Ongoing Work: Column Generation for the Aircraft Recovery Problem with Operational Restrictions

David Torres Sanchez

1. Modelling

To formulate the aircraft recovery (AR) problem, we introduce the following notation. Let $\mathcal{J}$ be the set of all flights, indexed by j. Let $\mathcal{K}$ be the set of aircraft, indexed by k. Let $\mathcal{S}$ be the set of schedules, indexed by s, with each schedule having a certain cost, c^s, associated with it. Each schedule is composed of a series of flights, which can be denoted with $\mathcal{J}^s$, hence, $j \in \mathcal{J}^s$ means that flight j is present in schedule s. We can define the cost of a schedule as the sum of the costs of the flights that compose it. Flight costs take into account operational and delay costs; see Section 1.3.1 for more information.

Let the parameter d^s_{jk}, be equal to 1 if flight j is assigned to aircraft k in schedule s, and it is 0 otherwise. Finally, we define an assignment variable, x^s_k gets the value 1 if schedule s is assigned to aircraft k, while it is 0 otherwise.

1.1. Definitions

Sets

$\mathcal{J}$: Set of flights indexed by j, with $\mathcal{J} = \{1, \dots, J\}$;
$\mathcal{K}$: Set of all aircraft indexed by k, with $\mathcal{K} = \{1, \dots, K\}$;
$\mathcal{S}$: Set of schedules indexed by s.
$\mathcal{J}^s$: Set of flights in schedule s, with $\mathcal{J}^s \subseteq \mathcal{J}$;

Parameters

c^s : Cost of schedule s;
d^s_{jk} : Parameter equal to 1 if flight j is assigned to aircraft k in schedule s, 0 otherwise.

Variables

x^s_k : 1, if aircraft k is assigned to schedule s, 0 otherwise.

1.2. Aircraft Recovery Problem Formulation

Using the notation defined, let us formulate the AR problem as follows,

Model 1.1. *AR using a set covering formulation.*

$$\min \sum_s \sum_k c^s x_k^s \tag{1}$$

Subject to

$$\sum_s \sum_k d_{jk}^s x_k^s \geq 1 \qquad \forall\ j; \tag{2}$$

$$\sum_s x_k^s \leq 1 \qquad \forall\ k; \tag{3}$$

$$x_k^s \in \{0, 1\} \qquad \forall\ k,\ s; \tag{4}$$

Objective 1 minimises the total cost of assigning aircraft to schedules. Constraints 2 ensure that each flight is assigned to at least one aircraft. Constraints 3 ensure that each aircraft is assigned to at most one schedule. Constraints 4 define the domain of the binary assignment variable.

To efficiently solve this problem, we can employ column generation. On the average case, column generation prevents from having to enumerate all schedules ($2^J - 1$).

1.3. Column Generation

To apply column generation, let us formulate a restricted and relaxed version of Model 1.1. For this purpose, we solve the problem on a reduced number of schedules. In every iteration, i, of the column generation scheme, we introduce a new schedule s, i.e. a column. The set of schedules $\mathcal{S}^i$ is then updated as $\mathcal{S}^i = \mathcal{S}^{i-1} \cup \{s\}$ (for $i \geq 1$). The corresponding schedule costs, c^s, are set appropriately (as discussed in Section 1.3.1). For the first iteration, we employ schedules $\mathcal{S}^0$, where each schedule contains a single and distinct flight. That is, for $s \in \mathcal{S}^0$, $\bigsqcup_s \mathcal{J}^s = \mathcal{J}$. Additionally, to avoid these single flight schedules to appear in the solutions, we set each schedule cost, c^s, to a large flight cancellation cost. For a given iteration i of the column generation scheme, we can use schedules $s \in \mathcal{S}^i$, to formulate the restricted and relaxed version of Model 1.1 as follows,

Model 1.2. *Restricted and relaxed AR problem.*

$$\min \sum_s \sum_k c^s x_k^s \tag{5}$$

Subject to

$$\sum_s \sum_k d_{jk}^s x_k^s \geq 1 \qquad \forall\ j; \tag{6}$$

$$\sum_s x_k^s \leq 1 \qquad \forall\ k; \tag{7}$$

$$x_k^s \geq 0 \qquad \forall\ k,\ s; \tag{8}$$

Aside from using a subset of schedules, the only other change with respect to Model 1.1, is the linear relaxation of the binary variable with constraints 8. Let π_j and μ_k be the dual variables for constraints 6 and 7, respectively. Hence, the reduced cost, for a given aircraft k, schedule s, is given as,

$$\overline{c}_k^s = c^s - \sum_j d_{jk}^s \pi_j + \mu_k \ . \tag{9}$$

Column generation involves solving a subproblem that will provide the column with the most negative reduced cost, or equivalently, finding a schedule s such that,

$$\overline{c}_k^* = \min_s \{\overline{c}_k^s\} = \min_s \left\{ c^s - \sum_j d_{jk}^s \pi_j \right\} + \mu_k \ . \tag{10}$$

Hence, generating a column for aircraft k corresponds to solving a subproblem and creating a schedule s with its corresponding flights. Specifically, each subproblem is modelled by employing the shortest path problem with resource constraints (SPPRC). Given that airlines' operational factors can be taken on an individual aircraft basis, we can model them in the subproblems as independent resource constraints.

1.3.1. Subproblems: Shortest Path Problem with Resource Constraints

To solve the subproblems for each aircraft k, given in Equation 10, we make use of the SPPRC. To model the subproblems accurately, let us define a network with some specific properties. Let $G = (N, A)$ be directed activity-on-arc or a TSN where the set of nodes N represents the airports under consideration at different time points, and the arcs, A, correspond to connections. Let G have a single source and sink nodes with no incoming or outgoing connections respectively. For two nodes $i, j \in N$ there is an arc, i.e connection, $g = (i, j) \in A$, if and only if it corresponds to a scheduled connection (flight or ground with $g \in \mathcal{J}$), or a non-scheduled connection ($g \notin \mathcal{J}$). We distinguish two types of non-scheduled connections: a **flight delay copy**, a delayed copy of a scheduled flight; and a **non-scheduled ground connection**. See Section 2.1 for more details.

For any connection g, let w_g denote the weight, h_g and t_g denote the head and tail nodes (respectively), a_g and d_g denote the arrival and departure times (respectively), r_g denote the flight range, and lastly, let T_g denote the type of connection (scheduled: flight or ground connection, non-scheduled: flight delay copy or ground connection).

Additionally, every connection has a corresponding operational cost. For a subproblem k and any scheduled connection, the operating cost is defined to be proportional to a standard operating cost $c_k^{\mathtt{st}}$. Where $c_k^{\mathtt{st}}$ is either, the standard total operating cost (per unit time) for aircraft k, if the connection corresponds to a flight, or 0 otherwise. The costs associated with non-scheduled connections (per unit time) are, $c_k^{\mathtt{od}}$ for flight delay copies (operating costs under delay), and $c_k^{\mathtt{gw}}$ for ground connections (ground waiting costs). See Appendix A for the values used in the computational tests.

To define resource consumption and constraints, let $\mathcal{R} = \{1, \ldots, R\}$ be a set of resources. Let $\underline{\mathbf{L}} = (\underline{L}^1, \ldots, \underline{L}^R)$ and $\overline{\mathbf{L}} = (\overline{L}^1, \ldots, \overline{L}^R)$ be vectors for minimum and maximum resources, respectively. For each connection g, we denote the associated weight by w_g, and the resource consumption vector by $\mathbf{f}_g = (f_g^1, \ldots, f_g^R)$. Each component in this vector is referred to as a resource extension function (REF) [6]. For a given path, p, we denote the set of connections by $A(p)$, the weight of the path by

$w(p)$, and the resource consumed along the path by $\mathbf{f}(p) = (f^1(p), \ldots, f^R(p))$, where,

$$w(p) = \sum_{g \in A(p)} w_g \ \ ; \ f^r(p) = \sum_{g \in A(p)} f_g^r \ . \tag{11}$$

To define the weight w_g, for an arc g and subproblem k, we use,

$$w_g = \begin{cases} c_k^{\mathtt{st}}(\mathrm{a}_g - \mathrm{d}_g) - \pi_g & \text{if } g \in \mathcal{J} \\ c_k^{\mathtt{st}}(\mathrm{a}_g - \mathrm{d}_g) + c_k^{\mathtt{od}}(\mathrm{d}_g - \mathrm{d}_j) - \pi_j & \text{if } g \notin \mathcal{J} \wedge COND_0 \\ c_k^{\mathtt{gw}}(\mathrm{a}_g - \mathrm{d}_g) & \text{if } g \notin \mathcal{J} \wedge COND_0' \end{cases} ; \tag{12}$$

where π_g is the dual variable for constraints 7. Here, condition $COND_0$ is true only if T_g corresponds to a flight delay copy of a scheduled flight j, where,

$$j = \sup \{ j \in \mathcal{J} : \ h_j = h_g, \ t_j = t_g, \text{ and } \mathrm{d}_j < \mathrm{d}_g \} \, . \tag{13}$$

Condition $COND_0'$ is true only if T_g corresponds to a non-scheduled ground connection. Hence, Equation 12 defines the weights of scheduled and non-scheduled connections separately. For a given subproblem k and any scheduled connection g, the weight is proportional to the standard operating cost for aircraft k ($c_k^{\mathtt{st}}$) times the duration of the connection, minus the corresponding dual value. Note that this may lead to negative weights. For a given subproblem k and any non-scheduled connection g, if the connection corresponds to a flight delay copy, then the weight is equal to the standard operating cost ($c_k^{\mathtt{st}}$) times the duration of the connection, plus the operating cost under delay ($c_k^{\mathtt{od}}$) times the delay with respect to the original flight, minus the dual value of the original flight. If the non-scheduled connection corresponds to a non-scheduled ground connection, then the weight is equal to the ground waiting cost ($c_k^{\mathtt{gw}}$) times the duration of the connection.

Using these definitions, generating a column for each aircraft k corresponds to a schedule t produced by the connections of a path produced by solving an appropriate SPPRC. Such path is one that minimises delay while satisfying operational restrictions. That is, finding a path in G, p, which minimises $w(p)$ subject to resource constraints $\underline{\mathbf{L}} \leq \mathbf{f}(p) \leq \overline{\mathbf{L}}$, i.e. $\underline{L}^r \leq f^r(p) \leq \overline{L}^r$ for every resource r. To evaluate the column, the reduced cost can be calculated by updating the parameters d_{jk}^t, and using Equation 9. If the reduced cost is negative, the column is added with a cost c^t. To minimise operational costs, c^t is set to be equal to the total operating costs of the schedule.

1.3.2. Resource Extension Functions

To take airline operational factors into account, we define appropriate resource extension functions (REFs) employed in the SPPRC. For this purpose, it is convenient to define some attributes for connections and aircraft. Each connection g, in a TSN, G, has attributes regarding the arrival and departure airports (h_g, t_g), departure and arrival times (d_g, a_g), and flight range (r_g). Each aircraft k has an attribute r_k that represents the range that a certain aircraft type can operate. Let us define a set of seven resources,

$$\mathcal{R} = \{\mathtt{mono}, \mathtt{air}, \mathtt{maint}, \mathtt{crew\text{-}d1}, \mathtt{crew\text{-}d2}, \mathtt{pass}, \mathtt{non\text{-}s}\} \ ,$$

corresponding to: an artificial monotone resource; operational restrictions: aircraft, maintenance, crew , and passenger delay; and an additional restriction for non-scheduled connections.

We can define REFs for each resource. For extending partial path p_i (a path from source to node i) along arc $g = (i, j)$, resulting in partial path p_j (a path from source to node j passing through node i), we can set the resource consumption vector, $\mathbf{f}(p_j)$, according to the following REFs.

For the artificial monotone resource,

$$f^{\texttt{mono}}(p_j) = f^{\texttt{mono}}(p_i) + 1 \tag{14}$$

For aircraft restrictions,

$$f^{\texttt{air}}(p_j) = \begin{cases} 1 & \text{if not } COND_1 \\ 0 & \text{otherwise} \end{cases}; \tag{15}$$

where condition $COND_1$ depends on the range of the aircraft under consideration, r_k. If r_k is short-haul or long-haul, then $COND_1$ is true only if r_g corresponds to the same range. If r_k is short/medium-haul, then $COND_1$ is true only if r_g is at most a medium-haul flight.

For maintenance restrictions,

$$f^{\texttt{maint}}(p_j) = \begin{cases} f^{\texttt{maint}}(p_i) & \text{if } h_g = t_g \\ 0 & \text{if } COND_2 \\ f^{\texttt{maint}}(p_i) + (\mathrm{a}_g - \mathrm{d}_g) & \text{otherwise} \end{cases}; \tag{16}$$

where condition $COND_2$ is true only if the origin and destination airports, $h_g = t_g$, corresponds to a hub airport and the turnaround time, $\mathrm{a}_g - \mathrm{d}_g$, is at least long enough for the shortest type of maintenance allowed. We denote this quantity with `maintMin`.

For crew duty restrictions 1 and 2,

$$f^{\texttt{crew-d1}}(p_j) = \begin{cases} f^{\texttt{crew-d1}}(p_i) + (\mathrm{a}_g - \mathrm{d}_g) & \text{if } h_g \neq t_g \\ f^{\texttt{crew-d1}}(p_i) & \text{if not } COND_2' \\ 0 & \text{if } COND_2' \end{cases}; \tag{17}$$

and

$$f^{\texttt{crew-d2}}(p_j) = \begin{cases} f^{\texttt{crew-d2}}(p_j) + (\mathrm{a}_g - \mathrm{d}_g) & \text{if } h_g \neq t_g \\ f^{\texttt{crew-d2}}(p_i) & \text{if not } COND_3 \\ 0 & \text{if } COND_3 \end{cases}. \tag{18}$$

Where condition $COND_2'$ is true only if the arrival and departure airport, $h_g = t_g$, corresponds to a hub airport. Condition $COND_3$ is true only if the origin and destination airports, $h_g = t_g$, corresponds to a hub airport and the turnaround time, $\mathrm{a}_g - \mathrm{d}_g$, is at least long enough for the shortest crew duty break. We denote this quantity with `breakMin`.

For the passenger delay restrictions,

$$f^{\,\texttt{pass}}(p_j) = \begin{cases} 1 & \text{if } COND_4 \\ 0 & \text{otherwise} \end{cases} ; \tag{19}$$

where condition $COND_4$ is true only if $g \notin \mathcal{J}$, corresponding to a flight delay copy of flight j (defined by Equation 13), and either of the following hold: if r_g is short-haul, $\mathrm{d}_g - d_j \geq 2$; if r_g is medium-haul, $\mathrm{d}_g - d_j \geq 3$; or if r_g is long-haul, $\mathrm{d}_g - d_j \geq 4$.

Lastly, for non-scheduled connection restrictions,

$$f^{\,\texttt{non-s}}(p_j) = \begin{cases} 1 & \text{if } g \in \mathcal{J} \\ f^{\,\texttt{non-s}}(p_i) - 1 & \text{otherwise} \end{cases} . \tag{20}$$

With these definitions, we can set the minimum and maximum resource levels for each resource as,

$$\underline{\mathbf{L}} = (0,0,0,0,0,0,0),\ \overline{\mathbf{L}} = (|A|, 0, \mathrm{FH}, \mathrm{CD1}, \mathrm{CD2}, 0, 1)\ , \tag{21}$$

Recall that the entries are in the same order as the resources, $\mathcal{R}$. This enforces the following restrictions, for each resource r,

$r =$ `mono`**:** the artificial monotone resource increases by 1 every time an arc is traversed, with an upper bound equal to the number of arcs in the network;

$r =$ `air`**:** aircraft can only perform flights that match with their type. Clearly, this can be included as part of the preprocessing stage, however, we incorporate it as this allows the modelling of more complex aircraft type constraints;

$r =$ `maint`**:** maintenance increases with flying hours unless either the path extension corresponds to a ground connection or a maintenance stop is done, in which case it is set to 0. A maintenance stop is one done at a hub airport with enough time to perform at least the shortest maintenance. The upper bound, FH, corresponds to the maximum flying hours allowed by regulations;

$r =$ `crew-d1`**:** crew duty rule 1, the maximum shift duration, must not exceed CD1;

$r =$ `crew-d2`**:** crew duty rule 2, the maximum flying hours without rest, must not exceed CD2;

$r =$ `pass`**:** passenger delay cannot exceed the limits imposed by regulations;

$r =$ `non-s`**:** enforces that no two consecutive non-scheduled connections are in a path. A scheduled flight sets the resource to 1, otherwise, it is set to its previous value minus 1. Since the minimum for this resource is 0, and two consecutive non-scheduled connections will produce a value of -1, the required relation is enforced.

See Appendix A for the values of FH, CD1, and CD2 used in the computational tests.

2. Solution Methodology

2.1. Network Generation

Flight schedules allow us to construct the necessary TSN network for the subproblems. The TSN network, G, as defined in Section 1.3.1, is constructed as follows. A node is included for every airport under consideration at each different time point. We include two artificial nodes that represent the source and sink nodes, with no incoming or outgoing arcs respectively. Arcs are split by scheduled flights/ground arcs (those present in the initial flight schedule), and non-scheduled connections which we generate. Non-scheduled connections are generated between all airports with a prior existing flight (flight delay copies) and all ground arcs. Another type on non-scheduled connection are ferry or empty flights (sometimes positioning). However, ferrying is not always considered in the AR problem [1], alternatively, some authors that account for crew recovery considered crew deadhead flights (where crew travel along with passengers), [8], for instance. [7] solved an extension to the TA problem (with maintenance) allowing ferry flights. They reported that the appearance of these types of connections in the solutions occurred only upon the decrease of the allocated ferrying costs (provided by an airline) and resulted in a small increase in profit. [10] solved the AR problem using a TSN and also including ferry flights. Results showed that the number of ferry flights in the solution decrease as the recovery horizon increases. Additionally, the recovery horizons in this work were under 24 hours. Given that the recovery horizons considered in this research are at least 2 days, we do not account for ferrying; nonetheless, flight delays and cancellations are included.

Figure 1 shows a two airport example of a TSN with two different routes. Given the routes in the pre-operational flight schedules (blue and red arcs), non-scheduled connections are generated (dashed arcs) for flight delay copies and ground arcs. Moreover, due to airline restrictions, flight delay copies are not generated prior to the original flight. For instance, a copy for the blue flight B_3 to A_4 is only generated after, creating the flight delay copy B_4 to A_5, but not B_2 to A_3. This creates a cross-hatched pattern between connected airports. Generating delay copies in this fashion, avoids creating additional arcs (as with most of the traditional methods), but fits more naturally into the original pre-operational plans. Clearly, larger networks benefit from more flight delay copies and therefore, more options when creating routes; hence, leading to better quality solutions. On the other hand, small networks provide less flight delay copies and typically correlate with worse quality solutions.

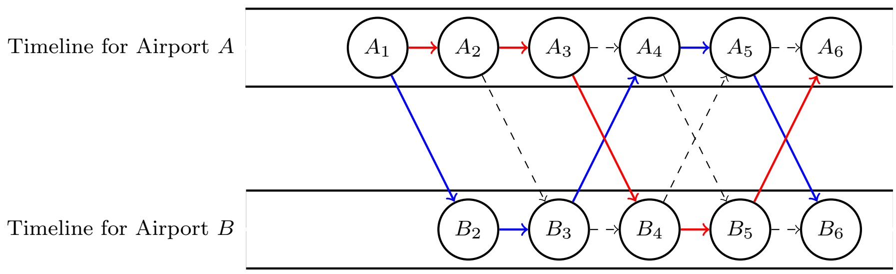

Figure 1. Generation of Time-Space Networks.

References

[1] J. Clausen, A. Larsen, J. Larsen, and N.J. Rezanova, *Disruption management in the airline industry—concepts, models and methods*, Computers & Operations Research 37 (2010), pp. 809–821.

[2] A. Cook and G. Tanner, *European airline delay cost reference values: Updated and extended values*, Tech. Rep., University of Westminster, 2015.

[3] A. Cook and G. Tanner, *Innovative Cooperative Actions of R&D in EUROCONTROL Programme CARE INO III: Dynamic Cost Indexing: Aircraft maintenance – marginal delay costs*, Tech. Rep., Transport Studies Group, University of Westminster, 2008.

[4] FAA, *Benefit-Cost Analysis. Section 4: Aircraft operating costs.*, Tech. Rep., Federal Aviation Administration, 2013.

[5] Federal Aviation Administration, *Flightcrew Member Duty and Rest Requirements*, Federal Aviation Administration (FAA), DOT. (2009). Available at `https://www.faa.gov/regulations_policies/rulemaking/recently_published/media/2120-AJ58-FinalRule.pdf`.

[6] S. Irnich and G. Desaulniers, *Shortest Path Problems with Resource Constraints*, in *Column Generation*, G. Desaulniers, J. Desrosiers, and M.M. Solomon, eds., Springer US, Boston, MA (2005), pp. 33–65.

[7] N. Kenan, A. Jebali, and A. Diabat, *The integrated aircraft routing problem with optional flights and delay considerations*, Transportation Research Part E: Logistics and Transportation Review 118 (2018), pp. 355–375.

[8] S.J. Maher, *A novel passenger recovery approach for the integrated airline recovery problem*, Computers & Operations Research 57 (2015), pp. 123–137.

[9] M. Palpant, M. Boudia, C. Robelin, S. Gabteni, and F. Laburthe, *ROADEF 2009 challenge: Disruption management for commercial aviation*, Sophia Antipolis, France: Operations Research Division (2009).

[10] B.G. Thengvall, G. Yu, and J.F. Bard, *Multiple fleet aircraft schedule recovery following hub closures*, Transportation Research Part A: Policy and Practice 35 (2001), pp. 289–308.

[11] D. Torres Sanchez, B. Boyacı, and K.G. Zografos, *An Optimisation Framework for Airline Fleet Maintenance Scheduling with Tail Assignment Considerations*, Transportation Research Part B: Methodological (n.d.).

Appendix A. Parameters in the Computational Tests

Table A1 reveals the precise aircraft makes and types for each test instance, their respective range, standard total operational costs (as given by 4) and operational costs under delay (as given by 2). For illustration and recalling the notation from Section 1.3.1, for a given subproblem k, with the make of aircraft k being DH8D, then, the standard total operating cost for a scheduled flight is $c_k^{\texttt{st}} = \$922$ (around €823 at the current exchange rate), the cost for non-scheduled ground connections is $c_k^{\texttt{gw}} = €540$, and, the operating cost under delay for flight delay copies is $c_k^{\texttt{od}} = €1630$.

Table A2 summarises the parameters employed in the computational tests. First, the cancellation cost, captured in the costs of the initial schedules in the column generation algorithm (in the notation of Section 1.3, c^s where $s \in \mathcal{S}^0$), were given by [9]. The minimum duration of a crew duty break, `breakMin`, is given by [5]. The minimum duration for a maintenance intervention, `maintMin`, is given by [11]. The remaining parameters concern to the resource limits introduced in Section 1.3.2. These ensure that the appropriate regulations are met. Please note that the values for maintenance (FH) and crew duty rules (CD1, CD2) were given by [3] and [8], respectively.

Table A1. Different aircraft types in test instances with their respective operating costs.

Aircraft Make	Aircraft Type (Range)	Inst.	Operating Costs (per hour)		Operating Costs (per hour)
			Standard ($c_k^{\mathtt{st}}$)	Ground ($c_k^{\mathtt{gw}}$)	Delayed Flight ($c_k^{\mathtt{od}}$)
DH8D	Turboprop (Short-haul)	I1	\$922	€540	€1630
JS32	Turboprop (Short-haul)	I2	\$2887	€540	€1630
A319	Narrow-body (Short/medium-haul)	I1, I4	\$9734	€810	€3420
A320	Narrow-body (Short/medium-haul)	I5	\$9734	€900	€3490
B738	Narrow-body (Short/medium-haul)	I6	\$10430	€1010	€3650
B752	Narrow-body (Short/medium-haul)	I6	\$10430	€720	€4210
A359	Wide-body (Long-haul)	I3, I4	\$13912	€1050	€4130
B77L	Wide-body (Long-haul)	I6	\$13912	€1310	€6230

Table A2. Parameters for computational tests.

Parameter	Representation	Value
c^s $(s \in \mathcal{S}^0)$	Flight cancellation cost (per hour)	€20000
`breakMin`	Minimum crew duty break	3 hours
`maintMin`	Minimum maintenance time	5 hours
FH	Maximum flying hours between maintenance (A check)	600 flying hours
CD1	Maximum crew shift duration (Duty rule 1)	13 hours
CD2	Maximum crew flying hours without rest (Duty rule 2)	8 hours