A COLUMN GENERATION APPROACH TO SOLVE AIRCRAFT RECOVERY PROBLEM

Nan Xu
China Southern Airlines Company Limited
Guangzhou, China
xunan@csair.com

ABSTRACT
It is common that an airline is faced with the necessity of reconstructing their schedules due to disruptions. In this paper, we propose a model of aircraft recovery problem to fulfill the airline constraints. An algorithm based on column generation is also presented to solve the problem, where the master problem is modeled as a set partition problem and the subproblem is modeled as a shortest path problem on a time-space network. Delaying flight, swapping flight and canceling flight are considered in our model. The tradeoff between canceling flight and delaying flight is modeled by the cancel-delay equivalent M and the optimization objective is to minimize the total recovery cost.

KEY WORDS
Airline optimization; Aircraft recovery problem; Disruption management; Column generation

1. Introduction
Airlines need to coordinate a relevant number of resources to provide their service to customers. To maximize the revenue, the resources of airlines are tightly coupled. However, on the day of operation, it is likely that the optimized schedule can become infeasible due to disruptions. The most common reasons for these disruptions are aircraft malfunction, absent crew, bad weather and air traffic control. When disruptions occur, operations personnel must make real-time decisions to bring the airline back to its original schedule as soon as possible. The operation personnel use a combination of the following options to recover from disruptions: delaying, swapping, canceling and ferrying. Ferrying involves flying an empty aircraft to an airport which is short of aircraft, and it is seldom used due to the excessive cost.

Recovery from disruption involves aircraft, crews and passenger itineraries. When only aircraft is involved in the recovery problem, we refer to it as the Aircraft Recovery Problem (ARP). We will focus on ARP in this paper.

2. Literature Survey
Teodorović and Guberinić [1] are the pioneers who study the ARP problem. They developed a graph construct in which nodes represent flights and arcs represent delays. They proposed a branch and bound algorithm and the optimization objective is to minimize the total customer delay by delaying and swapping flights.

Jarrah et al. [2] present two separate models to solve the ARP problem, one by canceling and swapping flights and the other by delaying and swapping flights. Ferrying is taken into account but aircraft maintenance is not considered. The major drawback is that the two models cannot work together.

Argüello et al. [3] present a heuristic to solve ARP, which is called Greedy Randomized Adaptive Search Procedure (GRASP). The heuristic can deal with cancellations, delays and flight swaps. The optimization objective is to minimize the total cost for delays and cancellations. An initial solution is obtained by canceling the affected flights. Three options are used to generate neighboring solutions of the incumbent: flight route augmentation, partial route exchange and simple circuit cancellation. Best neighbors are placed in the restricted candidate list and a new incumbent is randomly selected from the restricted candidate list. The iterations continue until a stopping criterion is encountered. Aircraft maintenance are not considered.

Thengvall et al. [4] propose a model to solve ARP. To minimize the deviations from the original aircraft routings, protection arcs are added in the network. By giving incentive to continuous combinations of the original flight path, deviations from the original aircraft routings are discouraged. Maintenance is not taken into account in this model. Different recovery schedules can be obtained by adjusting the parameters like the number of delay options, the cost of delaying flights and the incentive for protecting flights.

Løve et al. [7] implement a heuristic to solve ARP based on local search. They defined a kind of network where nodes are either aircraft or flights. Arcs connecting an aircraft node and a flight node represent an aircraft to flight assignment. By altering these aircraft-to-flight assignment arcs, the flight schedule is modified. Steepest Ascent Local Search heuristic (SALS) is used to find local optimum. The optimization objective is to maximize the revenue minus cost.

Anderson [8] present a heuristic which can generate a list of ranked non-dominated solutions. Tabu search and simulated annealing are used. Both models are based on a local search algorithm. The neighborhood is defined by selecting a pair of aircraft and generate new route for them.
A pool is formed of the flights of the two aircraft and the flights not assigned to any aircraft. New route for each aircraft is generated from this pool.

Eggenberg et al [9], propose a recovery algorithm based on column generation, where the master problem is modeled as a set partitioning problem and the subproblem is modeled a resource constrained shortest path problem. A dynamic programming algorithm for resource constrained elementary shortest paths is used to solve the pricing problem. Aircraft maintenance is incorporated by introducing maintenance arcs into the network. The optimization objective is to minimize the costs involves flight, delay, and maintenance.

Anderson et al [10] model the ARP as a mixed integer multicmodity flow model with side constraints. And Danzig-Wolfe decomposition is used to reformed the problem as a set packing problem with generalize upper bound (GUB) constraints, which ensure that each aircraft is assigned exactly one route. The problem is solved with a Lagrangian relaxation-based heuristic and column generation.

3. A Mathematical Model of ARP

In this paper, we focus on the aircraft recovery problem, where decisions are taken on the aircraft’s schedule. When disruption occurs, operational personnel make decisions to get back to the original schedule by delaying flights, canceling flights, aircraft swapping and ferrying. A small example is given to illustrate the aircraft recovery problem. In Table 1&2, the original schedule of aircraft a1&a2 are given. At time 1000, after aircraft a1 landed in HGH airport, it is reported to be unavailable from time 1000 to 1600 due to aircraft malfunction. A disruption is occurred and the schedule cannot be carried out as planned.

Assume the minimal turnover for a1 and a2 is 40 minutes. One possible recovery plan can be obtained by delaying and swapping flights. Flight f3&f4 are reassigned to a2 and flight 7&8 are reassigned to a1. New schedule of a1&a2 are given in table 3&4.

### 3.1 Mathematical Model

One difficult aspect of aircraft recovery problem is to determine the objectives. In general the objectives of operational personnel are to minimize the total delay, the number of canceled flights and the deviation from the original plan.

Here we adopt the objective to minimize the total delay. A user-defined parameter M is introduced to reflect the tradeoff between flight delay and cancellation, which means that a flight cancellation is equivalent to an M minute flight delay. Operation personnel often give different priority to different kind of flights, like VIP flights, long haul flights, medium haul fights, short haul flights, so a weight w associated with the category of the flight is defined.

Changes to aircraft routings may affect crew schedules and passengers, so recovery plan with minimal deviation is desired. To minimize the deviation from the original plan, an incentive (negative since our goal is to minimize the total delay) is add if the continuous combinations of the original flight path (started from the source node) is included in a path. The weight of an arc is the delay minutes of that flight multiply its weights and minus the incentive. For a canceled flight, the cost is its weight multiplies M.

### Definitions

Here we provide definitions that we use through the paper.

**Sets:**
- A = set of aircraft
- P=set of airports
- F = set of flights
- S = the set of required final states
- T = set of feasible aircraft path
- N = dynamic set containing flights f that have been canceled

**Index:**
- a = index of aircraft
- p=index of airports
- f = index of flights
- s = index of required final states
\[ t = \text{index of feasible aircraft path} \]

**Parameters:**

- \( b_{tp} = 1 \) when path \( t \) is terminated at airport \( p \), else 0
- \( b_{ta} = 1 \) when path \( t \) is serviced by aircraft \( a \), else 0
- \( b_{tf} = 1 \) when path \( t \) covers flight \( f \)
- \( h_p = \text{the number of aircraft at airport } p \text{ at the end of the recovery period to carry out the schedule after recovery period} \)
- \( d_{ta} = \text{the cost of aircraft path } t \text{ if it is serviced by aircraft } a \)
- \( w_f = \text{the weight of flight } f \)
- \( M = \text{user-defined cancel-delay equivalent, which represents that a cancellation of flight is equivalent to a } M \) minutes flight delay
- \( f_{Detch}^t = \text{the departure airport of the } i\text{-th flight in aircraft path } t \)
- \( f_{Arr}^t_i = \text{the departure airport of the } i\text{-th flight in aircraft path } t \)
- \( f_{Sched}^t_i = \text{the scheduled departure time the } i\text{-th flight in aircraft path } t \)
- \( f_{Arr}^t_i = \text{the scheduled arrival time the } i\text{-th flight in aircraft path } t \)
- \( i = \text{the sequence number of flight in aircraft path} \)
- \( TO_a = \text{the turn around time of aircraft } a \)
- \( p_{cf,s} = \text{the start of the curfew of airport } p \)
- \( p_{cf,e} = \text{the end of the curfew of airport } p \)
- \( p_{cl,s} = \text{the start of the closure of airport } p \)
- \( p_{cl,e} = \text{the end of the closure of airport } p \)

**Variables:**

- \( x_t^a = 1 \) if path \( t \) is assigned to aircraft \( a \), else 0
- \( y_f = 1 \) if flight \( f \) is canceled, else 0

We model the ARP with as a set partitioning problem.

\[
\text{Minimize } \sum_{a \in A, t \in T} d_{at} \cdot x_t^a + M \cdot w_f \cdot \sum_{f \in F} y_f
\]  \hspace{1cm} (1)

s.t.
\[
\sum_{a \in A, t \in T} b_{ta} \cdot x_t^a + y_f = 1 \quad \forall f \in F \hspace{1cm} (2)
\]
\[
\sum_{a \in A, t \in T} b_{tp} \cdot x_t^a = h_p \quad \forall s \in S \hspace{1cm} (3)
\]
\[
\sum_{t \in T} x_t^a \leq 1 \quad \forall a \in A \hspace{1cm} (4)
\]

**3.2 Network Representations**

Network representations are often used to describe aircraft recovery problem. The three most commonly used network representations are time-line network, time-band network and connection network [11].

We adopt time-line network for our model, one for each aircraft. Figure 1 is the recovery network for a2 in our example.

![Figure 1 Time-space recovery network of a2](image-url)

There are three kinds of nodes in the recovery network. Source node represent the initial state of aircraft. Sink nodes represent the final state of aircraft, they helps to ensure that the schedule after the recovery period can be performed as planned. The sink nodes in the recovery
network of an aircraft is the set of final states coverable by it. To reduce the number of sink nodes, sink nodes of the same airport can be clustered into a single sink node. The intermediate nodes are time-location of airplane, where it is ready to take off at an airport after having performed some flights. A minimal turnaround time is added to the arrival time of the last flight to avoid vertical arcs in the network.

An aircraft path must start with a source node and end up with a sink node.

The following constraints for aircraft path should be satisfied.

\( F_G \) is the set of flights in the recovery network \( G \), a subset of \( F \) which can be flown by aircraft \( p \). \( T_G \) is the set of path in \( G \), a subset of \( T \)

\[
f_{i,i}^{EsdD} \leq f_{i,j}^{SchD} \quad \forall t \in T_G
\]  
(7)

Constraint (7) ensures that a flight do not departure before its scheduled departure time.

\[
f_{i,i}^{ArrA} = f_{i,j}^{DepA} \quad \forall t \in T_G
\]  
(8)

\[
f_{i,i}^{EstA} + TO_a \leq f_{i,j}^{EstD} \quad \forall t \in T_G
\]  
(9)

Constraint (8) & (9) ensure the connectivity of flights. Along an aircraft path \( t \), the departure airport of a flight must be the arrival airport of its previous flight and the departure time must later than the arrival time of its previous flight plus a minimal turn around time.

\[
p_{f}^{d} \leq f_{i,i}^{EstD} \leq p_{f}^{d-e} \quad \text{if } p \in P \text{ and } p \text{ is the departure airport of the } i\text{-th flight in aircraft path } t
\]  
(10)

\[
p_{f}^{d} \leq f_{i,i}^{EstA} \leq p_{f}^{d-e} \quad \text{if } p \in P \text{ and } p \text{ is the arrival airport of the } i\text{-th flight in aircraft path } t
\]  
(11)

Constraint (10) & (11) ensures that during the curfew time of an airport, no take-off or landing takes place.

\[
p_{f}^{i} \leq f_{i,i}^{EstD} \leq p_{f}^{i-e} \quad \text{if } p \in P \text{ and } p \text{ is the departure airport of the } i\text{-th flight in aircraft path } t
\]  
(12)

\[
p_{f}^{i} \leq f_{i,i}^{EstA} \leq p_{f}^{i-e} \quad \text{if } p \in P \text{ and } p \text{ is the arrival airport of the } i\text{-th flight in aircraft path } t
\]  
(13)

Constraint (12) & (13) ensure that when an airport is close, no take-off or landing takes place.

In order to model delays, multiple arcs could identify the same flight in the recovery network. A rule check must be performed when construct the network to ensure that there are no two different arcs corresponding to the same flight traversed in the same path.

3.3 Column Generation Algorithm

Due to the huge dimension of the ARP problem, we recourse to column generation algorithm. By relaxing constraint (5) & (6), we get the master problem (MP) of column generation. We consider \( T' \), a subset of feasible aircraft path \( T \) to get the restrict linear master problem (RLMP).

We solve the RLMP to optimal. The subproblem involves generating columns for each aircraft. A column in our problem is a feasible aircraft path. The reduce cost of the column \( f \) is as equation (14), where \( \lambda_f \) is the dual variable.

\[
RPL_f = d_{nt} - \sum_{f \in F} b_{nf} \cdot \lambda_f
\]  
(14)

Columns with negative reduced cost are added to \( T' \) and solve the new RLMP. The above process continue until when no columns with negative reduced cost can be found. We adopt a labeling algorithm to solve the shorted path problem.

<table>
<thead>
<tr>
<th>Labeling Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sort the nodes in the time-space network in topological order</td>
</tr>
<tr>
<td>Foreach node ( n ) in the time-space</td>
</tr>
<tr>
<td>{</td>
</tr>
<tr>
<td>Foreach flight ( f ) departure from node ( n )</td>
</tr>
<tr>
<td>{</td>
</tr>
<tr>
<td>If the flight ( f ) satisfied the connectivity constraint</td>
</tr>
<tr>
<td>If the flight ( f ) is not in the aircraft path ( t )</td>
</tr>
<tr>
<td>If ( n \text{.cost} + f \text{.cost} &lt; n \text{.cost} )</td>
</tr>
<tr>
<td>Add flight ( f ) to aircraft path ( t )</td>
</tr>
<tr>
<td>}</td>
</tr>
<tr>
<td>}</td>
</tr>
</tbody>
</table>

4. Conclusion

Here we propose a model of aircraft recovery problem and present an algorithm to solve aircraft recovery problem based on column generation. The algorithm we propose is part of an ongoing project of China Southern Airline.

We aim to solve the ARP and propose an algorithm which fulfill the airline constraints. Delaying flight, swapping flight and canceling flight are considered in our model. Airport curfew and airport closure are also taken into account in our model. To minimize deviation from the original schedule, an incentive to discourage deviation is introduced in our model. The optimization objective is to minimize the total recovery cost.

Since it is part of ongoing project, the model is being refined and extended. In future work, crew and passenger will be taken into account and numerical result will be given to verify the efficiency and efficacy.
References