches used for generating flexible plans

A flexible plan is usually achieved by increasing the number of opportunities to easily recover from a disruption. Major modeling concepts proposed in the literature for enhancing flexibility in the planning stage are as follows.

(a) Short cycles and hub connectivity (Rosenberger et al., 2004): If a flight is canceled in case of a disruption, then the downstream legs on the same aircraft route must be canceled as well. In order to avoid the ensuing costly recovery ferrying option, airlines often choose to cancel all the sequence of legs in a rotation that departs and ends at the same station where the disruption occurred. Consequently, a fleet assignment and aircraft rotation with many short cycles is frequently less sensitive to flight cancellations. Furthermore, to reduce the cascading effect of a disruption at one hub on other hubs, airlines tend to reduce the hub connectivity. This network attribute refers to the number of legs in a rotation that are in a route that begins at a hub, ends at a different hub, and only stops at spokes in between.
(b) Station purity (Smith and Johnson, 2006): This concept of station purity, refers to limiting the number of fleets that serve each spoke station. In so doing, it is possible to create more opportunities to swap aircraft or crews to cover operational disruptions.

(c) Aircraft swap opportunities (Burke et al., 2010): This strategy seeks to maximize the number of aircraft swap opportunities to generate robust aircraft plans.

In this paper, we deal with building aircraft stable schedules while retiming flights. The proposed robustness measure is a new slack-based quantitative metric that is maximized within a mixed-integer linear programming formulation. Moreover, we present an effective network flow formulation that offers the significant advantage to be solvable using a general-purpose MIP solver and thus avoid implementing sophisticated branch-and-price solution procedures.

3. Surrogate robustness measures

3.1. The snowball effect in airline operations

The snowball effect is a figurative term that describes how a fact of small significance can lead to something much severe over time. In airlines operations, it refers to the situation where an initial delay on a single flight can propagate to the connecting flights, and spread the network causing more severe delays. An illustration of this effect is displayed in Fig. 2.

In this example, we assume that flight $f_1$ is planned to arrive at 9:10 and that:

- The cockpit crew of flight $f_1$ is assigned to flight $f_2$ departing at 9:50. The sit-time is 30 min.
- Some passengers on flight $f_1$ are connecting on flight $f_3$ departing at 10:00. The minimum connection time is 40 min.
- The aircraft that served $f_1$ is scheduled to serve flight $f_4$ at 10:20. The turn time is 50 min.

We assume that $f_1$ is delayed by 30 min. From Fig. 2, we see that the immediate consequences of this delay are the following:

- The crew on $f_1$ cannot be on time for its next scheduled flight $f_2$ and therefore this latter is delayed by 20 min.
- Passengers that are connecting from $f_1$ to $f_3$ would miss this latter flight. In some situations (for instance when there exists a large number of connecting passengers), $f_3$ would be delayed to wait 20 min for the connecting passengers.
- The aircraft that serves $f_1$ cannot be on time for its next planned flight $f_4$ and therefore this latter is delayed by 10 min.

Hence, while the primary delay is 30 min, we see that the reactionary (i.e. implied) delay is 50 min. Actually, the outcome might even prove more severe since the cascading delay effects exacerbate the total reactionary delay. That is, each delayed flight can disrupt subsequent downstream flights. Beatty et al. (1998) evaluated the flight delay propagation through Amer-
ican Airlines schedule, and found that the snowball effect amplifies with the size of the primary delay. For example, initial delays of 30 min and 1 h at 5:00 AM induce 1 h and 3 h in total flight delay, respectively. Therefore, it is of crucial importance to develop strategies to reduce delay propagation and mitigate the snowball effect. Wu (2005) observes that most disruptions in airline operations are unpredictable and stochastic in nature and that schedules may become robust and reliable, only if buffer times are embedded and designed properly in airline schedules. This is precisely the aim of model presented hereafter. This model is based on reallocating existing free slacks in a given planned schedule. In this context, the free slack time (or, “slack” for short) refers to the amount of time that an inbound flight can slip without delaying the start of the immediate successor outbound flights and while maintaining resource feasibility. Clearly, if the slacks are sufficiently large then they can absorb primary delays and therefore cancel reactionary delays. Nevertheless, a schedule that exhibits many large slacks is likely to result in a poor resource utilization.

In next section, we shall propose a model that aims at generating robust solutions through distributing slacks to connections where they are most needed operationally. In particular, we focus on mitigating the propagation of minor delays in a network. The significance of our model stems from the tremendous (yet surprising) economic impact of small-scale delays that account for 60% of total annual delay costs in the United States (Shavell, 2000). To effectively spread the slack times across connecting flights, we also propose to revise the flights departure times. To that aim, each flight is allowed to slightly deviate from its planned departure time while maintaining the designated time-slot assignments at airports, the expected demand and consequently the corresponding revenue, and the planned crew schedules. Consequently, we consider narrow flight departure time-windows, with a specified discrete set of alternative departure times. Thus, we propose an integrated model that simultaneously routes aircraft, revise flight departure times, and allocates slacks to connections while taking into account the minimum time for passenger connections and aircraft preparation. At this point, it is worth emphasizing that these decisions do not incur additional costs, while maintaining the corresponding revenue, since they neither affect the aircraft fleet assignment nor the crew schedules. Indeed, an important feature of the proposed model is that it considers as an input an optimized fleet assignment and then builds robust aircraft routes without altering the solution’s cost.

### 3.2. Slack-based robustness measures

In this section, we propose quantitative metrics that provide a good estimate of schedule robustness. These metrics will be used in the objective function of the proposed model that builds robust schedules by judiciously allocating slacks to vulnerable connections. For each pair of consecutive flights $j$ and $k$ that share a same resource (i.e. aircraft or passengers), the slack represents the scheduled connection time between flights $j$ and $k$ minus a minimum time for connecting between these flights. We remind here that the robustness of a schedule is related to its ability to absorb unanticipated disruptions and therefore mitigate the snowball effect. Moreover, a schedule can be considered as a set of consecutive connections that share a common resource (aircraft and/or passengers). Thus, the robustness of a schedule stems from the robustness of its consecutive connections, which in turn depends on the amount of the connections’ slacks. For some critical connections, large slacks are required whereas, for other connections, little slacks can be enough to maintain the departure time of its subsequent flight when disruptions occur.

For each pair of connecting flights, we define a function $g(\cdot)$, hereafter referred to as local robustness, that maps the slack into a scalar in [0,1] expressing the corresponding robustness. We propose to consider a concave piecewise linear function that captures the following observations (see Fig. 3):

- If the slacks are sufficiently large then they can absorb primary delays and therefore cancel reactionary delays. Hence, increasing the slack between two consecutive flights mitigates the snowball effect through improved delay absorption capability and therefore results in a more robust schedule.
- However, the marginal utility of the slacks decreases as slack increases until it becomes null when the slack exceeds some specified threshold value.

![Fig. 3. The robustness function.](image-url)
In practice, two to four segments would be sufficient to have a reasonable approximation of this slack/robustness relationship. In our implementation, we empirically defined a piecewise linear function having three breakpoints $\tau_1$, $\tau_2$ and $\tau_3$ that are defined as follows:

- $\tau_1$ corresponds to the estimated mean delay of flight $j$. We set $g(\tau_1) = 0.6$ for all pairs of consecutive flights sharing a same resource (aircraft or passengers).
- $\tau_2$ and $\tau_3$ correspond to the 80th and 95th percentile, respectively. We set $g(\tau_2) = 0.9$ and $g(\tau_3) = 1$ for all pairs of consecutive flights sharing a same resource (aircraft or passengers).

Hence, for a slack $s$ between $j$ and $k$ we have:

$$g(s) = \begin{cases} 
  g(\tau_1) \frac{s}{\tau_1} & \text{if } 0 \leq s \leq \tau_1 \\
  g(\tau_1) + (g(\tau_2) - g(\tau_1)) \frac{s-\tau_1}{\tau_2-\tau_1} & \text{if } \tau_1 < s \leq \tau_2 \\
  g(\tau_2) + (g(\tau_3) - g(\tau_2)) \frac{s-\tau_2}{\tau_3-\tau_2} & \text{if } \tau_2 < s \leq \tau_3 \\
  1 & \text{if } s > \tau_3
\end{cases}$$

We notice here that the number of breakpoints and their values are parameters of the robustness metrics and should be fixed in consistence with the historical data on delays. However, in our experiments we found that the robustness metrics are not sensitive to minor changes on breakpoints’ values.

Moreover, since some connections are particularly critical, we propose to assign a weight to each connection that reflects its specific criticality (or relative priority). For instance, we may assign a large weight to those flights that are scheduled early in the morning and whose delay might negatively impact a large number of subsequent flights. In doing so, the schedule planner is offered the opportunity to transfer to the model, via selecting appropriate weights, some information pertaining to relatively predictable large primary delays over certain specific connections. Aggregated historical data on the expected delays, estimated revenues generated by connecting passengers can be used to set the corresponding weights. Therefore, we define for each pair of connecting flights $j$ and $k$ sharing a same resource a weight $w_{jk}$. A surrogate measure of the robustness that is associated to a given resource (aircraft or passengers) is

$$\sum_{(j,k) \in CF} W_{jk}g_{jk}(S_{jk})$$

where $CF$ is the set of pairs of consecutive flights that share the resource, $S_{jk}$ represents the slack between flights $j$ and $k$, and $g_{jk}(\cdot)$ is the local robustness measure associated to connection $(j, k)$.

Further notation will be given in the next section, to develop more precisely these two surrogate robustness measures.

4. A MIP model for building robust schedules

In this section, we present a mixed-integer programming network-flow model that aims at maximizing an aggregated surrogate measure of robustness through appropriately revising the flights departure times and inserting slacks across the connecting flights to provide protection against possible delays. More precisely, the model determines the departure time of each scheduled flight as well as the specific route of each individual aircraft. Prior to describing the proposed model, we shall introduce the notation.

4.1. Notation

Let $F$ be a set of $m$ available aircraft with $\varphi$ different aircraft types, in which all aircraft have the same seating configuration. Denote by $F_f \subseteq F$, the subset of aircraft that belong to aircraft type $f$ and by $I_h$ their corresponding turn-time ($f = 1, \ldots, \varphi$). Let $L$, indexed by $j$ and $k$, denotes a set of $n$ flights to be flown during the planning horizon (typically, one day). Each flight $j \in L$ is defined by its origin and destination stations, its duration $d_j$ and a set of possible departure times $I_p$. If departure time $p \in I_p$ is selected, then the aircraft will take off at $t_{pf}$ and will land at $t_{pf} + d_j$. Moreover, we define $L_f \subseteq L$ as the set of flights $j \in L$ that are assigned to aircraft type $f$. Let $S$, indexed by $s$, be the set of stations. In order to take into account the aircraft origin location, define, for each station $s$ and each aircraft type $f$, two sets of zero-duration dummy flights $L_f^{(s)}$ and $L_f^{(s)}$ that correspond to flight landing at and departing from station $s$, respectively. For each flight $j \in L_f^{(s)}$ ($j \in L_f^{(s)}$), we associate a single departure time that corresponds to the beginning (the end) of the planning horizon. Moreover, for each station $s$, we define $A_{S}$ and $B_{S}$ as the sets of flights that are assigned to aircraft of type $f$, and that arrive to and depart from station $s$, respectively. These sets include the dummy flights $L_f^{(s)}$, $L_f^{(s)}$, respectively.

For each pair of flights $(j, k) \in L_f \times L_f$ that are assigned to the same aircraft type $f$ while having compatible stations (that is, the arrival station of flight $j$ coincides with the departure station of flight $k$), we define $I_{jk}$ as the set of compatible departure time pairs $(p, q)$, $p \in I_p$, $q \in I_q$ (that is, those that allow an aircraft to serve flight $k$ departing at $t_{qf}$ immediately after flight $j$ departing at $t_{pf}$). If the connecting flights $j$ and $k$, that are assigned to an aircraft of family $f$, depart at times $p$ and $q$, respectively (with $(p, q) \in I_{jk}$), then the corresponding utile slack $\sigma_{pq}$ is given by
In our experiments we set
\[
\sigma_{pq}^j = \min \left\{ t_j, \tau^k_j - \left( t_j^k + \delta_i + \theta_j \right) \right\}, \quad \forall f = 1, \ldots, \varphi, \ j \in L_f, \ k \in L_f, \ (p, q) \in I_{jk}
\]  
(3)

For each pair of connecting flights \( j \) and \( k \) that are assigned to the same aircraft, let \( g_{jk}(\cdot) \) be the local robustness measure and \( w_{jk} \) its weight according to Eq. (1). Let \( \tau^{j,1}_{jk}, \tau^{j,2}_{jk} \) and \( \tau^{j,3}_{jk} \) be the three breakpoints of \( g_{jk}(\cdot) \). Hence, the associated robustness is given by
\[
\gamma_{jk} = w_{jk} g_{jk}(\sigma_{pq}^j). \quad \forall f = 1, \ldots, \varphi, \ j \in L_f, \ k \in L_f, \ (p, q) \in I_{jk}
\]  
(4)

To account for the flow of connecting passengers, we denote by \( T \) the set of connecting flights (in practice, \( T \) may be restricted to the set of key passenger connections), and \( \varepsilon_{jk} \) the estimated minimum time for passengers connecting between flights \( j \) and \( k \), for \( (j, k) \in T \). Furthermore, we assume that for each pair \( (j, k) \in T \) the corresponding local robustness measure \( h_{jk}(\cdot) \) is a piecewise linear function, according to Eq. (1), having three breakpoints \( \tau^{j,1}_{jk}, \tau^{j,2}_{jk} \) and \( \tau^{j,3}_{jk} \) that are defined in a similar way as \( \tau^{1,4}_{jk}, \tau^{2,4}_{jk} \) and \( \tau^{3,4}_{jk} \), respectively.

Let \( y_{jk} \) denote the util slack between flights \( j \) and \( k \) \( ((j, k) \in T) \). That is, we have
\[
y_{jk} = \min \left\{ \tau^{k,3}_{jk}, \tau^k - (t_j + \delta_j + \varepsilon_{jk}) \right\}, \quad \forall (j, k) \in T
\]  
(5)

where \( t^k \) and \( t^j \) represent the selected departure times of flights \( j \) and \( k \), respectively, while ensuring that \( y_{jk} \geq 0 \). Hence, we have
\[
h_{jk}(y_{jk}) = \begin{cases} 
      h_{jk}\left( \tau_{jk}^{k,1} \right) y_{jk} & \text{if } 0 \leq y_{jk} \leq \tau_{jk}^{k,1} \\
      h_{jk}\left( \tau_{jk}^{k,2} \right) + \left( h_{jk}\left( \tau_{jk}^{k,1} \right) - h_{jk}\left( \tau_{jk}^{k,2} \right) \right) \frac{y_{jk} - \tau_{jk}^{k,2}}{\tau_{jk}^{k,1} - \tau_{jk}^{k,2}} & \text{if } \tau_{jk}^{k,1} < y_{jk} \leq \tau_{jk}^{k,2} \\
      h_{jk}\left( \tau_{jk}^{k,2} \right) + \left( h_{jk}\left( \tau_{jk}^{k,3} \right) - h_{jk}\left( \tau_{jk}^{k,2} \right) \right) \frac{y_{jk} - \tau_{jk}^{k,2}}{\tau_{jk}^{k,3} - \tau_{jk}^{k,2}} & \text{if } \tau_{jk}^{k,2} < y_{jk} \leq \tau_{jk}^{k,3} \\
      1 & \text{if } y_{jk} > \tau_{jk}^{k,3}
\end{cases}
\]  
(6)

In our experiments we set \( h_{jk}\left( \tau_{jk}^{k,1} \right) = 0.6, \ h_{jk}\left( \tau_{jk}^{k,2} \right) = 0.9 \) and \( h_{jk}\left( \tau_{jk}^{k,3} \right) = 1.0 \).

Let \( y_{jk}^1 = \min \left( y_{jk}, \tau_{jk}^{k,1} \right), \ y_{jk}^2 = \max \left( 0, \min \left( y_{jk} - \tau_{jk}^{k,1}, \tau_{jk}^{k,2} - \tau_{jk}^{k,1} \right) \right) \) and \( y_{jk}^3 = \max \left( 0, \min \left( \tau_{jk}^{k,2} - \tau_{jk}^{k,1}, \tau_{jk}^{k,3} - \tau_{jk}^{k,2} \right) \right) \) then we can restate (6) in the following equivalent form:
\[
h_{jk}(y_{jk}) = \frac{0.6}{\tau_{jk}^{k,1}} y_{jk}^1 + \frac{0.3}{\tau_{jk}^{k,2} - \tau_{jk}^{k,1}} y_{jk}^2 + \frac{0.1}{\tau_{jk}^{k,3} - \tau_{jk}^{k,2}} y_{jk}^3
\]  
(7)

Fig. 4. Computation of slack \( y_{jk} \) and associated values \( y_{jk}^1, y_{jk}^2 \) and \( y_{jk}^3 \).
Let $v_{jk}$ denote the weight associated with the pair $(j, k) \in T$, we set $y_{jk}^1 = \frac{0.6v_{jk}}{t^k_{jk}}$, $y_{jk}^2 = \frac{0.3v_{jk}}{t^k_{jk}}$ and $y_{jk}^3 = \frac{0.1v_{jk}}{t^k_{jk}}$.

Fig. 4 illustrates computation of slack $y_{jk}$ and associated values $y_{jk}^1$, $y_{jk}^2$, and $y_{jk}^3$. For example, for $t^1_{jk} = 15\text{ min}$, $t^2_{jk} = 30\text{ min}$ and $t^3_{jk} = 50\text{ min}$, if the slack $y_{jk} = 25\text{ min}$ then $y_{jk}^1 = 15\text{ min}$, $y_{jk}^2 = 10\text{ min}$ and $y_{jk}^3 = 0$.

### 4.2. Model formulation

To formulate the robust aircraft routing and flight retiming problem, we propose a network flow-based model using aircraft-connection variables. Alternative formulations have been presented in the literature. They usually define a string as a sequence of connected flight legs that begins and ends at maintenance stations (possibly different ones) and then use assignment models that are solved with Branch-and-Price algorithms (Lan, 2003; Lan et al., 2006; Dunbar et al., 2012). In addition to the $y$ slack variables, we introduce the following further notation:

- $x_{jk}^p$: binary decision variable that takes the value 1 if the departure time $p \in I_j$ is selected for flight $j \in L$, and 0 otherwise.
- $z_{pq}^k$, $f = 1, \ldots, q$, $(j, k) \in L_f \times L_k$, $(p, q) \in I_{k}$: binary decision variable that takes the value 1 if the same aircraft is assigned to flight $k$ departing at time $q$ immediately after flight $j$ departing at time $p$, and 0 otherwise.
- $y_{jk}^1$, $y_{jk}^2$ and $y_{jk}^3$: $(j, k) \in T$: continuous decision variables that define the utility slack between passengers connecting flights $j$ and $k$.

Fig. 5 illustrates some notations and decision variables $x_{jk}^p$ and $z_{pq}^k$ using an example of a station $s \in S$. In this figure, Beginning and End are two dummy flights that correspond to the beginning and the end of the planning horizon respectively. $L_f^{(S)} = \{\text{Beginning}\}$, $L_k^{(S)} = \{\text{End}\}$, $A_f = \{\text{Beginning}, f_1, f_2\}$ and $B_k = \{f_3, f_4, \text{End}\}$.

The four flights have all three possible departure times. For flight $f_1 : x_{jk}^1 = 1$ and $x_{jk}^2 = x_{jk}^3 = 0$. If a same aircraft of type $f$ is assigned to flight $f_3$ immediately after flight $f_1$ (as displayed in Fig. 5) then $z_{jk}^1 = 1$.

Using these definitions, the robust aircraft routing and flight retiming problem can be formulated using a network flow formulation as follows:

\[
\text{(RARFR): Maximize } \sum_{j, k \in T} \left( z_{jk}^1 y_{jk}^1 + z_{jk}^2 y_{jk}^2 + \alpha y_{jk}^3 \right) \quad + \sum_{j, k, l, m} \sum_{(p, q) \in I_{k}} \sum_{(s, t) \in I_{k}} \sum_{(u, v) \in I_{k}} z_{pq}^k y_{jk}^1 \quad \tag{8}
\]

subject to:

\[
0 \leq y_{jk}^1 \leq t^1_{jk}, \quad (j, k) \in T \quad \tag{9}
\]

\[
0 \leq y_{jk}^2 \leq t^2_{jk} - t^1_{jk}, \quad (j, k) \in T \quad \tag{10}
\]

\[
0 \leq y_{jk}^3 \leq t^3_{jk} - t^2_{jk}, \quad (j, k) \in T \quad \tag{11}
\]

\[
\sum_{p \in I_j} x_{jk}^p = 1, \quad j \in L \quad \tag{12}
\]

\[
\sum_{j, k \in T} \sum_{(p, q) \in I_{k}} z_{pq}^k = \chi^k_q, \quad f = 1, \ldots, q, \quad s \in S, \quad k \in B_f^q, \quad q \in I_{k} \quad \tag{13}
\]

\[
\sum_{j \in B_f^q} \sum_{(p, q) \in I_{k}} z_{pq}^k = \chi^p_f, \quad f = 1, \ldots, q, \quad s \in S, \quad j \in A_f^q, \quad p \in I_j \quad \tag{14}
\]

\[
x_{jk}^p \in \{0, 1\}, \quad j \in L, \quad p \in I_j \quad \tag{15}
\]

\[
z_{pq}^k \in \{0, 1\}, \quad f = 1, \ldots, q, \quad j \in L_f \cup_{k \in S} L_f^{(S)}, \quad k \in L_f \cup_{k \in S} L_k^{(S)}, \quad (p, q) \in I_k \quad \tag{16}
\]

\[
y_{jk}^1, y_{jk}^2, y_{jk}^3 \geq 0, \quad (j, k) \in T \quad \tag{17}
\]

The objective function (8) maximizes the sum of two robustness measures. While the first term pertains to passenger connections, the second one refers to aircraft connections. In constraint (9), $\sum_{q \in I_{k}} z_{pq}^k$ represents the departure time $t^k$ of flight $k$. Constraints (9)-(12) define the utility slack for passenger connections. In particular, Constraint (9) enforces that after retiming the passengers can still follow their planned connections. Note that the utility slack for connection $(j, k) \in T$ is $y_{jk} = y_{jk}^1 + y_{jk}^2 + y_{jk}^3$. Since $y_{jk} > y_{jk}^2$ then we have $y_{jk}^2 > 0 \Rightarrow y_{jk}^1 = t_{jk}^1$. Similarly, since $y_{jk}^2 > y_{jk}^3$ then we have $y_{jk}^3 > 0 \Rightarrow y_{jk}^2 = t_{jk}^2$. Constraints (13) ensure that exactly one departure time will be assigned to each flight. Constraints (14) and (15) require that on each aircraft route, each flight has exactly one successor and one predecessor, respectively.
Remark 1. Observe that for given $x$ binary, Constraints (14) and (15) exhibit an assignment structure. Therefore, (17) can be relaxed to simply $z^k_{pj} \geq 0$. This leads to a formulation that is amenable to be solved using Benders decomposition. This may be useful for solving larger problems.

Remark 2. In Model RARFR, the maintenance constraints are not accommodated. Nevertheless, for the (frequent) case where the flight network includes one single hub where all maintenance checks are achieved overnight, the flights start and end in the hub and in this case, our model is particularly suitable. This simplifies the optimization but the model still makes sense in practice.

In our computational experiments, we found that Model RARFR is relatively easy to solve and that large instances can be optimally solved using a general-purpose MIP solver.

5. Computational study

To provide a proof-of-concept of the tractability and relevance of the proposed model and robustness measure, we carried out a computational study. In this section, we describe the set of instances. Next, we will investigate the model tractability. Finally, we will present the results of a Monte Carlo simulation study that demonstrates that the solutions yielded by the model are robust.

5.1. Generation of the problem instances

We considered two real-world instances hereafter referred to by A and B, respectively. These instances were previously proposed by Amadeus, SAS for the ROADEF Challenge 2009. Both instances correspond to one-day schedules. Table 1 displays the characteristics of both instances.

Starting from instance A, we built six instances $A_1, \ldots, A_6$ that are obtained by defining for each leg $j \in L$ a time-window $[a_j - \delta, a_j + \delta]$ where $a_j$ represents the genuine departure time and parameter $\delta \in \{0, 5, 10, 15\}$ represents half of the time-window width. Note that instance $A_0$ (that corresponds to $\delta = 0$) refers to the genuine schedule. In Instance $A_i$, the flights’ departure times are similar to those in $A_0$, but aircraft routes are not fixed and will be computed by RARFR model. Next, we define for each instance $A_i$ ($i = 2, \ldots, 6$) a number $v_i$ of alternative departure times (within the time-window) that are evenly spaced by $r$ minutes (with $r \in \{5, 10, 15\}$). Similarly, we used the same procedure to build six instances $B_1, \ldots, B_6$ that are derived from the genuine instance B. Table 2 displays a summary of the characteristics of the generated scenarios.

To complete the instances construction, we generated the weights that are used in the proposed robustness measure according to the number of passengers. In so doing, we generated a total number of 14 instances (including the genuine schedules). Finally, recall that in our model, the breakpoints of the local robustness measure should be defined using the distribution of flight’s delays. However, since these distributions were not available to us, we used the same robustness function for all pairs of flights, and we set for all instances $\tau^{\text{mk}}_1 = 15$ min, $\tau^{\text{mk}}_2 = 30$ min, and $\tau^{\text{mk}}_3 = 50$ min for all aircraft connections and for all passenger connections.

5.2. Computational tractability of Model RARFR

Model RARFR was solved using CPLEX 12.2 solver with the default setting. All our experiments were run on an i5 dual core 2.3 GHz Personal Computer with 4 GB RAM. A summary of the results are displayed in Table 3. This table provides the CPU

time in seconds, the total robustness of the optimal solution and the percentage robustness improvement with respect to the fixed flight schedule (scenarios \(A_0\) and \(B_0\)).

Looking at Table 3, a first striking observation is that, despite their relatively large size, all problem instances were optimally solved within reasonable CPU times. Indeed, the largest instance \((B_6)\) required less than 790 s. Hence, an appealing feature of Model \(RARFR\) is that it exhibits a structure that is amenable to be quickly solved without resorting to a tailored optimization algorithm.

Not surprisingly, we also observe that the CPU time increases as the number of alternative departure times increases. However, we see that while the increase of the computational effort is significant, the corresponding objective function improvement is only marginal (see for example the third and the fourth scenarios as well as the fifth and the sixth scenarios for both instances).

5.3. How robust are the generated solutions?

To assess the robustness of the solutions derived by Model \(RARFR\), we carried out a Monte Carlo simulation study. To that aim, we use the following 4-step methodology.

**Step 1: Random generation of primary delays.** To the best of our knowledge, there is no public available data set that includes a description of the primary delay distribution (and not all of the delays). Thus, primary delays were randomly generated to carry out the simulation study. We make the following assumptions.

**A1** The primary delay distribution is unimodal, strongly assymetric, and skewed toward short delays.

**A2** Primary delays lie within a bounded interval \([\xi_{\text{min}}, \xi_{\text{max}}]\) (that is, \(\xi_{\text{min}}\) and \(\xi_{\text{max}}\) are the minimum and maximum primary delays, respectively).

Therefore, a natural candidate for the primary delay distribution is beta distribution. Each delayed flight \(j\) is assigned a delay \(\xi_j\) that is computed as follows:

\[
\xi_j = (\xi_{\text{max}} - \xi_{\text{min}}) r_j + \xi_{\text{min}}
\]

where \(r_j \in [0, 1]\) is randomly drawn from beta distribution \(Be(\alpha, \beta)\). The distribution parameters \(\alpha\) and \(\beta\) are set as follows.

Recall that the mean \(\mu\) and the mode \(\lambda\) of the beta distribution are given by

\[
\mu = \frac{\alpha}{\alpha + \beta} \quad \text{and} \quad \lambda = \frac{\alpha - 1}{\alpha + \beta - 2}
\]

Hence, given two estimates \(\xi_{\mu}\) and \(\xi_{\lambda}\) of the mean and mode of the primary delays, we derive the corresponding parameters \(\alpha\) and \(\beta\).

In our experimentations, starting from the genuine instance \(A\) (\(B\)), we delay the departure time of each scheduled flight with a probability of \(p_d = 0.15\) (this value is consistent with probabilities of primary delays used by Düeck et al. (2012)). Furthermore, we used the following parameter setting: \(\xi_{\text{min}} = 5\) min, \(\xi_{\text{max}} = 80\) min, \(\xi_{\mu} = 25\) min, and \(\xi_{\lambda} = 15\) min. With this setting, we obtain \(\alpha = 1.467\) and \(\beta = 4.033\). The resulting delay distribution is displayed in Fig. 6.
Step 2: Generation of reactionary delays. If flight $j$, that is assigned to an aircraft of family $f$, is delayed by $n_j$ then the subsequent downstream flight $k$ cannot depart earlier than $d_j + h_f + n_j$. In this case, flight $k$ might be delayed by up to $n_k = \max (t_j + d_j + h_f + n_j, 0)$. Hence, using a right-shift recovery strategy (also referred to as push-back), that delays a leg until its scheduled plane is ready, we can sequentially compute the updated departure times of all the legs. Clearly, more involved recovery strategies might be used. However, we have focused on the (simple) right-shift strategy because in cases of delays of short duration it is often used in practice (Schaefer et al., 2005).

Step 3: Computation of performance measures. We evaluate the robustness of each solution $S(I)$ of instance $I$ ($I \in \Sigma = \{A_i, B_i; i = 0, \ldots, 6\}$) by computing the following performance measures:

- $PM_1$: Total reactionary delay, when the delay is more than 15 min.
- $PM_2$: Number of flights delayed by more than 15 min.
- $PM_3$: Number of passengers delayed by more than 15 min.
- $PM_4$: Number of passengers missing their connections.

Steps 1–3 are replicated $N$ times. In our implementation, we set $N = 100$.

Step 4: Correlation analysis. For each solution $S(I)$ of instance $I \in \Sigma$, we compute the average performance measure $PM(I)$ ($I = 1, \ldots, 4$). The results are displayed in Table 4.

Let $Z(I)$ denote the value of the objective function of $S(I)$. We measure the strength of the relationship between the surrogate robustness measure and the performance measure by computing the correlation coefficient between $Z(A_i)$ and $PM(I)$ ($I = 1, \ldots, 6, I = 1, \ldots, 4$). We analyze in a similar way the correlation between $Z(B_i)$ and $PM(I)$ ($I = 1, \ldots, 6, I = 1, \ldots, 4$). The results are displayed in Table 5.

Looking at these tables, we can make the following observations:

- The correlation analysis reveals that there is a strong negative correlation between the proposed surrogate robustness measure and all considered performance measures. In particular, we observe that $R^2 = -0.9996$ for $PM_1$ and the instances of Classes $A$ and $B$. Therefore, there is a strong empirical evidence that the value of the objective function increases as the performance measure decreases.

Table 3
Results of the robust aircraft routing and flight retiming model.

<table>
<thead>
<tr>
<th>Instance</th>
<th>Scenario</th>
<th># Of departure times</th>
<th>$r$ (min)</th>
<th>$\delta$ (min)</th>
<th>CPU (s)</th>
<th>Total robustness</th>
<th>% Of improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$A_0$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>–</td>
<td>21974.97</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$A_1$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.48</td>
<td>24450.75</td>
<td>11.27</td>
</tr>
<tr>
<td></td>
<td>$A_2$</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>6.65</td>
<td>27391.28</td>
<td>24.65</td>
</tr>
<tr>
<td></td>
<td>$A_3$</td>
<td>3</td>
<td>10</td>
<td>10</td>
<td>6.46</td>
<td>29928.80</td>
<td>36.19</td>
</tr>
<tr>
<td></td>
<td>$A_4$</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>35.45</td>
<td>30123.78</td>
<td>37.08</td>
</tr>
<tr>
<td></td>
<td>$A_5$</td>
<td>3</td>
<td>15</td>
<td>15</td>
<td>6.39</td>
<td>30520.88</td>
<td>38.89</td>
</tr>
<tr>
<td></td>
<td>$A_6$</td>
<td>7</td>
<td>5</td>
<td>15</td>
<td>124.97</td>
<td>30805.76</td>
<td>40.19</td>
</tr>
<tr>
<td>B</td>
<td>$B_0$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>–</td>
<td>59194.35</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>$B_1$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3.81</td>
<td>62693.32</td>
<td>5.91</td>
</tr>
<tr>
<td></td>
<td>$B_2$</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>60.62</td>
<td>72986.55</td>
<td>23.30</td>
</tr>
<tr>
<td></td>
<td>$B_3$</td>
<td>3</td>
<td>10</td>
<td>10</td>
<td>60.53</td>
<td>79299.35</td>
<td>33.96</td>
</tr>
<tr>
<td></td>
<td>$B_4$</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>287.76</td>
<td>79575.55</td>
<td>34.43</td>
</tr>
<tr>
<td></td>
<td>$B_5$</td>
<td>3</td>
<td>15</td>
<td>15</td>
<td>62.81</td>
<td>83331.63</td>
<td>40.78</td>
</tr>
<tr>
<td></td>
<td>$B_6$</td>
<td>7</td>
<td>5</td>
<td>15</td>
<td>786.91</td>
<td>84051.33</td>
<td>41.99</td>
</tr>
</tbody>
</table>
The optimized schedules exhibit good performance and consistently outperform the genuine ones. For instance, if we compare $S(A_1)$ to $S(A_0)$ and $S(B_1)$ to $S(B_0)$, we remark that the number of delayed flights, and the number of delayed passengers was divided by a factor of around 10. This ratio is more than 25 and 124 for the number of passengers missing their connections and the total reactionary delay, respectively. This proves that aircraft routing has a major importance on the stability of the generated schedules. Besides, when allowing retiming flights’ departure dates, the performance measures $PM_i$, $i = 1, \ldots, 4$, are divided by a factor greater than 2.29 for solution $S(A_6)$ with respect to solution $S(A_1)$, and by a factor greater than 3.15 for solution $S(B_6)$ with respect to solution $S(B_1)$. This is a clear indication of the ability of Model RARFR to effectively generate robust solutions that are less vulnerable to disruptions.

### 6. Conclusion

In this paper, we addressed the issue of building robust aircraft routes that are less vulnerable to delays. We presented slack-based quantitative metric to measure the ability of a schedule to absorb delays. Next, we developed a mixed-integer programming formulation that enables to determine the flights departure times and aircraft's routes while maximizing the proposed surrogate robustness measure. We presented the results of computational experiments carried out on real-world instances with up to 1278 flights and 251 aircraft. These results show that the proposed model is efficiently solvable using a commercial MIP solver. Also, we assessed the robustness of the proposed solutions through a Monte Carlo simulation. The results of this simulation provide strong empirical evidence that the optimized schedules are robust and thereby improving the on-time performance. A nice feature of the proposed model is that only departure times and aircraft routes might be revised and not the fleeting decisions. Therefore, the increased schedule’s ability to absorb disruptions is achieved with no significant additional cost. Hence, the proposed model can be usefully used at a post-optimization step (that is, after generating an optimized flight schedule and fleet assignment) to generate robust plans.

Recognizing the interplay of recovery process with the planning process, it is desirable to further integrate in the model recovery decisions together with the corresponding costs (e.g. swapping aircraft, canceling or diverting flights, etc.). This is the subject of our ongoing research.

### References


### Table 4

Average performance measures.

<table>
<thead>
<tr>
<th>Instance</th>
<th>Scenario</th>
<th>PM1: Total reactionary delay (min)</th>
<th>PM2: Number of delayed flights</th>
<th>PM3: Number of delayed passengers</th>
<th>PM4: Number of passengers missing their connections</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$A_0$</td>
<td>100004.95</td>
<td>286.9</td>
<td>21708.79</td>
<td>1145.54</td>
</tr>
<tr>
<td></td>
<td>$A_1$</td>
<td>800.42</td>
<td>28.22</td>
<td>2337.46</td>
<td>43.8</td>
</tr>
<tr>
<td></td>
<td>$A_2$</td>
<td>591.81</td>
<td>21.75</td>
<td>1774.17</td>
<td>27.63</td>
</tr>
<tr>
<td></td>
<td>$A_3$</td>
<td>384.96</td>
<td>14.34</td>
<td>1152.25</td>
<td>18.19</td>
</tr>
<tr>
<td></td>
<td>$A_4$</td>
<td>379.14</td>
<td>14.25</td>
<td>1152.78</td>
<td>18.88</td>
</tr>
<tr>
<td></td>
<td>$A_5$</td>
<td>345.88</td>
<td>12.91</td>
<td>1057.68</td>
<td>16.08</td>
</tr>
<tr>
<td></td>
<td>$A_6$</td>
<td>324.48</td>
<td>12.27</td>
<td>999.1</td>
<td>17.43</td>
</tr>
<tr>
<td>B</td>
<td>$B_0$</td>
<td>268612.57</td>
<td>772.18</td>
<td>59979.08</td>
<td>4490.57</td>
</tr>
<tr>
<td></td>
<td>$B_1$</td>
<td>1812.76</td>
<td>64.91</td>
<td>5192.41</td>
<td>179.51</td>
</tr>
<tr>
<td></td>
<td>$B_2$</td>
<td>1166.32</td>
<td>43.79</td>
<td>3438.91</td>
<td>108.38</td>
</tr>
<tr>
<td></td>
<td>$B_3$</td>
<td>770.38</td>
<td>29.69</td>
<td>2257.83</td>
<td>80.35</td>
</tr>
<tr>
<td></td>
<td>$B_4$</td>
<td>760.66</td>
<td>29.4</td>
<td>2222.26</td>
<td>79.31</td>
</tr>
<tr>
<td></td>
<td>$B_5$</td>
<td>538.37</td>
<td>21.13</td>
<td>1575.24</td>
<td>52.18</td>
</tr>
<tr>
<td></td>
<td>$B_6$</td>
<td>518.57</td>
<td>20.56</td>
<td>1508.26</td>
<td>51.84</td>
</tr>
</tbody>
</table>

### Table 5

Correlation coefficients between robustness measures and performance measures.

<table>
<thead>
<tr>
<th>Instance</th>
<th>PM1: Total reactionary delay (min)</th>
<th>PM2: Number of delayed flights</th>
<th>PM3: Number of delayed passengers</th>
<th>PM4: Number of passengers missing their connections</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$-0.9996$</td>
<td>$-0.9982$</td>
<td>$-0.9987$</td>
<td>$-0.9888$</td>
</tr>
<tr>
<td>B</td>
<td>$-0.9996$</td>
<td>$-0.9987$</td>
<td>$-0.9987$</td>
<td>$-0.9965$</td>
</tr>
</tbody>
</table>