Optimally mitigating risk for airlines

Gary Froyland

School of Mathematics and Statistics
University of New South Wales, Sydney

UNSW-CSIRO Workshop on Risk:
Modelling, Optimization, Inference
11–12 December 2014

Joint work with
Michelle Dunbar (Wollongong) and **Richard Wu** (UNSW)

Airline delays caused by propagation in complex networks

Motivation: Robustness

Historically, the construction of airline schedules has been undertaken with a view towards maximizing profit.

- This usually results in brittle schedules with minimal slack time between connections.
- This seeks to minimise planned cost rather than the *realised cost*.
- Delays have a tendency to propagate rapidly throughout the network.

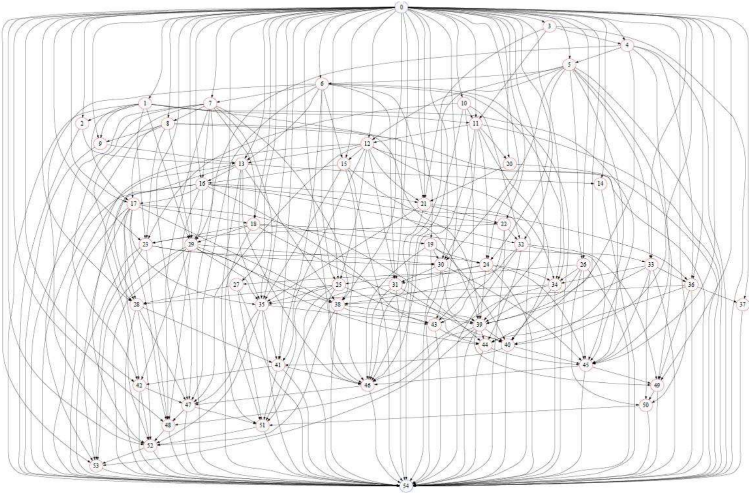
– In 2000, 30% of flight legs operated by a U.S. airline were delayed (Lan *et al.* (2006)).

– It was estimated that in 2006, the U.S. airline industry experienced **116.5 million minutes of delay** $\Rightarrow$ **\$7.7 billion increase in operating costs** (AhmadBeygi *et al.* (2008)).

Connection networks

- Let $G = (\mathcal{N}, \mathcal{A})$ be a directed acyclic graph with a single source node s, and a single terminal node t.
- The source and terminal nodes are dummy nodes that link to both the morning and evening flights, respectively.
- In this graph, nodes correspond to flights and arcs correspond to possible feasible connections between flight nodes.

A flight connection network



Flight connections for 53 flights in an Australian network. Crew connections are described by similar networks.

Primary delays and slack

- Each connection $(i, j) \in \mathcal{A}$, will have associated with it two *primary delays*.
- The total **primary delay** for aircraft connection (i, j) is denoted p_{ij}^R and is the sum of the expected en-route delay for flight i (estimated from historical data), and primary delays during aircraft turnaround operations, such as passenger connection delay, and ground handling delay.
- The **slack** s_{ij} for a connection (i, j) is the difference between the scheduled arrival time of flight i and the scheduled departure time of flight j, minus the mean turn-around time for the relevant aircraft type under the specific ground handling procedure of the airline.

Propagated delay

- We now come to the *propagated delay* at node i, denoted d_i.
- We fix the initial delay at the source node $d_s = 0$ and inductively apply the formulae below to calculate propagated delay along a path in the aircraft connection network:

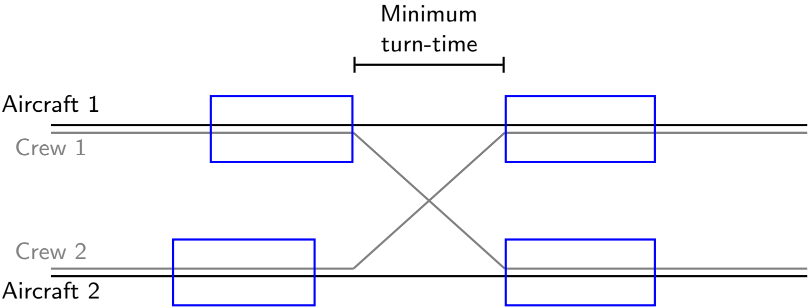
$$d_j^R = \max\left\{d_i^R - (s_{ij} - p_{ij}^R), 0\right\}, \quad j \neq s, \tag{1}$$

and in the crew connection network:

$$d_j^P = \max\left\{d_i^P - (s_{ij} - p_{ij}^P), 0\right\}, \quad j \neq s. \tag{2}$$

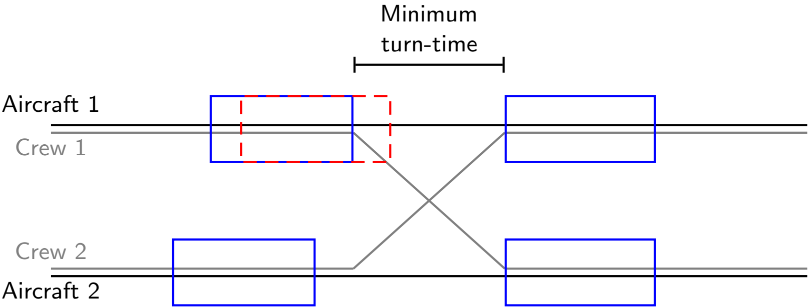
Example: Propagation of Delay from Crew to Aircraft

The crew pairing may contribute to the delays associated with aircraft routing (Weide *et al.* (2010)).



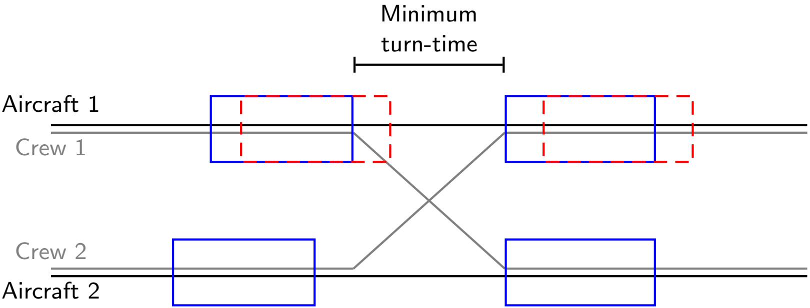
Example: Propagation of Delay from Crew to Aircraft

The crew pairing may contribute to the delays associated with aircraft routing (Weide *et al.* (2010)).



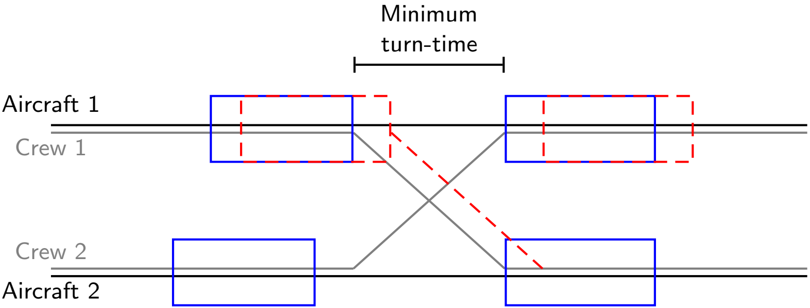
Example: Propagation of Delay from Crew to Aircraft

The crew pairing may contribute to the delays associated with aircraft routing (Weide *et al.* (2010)).



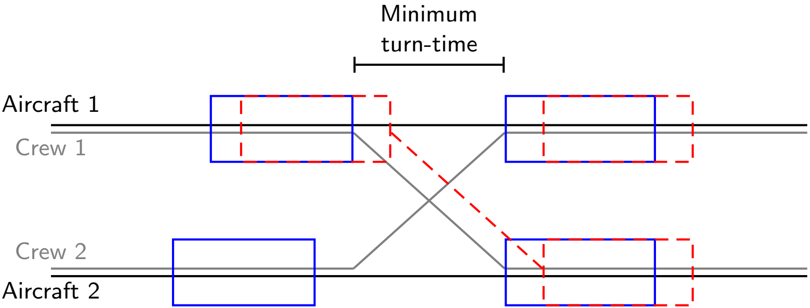
Example: Propagation of Delay from Crew to Aircraft

The crew pairing may contribute to the delays associated with aircraft routing (Weide *et al.* (2010)).



Example: Propagation of Delay from Crew to Aircraft

The crew pairing may contribute to the delays associated with aircraft routing (Weide *et al.* (2010)).



The effect of crew delay on aircraft delay

To calculate the propagated delay along an aircraft string, **taking into account propagated delays from crew** we inductively apply:

$$d_j^R = \max\left\{d_i^R - (s_{ij} - p_{ij}^R), d_k^P - (s_{kj} - p_{kj}^P), 0\right\}, \quad j \neq s, \quad (3)$$

where the connection (i, j) is part of the aircraft string and the connection (k, j) is part of the crew string that includes flight j.

Delay evaluation in a combined network

Propagated Delay Evaluation: Carry out the following evaluations in a single topologically ordered sweep through the connection networks.

1. Using the strings from the incumbent routing and crew pairing solutions, update d_j^R inductively along each string in the incumbent routing solution using
$$d_j^R = \max\left\{d_i^R - (s_{ij} - p_{ij}^R), d_k^P - (s_{kj} - p_{kj}^P), 0\right\}.$$
2. Using the strings from the incumbent routing and crew pairing solutions, update d_j^P inductively along each string in the incumbent crew pairing solution using
$$d_j^P = \max\left\{d_i^P - (s_{ij} - p_{ij}^P), d_k^R - (s_{kj} - p_{kj}^R), 0\right\}.$$

Mean primary delay and scenarios of primary delay

1. The first part of the talk discusses propagated delay reduction when only the mean primary delay for each connection is known.
2. The second part of the talk considers the situation where scenarios of primary delay are known or can be reasonably generated and shows further improvements that can be made in total propagated delay reduction.

USING MEAN PRIMARY DELAY INFORMATION

An "ideal" integrated model

With our understanding of delay propagation, an ideal integrated model might look like:

$$
\begin{aligned}
\text{Minimise:} \quad & (\mathbf{c}^R)^T \mathbf{x}^R + (\mathbf{c}^P)^T \mathbf{x}^P \\
\text{Subject to:} \quad & A^R \mathbf{x}^R = \mathbf{e} \\
& A^P \mathbf{x}^P = \mathbf{e} \\
& \sum_{i=1}^{n_R} x_i^R \leq N \\
& \sum_{i=1}^{n_P} x_i^P \leq M \\
& \mathbf{x}^R \in \{0,1\}^{n_R}, \mathbf{x}^P \in \{0,1\}^{n_P}
\end{aligned} \qquad (4)
$$

where c_j^R is the total propagated delay cost along routing string j and c_j^P is the total propagated delay cost along pairing string j.

Problem: $\mathbf{c}^R$ depends on $\mathbf{x}^P$ and $\mathbf{c}^P$ depends on $\mathbf{x}^R$.

Related work

- Lan *et al.* (2006) develop robust aircraft routings, minimising expected delay. Propagated delay is used for evalution of column costs, but not for column selection. Crew interactions are not considered.
- Weide *et al*. (2010) propose an iterative integrated aircraft routing and crew pairing model, which attempts to reduce the number of restricted (short) connections and minimise a non-robustness measure.
- Borndörfer *et al*. (2010) consider robust aircraft routings, minimising delay based on iid primary flight delay distributions.
- Work on integrating retiming includes Desaulniers *et al.*, Rexing *et al.*, Klabjan *et al.*, Lohatepanont *et al.*, Samardi, Belanger *et al.*, AhmadBeygi *et al.*, Papadakos.

Our first approach

Our aim is to improve upon the work of Lan *et al.*, AhmadBegi *et al.*, and Weide *et al.* We use the iterative theme of Weide *et al.*, and alternate between aircraft routing and crew pairing master problems until no further reduction in total propagated delay is observed. The key ingredients of our approach are:

1. the accurate calculation of the **combined effects** of propagation of delay along aircraft routing strings **and** crew pairing strings,
2. the use of this information for both the calculation of the **cost** of columns and the dynamic **selection** of optimal columns, and
3. the embedding of these calculations in an iterative scheme for the master routing and crew pairing problems, to optimise the **true cost** due to total propagated delays of aircraft and crew.

Iterative Algorithm

1. INITIALISATION:
   1. Create incumbent routing and crewing solutions.
   2. For each arc $(i,j) \in \mathcal{A}$, independently randomly sample a large number of scenarios, consisting of primary delays p_{ij}^R and p_{ij}^P from a given probability distribution.
   3. Set $d_k^P = 0$ for all $k \in \mathcal{N}$ and $d_s^R = 0$. Set an counter $c = 0$.
2. MINIMUM DELAY AIRCRAFT ROUTING:
   1. Apply Propagated Delay Evaluation Algorithm.
   2. Assign expected delay costs to aircraft strings; (not only do we use the existing crew strings, but we also estimate the **crew delay impact** of including a particular aircraft string in the restricted master basis).
   3. Solve master routing problem, minimising the total delay cost to produce a new incumbent routing solution.

Iterative Algorithm (cont...)

3 MINIMUM DELAY CREW PAIRING:

1. Apply Propagated Delay Evaluation Algorithm.
2. Assign delay costs to crew strings; (not only do we use the existing routing strings, but we also estimate the **routing delay impact** of including a particular crew string in the restricted master basis).
3. Solve master crew pairing problem, minimising the total delay cost to produce a new incumbent crew pairing solution.

4 If either the aircraft routing or crew pairing solutions have changed, increment counter $c \to c + 1$ and return to Step 2. Otherwise, STOP.

The pricing problem: Routing

- We assume that for every unit of time an aircraft (resp. crew) is late at node i a dollar cost $a_i^R > 0$ is incurred.
- Thus, for the aircraft routing pricing problem we wish to find a path $\pi = \{s, i_1, i_2, \dots, t\}$ from s to t that minimises

$$z^R = \min \left\{ \sum_{i \in \pi} \left(a_i^R d_i^R + w_i^R \right) \; : \; \pi \text{ is a path from } s \text{ to } t \right\},$$

  where the weight $-w_i^R$ is the dual multiplier for constraint i in the master problem.
- Because the delay d_i^R **not a simple sum of delays** along the path from s to i, the routing pricing problem is not easily cast as a minimum cost network flow.
- We use a label setting algorithm, augmented by a notion of label dominance, modified from related problems in Desrochers and Soumis (1988) and Dumitrescu and Boland (2003), that works efficiently in the cases tested.

Numerical Results

- We apply our algorithms to a one-day schedule on a real airline network consisting of 53 flights and 126 feasible connections.
- We determine that the minimum number of aircraft and crew pairs required to cover this network are 10 and 16, respectively.
- The mean primary aircraft and crew pairing delays p_{ij}^R and p_{ij}^P are randomly sampled from four different probability distributions.
- To capture the asymmetric nature of the aircraft and crew delays, we sample from an exponential distribution $E(\lambda)$ with mean $1/\lambda$ in minutes and a truncated normal distribution (truncated to non-negative delays), denoted $tN(\mu, \sigma^2)$ with mean μ and variance σ^2, in minutes and minutes^2.
- Each instance consists of at least 1000 scenarios (1000 random copies of 126 primary connection delays).

Numerical Results

We study two simplified approaches in addition to our base case and proposed approach:

1. **Base (B):** - Propagated Delay Evaluation applied to initial "min cost" routing and crewing solution.
2. **Routing and Crewing Solved Sequentially, Simple Delay (SSD):** - One iteration of the Iterative Algorithm where we remove the "max" in the delay propagation estimates so that $d_j^R = d_i^R - (s_{ij} - p_{ij}^R)$ and $d_j^P = d_k^R - (s_{kj} - p_{kj}^R)$. Apply Propagated Delay Evaluation Algorithm.
3. **Routing and Crewing Solved Sequentially, Propagated Delay (SPD):** - One iteration of the Iterative Algorithm. Apply Propagated Delay Evaluation Algorithm.
4. **Routing and Crewing Integrated, Propagated Delay (IPD):** - Iterative Algorithm.

Numerical Results

- The SPD approach will demonstrate the value of calculating the more accurate, nonlinear, *propagated delay* over the simpler, less accurate linear delay of the SSD approach.
- Our proposed IPD approach will demonstrate the value of *integrating* routing and crewing, rather than simply performing them sequentially as in the SPD approach.

A "typical" result

Exponential distribution with mean $\lambda = 10$.

Instance 4:

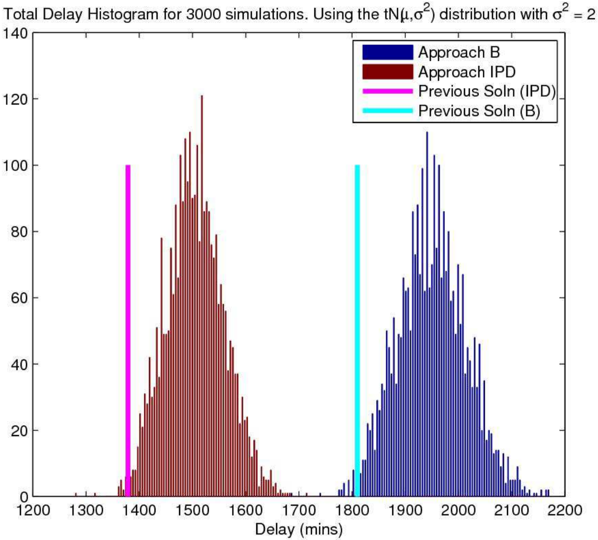
Approach	Aircraft Delay (mins)	Crew Delay (mins)	Total Delay Delay (mins)	% improv.
B	704	1749	2453	0.00
SSD	633	1176	1809	26.25
SPD	610	1079	1689	31.15
IPD	540	1079	1619	34.00

Numerical Results

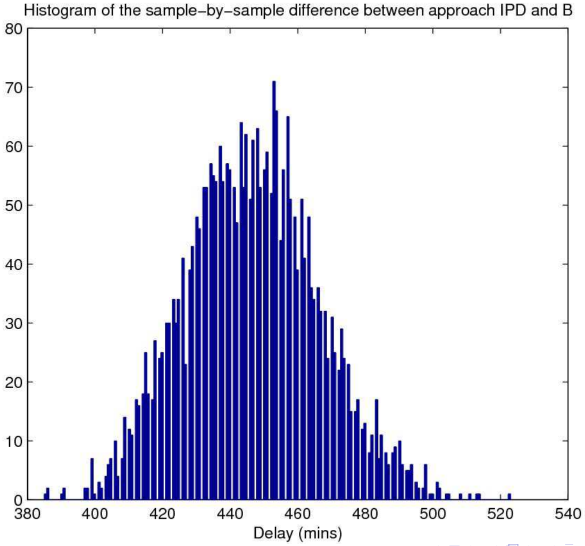
Table : Relative improvements of the algorithms SSD, SPD, and IPD.

Instance	(SSD-SPD)/SSD *100%	(SPD-IPD)/SPD *100%	(SSD-IPD)/SSD *100%
1	30.03	6.97	34.90
2	−4.38	11.69	7.82
3	15.14	4.32	18.80
4	6.63	4.14	10.50
5	1.66	9.96	11.45
6	−20.72	21.04	4.69
7	−0.71	3.87	3.19
8	13.52	5.13	17.95
9	−2.68	4.96	2.41
10	−1.83	4.77	3.03
11	4.16	2.83	6.87
12	−1.95	6.17	4.34
Average	3.24	7.15	10.50

Results from simulations: 53 flights (3000 simulations) $tN(5, 2)$



Results from simulations: Delay reductions on 53 flights (3000 simulations) $tN(5,2)$



Contributions of our Integrated Scheduling Model

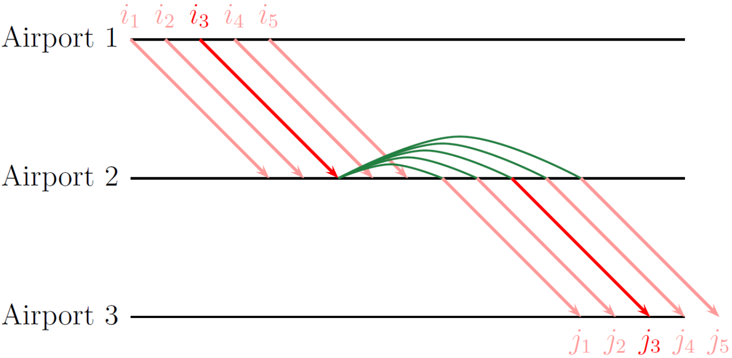
The key contributions of our Integrated Model are:

1. The **accurate calculation of propagated delay** throughout the aircraft and crew networks.
2. Calculating the **combined effects** of delays from aircraft and crew.
3. Attempting to choose aircraft and crew connections that **minimise the overall combined propagated delay**.

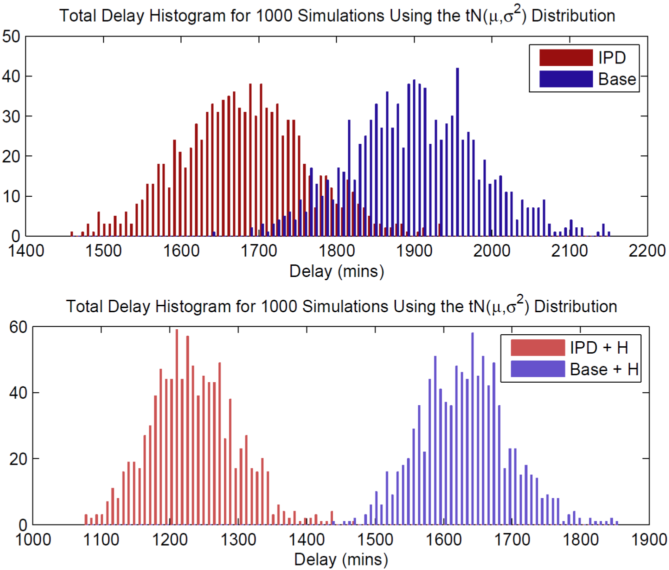
USING RETIMING

Retiming algorithms and other combinations

1. We apply a simple greedy retiming algorithm, which adjusts flight departure times in 5 minute increments in a ±10-minute window, and calculates expected propagated delay.
2. This retiming algorithm is applied after the combined routing/crewing optimisation.



Improvements over base with retiming



USING SCENARIO INFORMATION TO OPTIMIZE ROUTING AND CREWING

Use of scenario information

We have three methods of incorporating scenarios into the pricing.

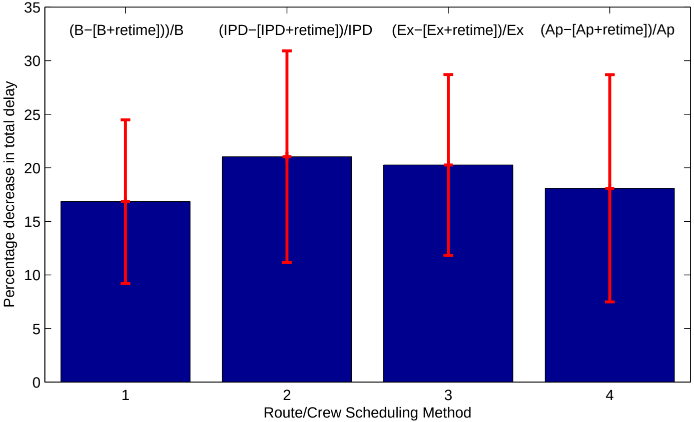
1. We use the **mean value of primary delay** for each connection and price using a label setting algorithm to calculate propagated delay. We denote this method "**IPD**". [Dunbar *et al.* 2012] showed an average improvement of 8.91% over the method of Weide *et al.* This was achieved by accurately estimating and explicitly optimising propagated delay.
2. We exhaustively construct all possible paths in the connection network, and **individually price each path for each delay scenario**, then use the scenario average of these prices to produce a price for each path. Then select the path with the minimal cost. We denote this method "**Ex**".
3. We incorporate a **local estimate of mean propagated delay based on scenario information** into the label setting algorithm; this provides better accuracy than using the mean value with almost no extra computational cost. We denote this method "**Ap**".

Retiming algorithms with scenario information and other combinations

1. We now apply the retiming algorithm after the combined routing/crewing optimisation using scenario information. **[Alg+retime]**
2. We also re-apply the combined routing/crewing optimisations after retiming to obtain further improvements. **[Alg+retime+Alg]**
3. These are compared with a simultaneous optimisation of retiming included in the combined routing/crewing optimisation. **[AlgSretime]**

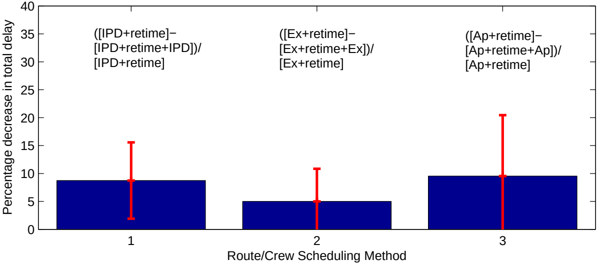
Improvements achievable with the retiming algorithm

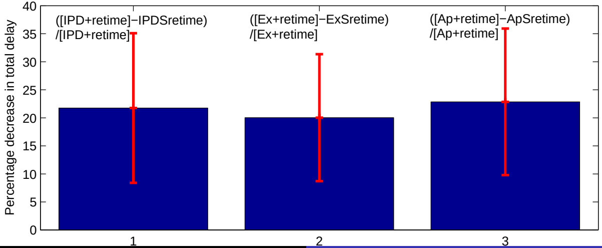
We plot the percentage delay reductions achieved by applying the retiming algorithm to the Base, IPD, Ex, and Ap solutions.



[Alg+retime+Alg] vs [AlgSretime]

There are benefits in rescheduling after retiming, but much greater benefits to integrating retiming in the scheduling algorithms.





In summary

1. **Integrating routing and crewing** gave an average 10.5% decrease in total delay over sequential optimisation.
2. **Retiming** produced around an additional 18-20% decrease.
3. **Using scenario information** in the column pricing produced an additional 2-4% decrease.
4. **Simultaneous retiming with routing/crewing optimisation** produced a 22.8% reduction over retiming after routing/crewing optimisation.

Summary of contributions

1. The **accurate calculation of propagated delay** throughout the aircraft and crew networks.
2. Calculating the **combined effects** of delays from aircraft and crew in the pricing and column selection steps.
3. **Incorporating retiming**, in a way that also accurately uses scenario information.
4. **Utilising scenario information** in the column pricing and selection.
5. Choosing aircraft and crew connections, that **minimise the overall combined propagated delay**.

See [Dunbar/F/Wu, *Transportation Science*, 2012] for the work on mean primary delays (items 1,2,5).
See [Dunbar/F/Wu, *Computers and Operations Research*, 2014] for the work on using scenario information, the scenario-based pricing methods, and incorporating retiming (items 3,4,5).