PB2000-106422

Planning and Control of Transportation Systems: Robust Airline Planning*

Professor Cynthia Barnhart

Massachusetts Institute of Technology
Cambridge, Massachusetts
77 Massachusetts Avenue
Cambridge, MA 02139

A Final Report to:
The New England (Region One) UTC
Massachusetts Institute of Technology
77 Massachusetts Avenue, Room 1-235
Cambridge, Massachusetts

May 10, 2000

* Supported by a grant from the US Department of Transportation, University Transportation Centers Program

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Planning and Control of Transportation Systems: Robust Airline Planning		5. Report Date May 10, 2000 6. Performing Organization Code
7. Author(s) Prof. Cynthia Barnhart		8. Performing Organization Report No.
9. Performing Organization Name and Address Massachusetts Institute of Technology 77 Massachusetts Avenue Cambridge, MA 02139		10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTRS95-G-0001
12. Sponsoring Agency Name and Address New England (Region One) UTC Massachusetts Institute of Technology 77 Massachusetts Avenue, Room 1-235 Cambridge, MA 02139		13. Type of Report and Period Covered Final Report 9/1/98-9/30/99 14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the US Department of Transportation, University Transportation Centers Program

16. Abstract

Airline operations are made up of many interdependent components. Both aircraft and crews are limited and costly resources. Aircraft are subject to strict maintenance rules; crews must follow complex FAA and union restrictions. Coupled together, they form a complex network over which passengers flow. Each planning decision, such as, the departure time of a flight, the type of aircraft flown, the pairing of flights for direct flights, can have a dramatic impact on other aspects of the system.

Decisions are usually made with the expectation that they will cover an extended period where each day has the same schedule and fleeting. In reality, this is rarely the case. In addition to planned deviations (for example, additional flights in business markets on busier days of the week), there are a vast array of unplanned daily deviations. These range from a crowded flight needing a few extra unplanned minutes to load and unload passengers, to weather-induced ground holds which may significantly delay departures of a large number of flights, to an aircraft being grounded for unplanned maintenance, to a crew member not being available for his or her planned flight. Deviations such as these can ripple throughout the system, causing wide-spread disruptions.

We look at how robustness may be measured and incorporated in the airline planning phase. Our focus is on the fleet assignment problem, with some limited schedule adjustment. We focus on three different areas including the identification of metrics to measure how operations deviate from the plan and how this impacts the rest of the system; the development of tools for comparing different plans given the same scenario of operational conditions; and the existence of alternate approaches to the fleet assignment problem that encourage robustness.

17. Key Words		18. Distribution Statement	
robust planning, airline schedule planning, fleet assignment, airline operations			
19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages 33	22. Price $45,000.00

Form DOT F 1700.7 Reproduction of form and completed page is authorized

Robust Airline Planning

Amy Mainville Cohn
Professor Cynthia Barnhart
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

May 10, 2000

Abstract

Airline operations are made up of many interdependent components. Both aircraft and crews are limited and costly resources. Aircraft are subject to strict maintenance rules; crews must follow complex FAA and union restrictions. Coupled together, they form a complex network over which passengers flow. Each planning decision – the departure time of a flight, the type of aircraft flown, the pairing of flights for direct flights – can have a dramatic impact on other aspects of the system. Decisions are usually made with the expectation that they will cover an extended period where each day has the same schedule and fleeting. In reality, this is rarely the case. In addition to planned deviations (for example, additional flights in business markets on busier days of the week), there are a vast array of unplanned daily deviations. These range from a crowded flight needing a few extra unplanned minutes to load and unload passengers, to weather-induced ground holds which may significantly delay departures of a large number of flights, to an aircraft being grounded for unplanned maintenance, to a crew member not being available for his or her planned flight. Deviations such as these can ripple throughout the system, causing wide-spread disruptions. We look at how robustness may be measured and incorporated in the airline planning phase. Our focus is on the fleet assignment problem, with some limited schedule adjustment. We focus on three different areas including the identification of metrics to measure how operations deviate from the plan

and how this impacts the rest of the system; the development of tools for comparing different plans given the same scenario of operational conditions; and the existence of alternate approaches to the fleet assignment problem that encourage robustness.

1 Robust Airline Planning

Airline operations are made up of many interdependent components. Both aircraft and crews are limited and costly resources. Aircraft are subject to strict maintenance rules; crews must follow complex FAA and union restrictions. Coupled together, they form a complex network over which passengers flow. Each planning decision – the departure time of a flight, the type of aircraft flown, the pairing of flights for direct flights – can have a dramatic impact on other aspects of the system.

In order to handle such complexity, many airline decisions are made, at least in part, far in advance. Furthermore, they are often made sequentially in spite of their interdependence. Most carriers develop a tactical plan by first setting the flight schedule, deciding on markets as well as frequency and times of flights. Then they decide what aircraft type to assign to each flight, trading off between operational costs and revenue issues associated with seating capacity. Next, rotations of individual aircraft are designed so as to ensure that aircraft maintenance is carried out as needed. Finally, crews are scheduled to provide full coverage for all flights while honoring union and FAA labor restrictions.

These decisions are usually made with the expectation that they will cover an extended period where each day has the same schedule and fleeting. In reality, this is rarely the case. In addition to planned deviations (for example, additional flights in business markets on busier days of the week), there are a vast array of unplanned daily deviations. These range from a crowded flight needing a few extra unplanned minutes to load and unload passengers, to weather-induced ground holds which may significantly delay departures of a large number of flights, to an aircraft being grounded for unplanned maintenance, to a crew member not being available for his or her planned flight. Deviations such as these can ripple throughout the system, causing wide-spread disruptions.

In the current airline planning literature, these issues of disruption are rarely considered. Problem inputs are assumed to be deterministic and static over the operational horizon. The success of a new approach or technique is usually gauged by comparing the cost of different plans relative to this input, disregarding the fact that the "better" plan may be less robust and thus perform worse operationally.

We look at how robustness may be measured and incorporated in the airline planning phase. Our focus is on the fleet assignment problem, with some limited schedule adjustment. We focus on three different areas.

First, we identify metrics with which to analyze data from a domestic operation to gain insights as to how operations deviate from the plan and how this impacts the rest of the system.

Second, we develop tools for comparing different plans given the same scenario of operational conditions. In particular, we develop a delay and cancellation recovery model that provides us with a means for estimating the cost of a plan under specific operational circumstances.

Third, we consider alternate approaches to the fleet assignment problem which encourage robustness. We present a matrix framework in which multiple scenarios are used to compare and contrast different fleet assignments under varying operational conditions.

We conclude by identifying enhancements to each of these stages that will widen the scope of the network under evaluation.

2 Fleet Assignment and the Airline Planning Process

The fleet assignment problem (FAM) is a key step in the airline planning process. This planning process is often divided into four sequential stage – scheduling, fleet assignment, aircraft routing, and crew scheduling.

In the scheduling stage, the airlines determine what flights to offer. This includes determining how many flights to offer in each market, when to offer these flights, and how to pair flights for direct (but not non-stop) service. The scheduling stage is often based primarily on market analysis with limited use of optimization.

The fleet assignment problem is usually solved next. In this stage, each of the scheduled flights is assigned an aircraft type. We will discuss this problem in further detail in the next section.

In the aircraft routing stage, flights are strung together to be flown sequentially by a particular aircraft. The purpose of this stage is to ensure that each aircraft is maintained as required by carrier and FAA requirements.

In the crew scheduling stage, blocks of work are designed for individual crew members over an extended period of time. These blocks ensure that all flights are covered while FAA and union constraints are satisfied.

2.1 The Fleet Assignment Problem

In its most basic form, the fleet assignment problem can be stated as follows: *Minimize the operating and spill costs associated with assigning aircraft fleet types to flights, subject to the constraints that each flight be assigned exactly one fleet type ("cover"), that the available number of each aircraft type not be exceeded ("count"), and that aircraft connectivity be maintained in time and space ("balance").*

Since the first basic FAM papers, significant enhancements have been added to the model. FAM with time windows allows for minor changes to scheduled departure times in order to improve the quality of the fleeting. Combined fleeting and routing incorporates maintenance requirements. Origin-destination FAM and the passenger mix problem take into account the network impacts caused by passengers connecting through hub airports.

However, little work has been done to analyze how these fleetings perform under operational conditions. In particular, the static/deterministic formulation encourages a tight schedule, with maximum utilization of aircraft. Such a schedule may be easily disrupted by a minor deviation – one flight early in the day with a minor delay in departure can snowball throughout the remainder of the day, affecting other flights, other aircraft, crew members, and connecting passengers. Such downstream impacts may be difficult to recover from efficiently in a timely manner. On the other hand, a more costly plan may actual perform better in practice because of its ability to absorb a greater degree of disruption.

Airline statistics on the cost of irregular operations suggest that a more robust plan, if such a plan can be found, could dramatically enhance operational performance.

3 Operational Analysis/Case Study

We begin by proposing metrics for comparing in detail the variations between planned and actual operations at a domestic US carrier. For a fixed time period, the fleeted schedule is compared to the daily flights that were flown, the fleet type used, the departure and arrival times, and the set of cancellations for the day. In addition, the optimal fleet schedule for a given day is determined, assuming a priori knowledge of all disruptions, in order to compute lower bounds.

Due to the complexity of the network, with its interdependence of aircraft, crews, and passengers, it can be very difficult to isolate the cause and effect of individual disruptions. For example, if we are told that a flight departed ten minutes late, it is not clear whether this is an "initial disruption," perhaps caused by unusually long passenger loading time or a minor maintenance problem, or a downstream effect, perhaps due to delays in the arrival of the prior flight flown by that aircraft. Furthermore, there are no iron-clad rules governing how the carrier recovers from disruptions. Instead, rules-of-thumb, intuition, and experience are often the key components in adjusting the day's plan. This makes it difficult to make definitive statements about how well a plan performs operationally. Instead, we suggest focusing on quantifying variability without making any claims about the quality of the initial plan.

3.0.1 Aircraft Metrics

The following metrics relate to the impact of variability on aircraft utilization.

- Average delay minutes per flight
- Total delay minutes
- Number/Percent of flights delayed
- Number/Percent of flights canceled
- Total aircraft delay cost
- Total aircraft cancellation cost

3.0.2 Passenger Metrics

The following metrics relate to the impact of variability on passengers.

- Average passenger delay minutes
- Number/Percent of passengers delayed
- Number/Percent of missed connections
- Number/Percent of passengers with changed itineraries
- Number/Percent of passengers not transported
- Total passenger delay cost
- Total passenger cancellation cost
- Total aircraft delay cost due to holding for connecting passengers

3.0.3 Crew Metrics

The following metrics relate to the impact of variability on flight crews.

- Number/Percent of missed crew connections
- Total delay minutes associated with late crew connections
- Total delay cost associated with late/missed crew connections
- case study of a disrupted day – comparison of actual to planned
- discussion of costs involved in disruption

3.0.4 Inputs

The input data for the analysis would include the following:

- For each time period (perhaps one peak month, one off-peak month), the original fleeting and schedule. Specifically, for each flight we require the flight number, assigned aircraft type, scheduled departure time and location, and scheduled arrival time and location.

- For each day in the time period, for each flight, the actual aircraft type, departure time, and arrival time. If the flight were canceled, this must be indicated and the reason for cancellation given. In addition, reasons for delay (late crew, late aircraft, maintenance, etc.) must be specified. We also require tail numbers to piece together network effects.
- For each day in the time period, we require information about aircraft taken out of rotation for unplanned maintenance and some indication of the cause. In particular, we want some sense of how rescheduling would affect that maintenance need.
- In order to maintain maintenance and crew feasibility without addressing them explicitly, we might be able to establish some feasibility limits restricting the set of permissible changes.

4 The Recovery Model

As discussed in the previous section, it is difficult to quantify the robustness of a fleeted schedule. Robustness is a measure of how the plan performs under stochastic conditions for a continuous period of days. The stochasticity alone is difficult to quantify, incorporating weather conditions, airport congestion, unanticipated aircraft maintenance needs, crew personnel issues such as sick time, and so forth.

Even if the disruptions for every day in the relevant time period were known a priori, it would still be difficult to determine the optimal fleeted schedule for that time period. There are two key contributors to this difficulty. First, a variety of objectives may be used. For some carriers, it might be critical to return to the original plan as quickly as possible. Others may focus on minimizing the cost impact of disruption. Alternate approaches could be to minimize the number of delay minutes or passengers missing connections.

Second, even with a well-defined objective it is not clear how each day's recovery should be determined and measured against the objective. For example, a delayed aircraft may disrupt not only that aircraft's current flight, but all of its future flights for some time horizon. Furthermore, future flights staffed by the disrupted flight's crew may also be impacted either because the crew is not available at the appropriate time or because they have used too large a portion of their permitted work hours for the day. In addition, passengers flying on any of these impacted flights may miss connections or – if connections are held for them – cause further downstream disruptions. All of these changes must be addressed operationally.

Such operational decisions are continually being made by carriers to address disruptions. These decisions are often based largely on rules-of-thumb, experience, and intuition. There is no rigorous structure that would allow for unbiased comparison between two different fleet schedules given the same set of initial disruptions. Furthermore, it is intractable to incorporate all elements of the network into any such structure. In the same way that the planning process must be broken down into stages for tractable, so too should the recovery phase, particularly given the brief time period in which it must be solved. Therefore, we have chosen to focus on the aircraft impact of delaying and cancelling flights.

Even within the aircraft recovery model, there are many different layers of sophistication that could be used. We outline a number of these below:

1. At the most basic level, we assume that the locations of all aircraft at the beginning of the time horizon (this could be a few hours, a day, or a number of days) are known. In addition, we assume that the set of initial disruptions occuring during this time period are also known. [Given that our objective is to improve the robustness of the fleeted schedule, we focus specifically on non-catastrophic distruptions. That is, our model is not designed to address multi-airport, massive disruptions over an extended period of time (for example, a large winter storm).] We further assume that flights must be flown by their original fleet type. Finally, we assume aircraft delay costs and flight cancellation costs. Our objective is to determine which flights to cancel and at what time the remaining flights should depart so as to minimize delay and cancellation costs.

2. In the next stage, we add soft constraints specifying the end conditions. For example, it might be required that all stations have the planned number of each aircraft type on the ground at the end of a predetermined time. Or, a penalty might be assigned associated with not returning immediately to the original plan.

3. In stage three, we permit the swapping of aircraft types in order to enhance recoverability. Thus, if an aircraft is available due to other

disruptions, an otherwise delayed flight may be assigned to that aircraft even if it is not the originally fleet type.

4. In stage four, we look out over an extended horizon. We continually update the proposed recovery plan as subsequent informaiton about network disruptions is gained.

5. In stage five, after determining the recovery of the aircraft, we solve the passenger mix problem to reassign disrupted passengers.

6. In stage six, passenger mix is solved in conjunction with aircraft recovery.

Initially, we have focused on the basic model. This will allow us to develop and assess the tractability of algorithms for solving this model, gaining insight into the feasibility of more advanced models. In addition, this will provide us with a preliminary source of comparison between different schedules that will yield insight into analyzing the robustness of a fleet schedule.

Our basic model can be stated as follows: *Given a set of aircraft of a specific fleet type in specified locations at the beginning of the time period, requirements for where these aircraft must be at the end of the time period, a set of flights to be covered or cancelled, a penalty function capturing delay costs, and a set of cancellation costs, find the optimal set of cancellations and delays to minimize cost.*

5 Basic Approach

5.1 Formulation

Recall the problem description. We consider a fixed time period (perhaps, one day), and assume that all disruptions are known a priori at the beginning of the time period, and that the horizon for recovery is fixed. We are given the original schedule, some information about disruptions, and want to develop a new schedule for the day that minimizes any negative impacts.

Our inputs are the set of flights for the time period (origin, destination, original departure time, original travel time, original fleet type) and the number of aircraft of each type available in each city at the start of the window.

We also have a set of disruption data (flights with new earliest start times, flights with lengthened travel times, points in time at which an aircraft becomes unavailable, etc.).

Our goal is to build paths covering the time period's work which are optimal with regards to penalties for delayed arrivals, penalties for canceled flights, and penalties for fleet imbalances at the end of the day.

A feasible path must be sequential in time and space, and not violate airport curfews (we assume any flight can be delayed by any amount, and can use the penalty on this accordingly).

We have the following constraints - each flight is covered at most once, "lost" aircraft are accounted for, and we don't use more paths than we have aircraft.

The formulation for the basic recovery model is as follows:

$$x_f^t = \begin{cases} 1 \text{ if flight } f \text{ departs at time } t \\ 0 \text{ else} \end{cases}$$

$$z_f = \begin{cases} 1 \text{ if flight } f \text{ is cancelled} \\ 0 \text{ else} \end{cases}$$

y_j^{t+} = flow on ground arc out of time t at airport j

y_j^{t-} = flow on ground arc into time t at airport j

$A \equiv$ set of airports
$F \equiv$ set of flights
$F_{oj}^t \equiv$ set of flight copies originating at time t at airport j
$F_{dj}^t \equiv$ set of flight copies terminating at time t at airport j
$T_f \equiv$ set of times at which flight f can depart or arrive
$T \equiv \cup_{f \in F} T_f$
$p_f^t \equiv$ penalty cost of flight f at time t
$c_f \equiv$ cancellation cost of flight f
- $\equiv$ end of time horizon
$j_0 \equiv$ source of aircraft at j at time 0
$j_{\text{-}} \equiv$ demand for aircraft at j at time -

$$\min \sum_{f \in F} \sum_{t \in T_f} p_f^t x_f^t + \sum_{f \in F} c_f z_f$$
st
$$\sum_{t \in T_f} x_f^t + z_f = 1 \qquad \forall f \in F$$

$$\sum_{f \in F^t_{dj}} x^t_f + y^{t^-}_j = \sum_{f \in F^t_{oj}} x^t_f + y^{t^+}_j \qquad \forall j \in A, \forall t \in T/\{0, -\}$$

$$\sum_{f \in F^0_{oj}} x^0_f + y^{0^+}_j = j_0 \qquad \forall j \in A$$

$$\sum_{f \in F^-_{dj}} x^-_f + y^{-^-}_j = j_- \qquad \forall j \in A$$

$$x^t_f \in \{0,1\} \qquad \forall t \in T, \forall f \in F$$
$$z_f \in \{0,1\} \qquad \forall f \in F$$
$$y^{t^*}_j \geq 0 \qquad \forall j \in A, \forall t \in T$$

The first set of constraints ensures that each flight is either assigned to exactly one time or is cancelled. The second set of constraints ensures balance of aircraft in time and space. The third and fourth sets of constraints establish end conditions. The objective minimizes delay and cancellation costs.

5.1.1 Algorithmic Techniques

The algorithmic approach to this problem is based on the time window variation of FAM studied by Brian Rexing. The basic steps are as follows:

- Begin with one true arc for each earliest possible departure, plus *super arcs* covering the remainder of the day for some fixed interval (perhaps hourly). A super arc is an arc that, instead of representing a specific departure and arrival time, represents a range of times. It provides a less restrictive network of much smaller size.
- Solve the adjusted integer program using branch-and-bound.
- If only true arcs are chosen, stop – the solution is optimal. Otherwise, we need to address super arcs. A feasibility check on the set of chosen arcs is conducted first, possibly yielding an upper bound. Then, we expand chosen super arcs into finer gradations and iterate.

In the following section, we expand upon this approach.

Master problem There are a number of strategic decisions to be made when implementing the master problem. Recall that in this master, we seek to assign flights to new departure times (or cancel the flight) to abosrb the disruptions. The new departure time chosen might be a super arc, representing a range of times to be narrowed down in a future iteration. The following decisions will significantly impact performance and tractability:

- **How do we place original arcs?** At one extreme, we could have one super arc for each minute. This would simply be FAM with time windows, and would be intractable in most instances. At the other extreme, we could have one super arc for each flight, spanning the entire feasible time window. This would provide no information in the initial iteration, since there is exactly one feasible solution (assuming no benefit in cancelling a flight), and would lead to a very large number of iterations of the master problem.

- **How do we expand the network at each iteration?** Given that a super arc has been chosen, should it be expanded to individual copies for each minute in the covered time period, or broken down into finer super arcs? Can we determine whether the super arc contains an arc in the optimal solution?

- **What can we say about optimality given a master solution with super arcs and a feasible solution with only true arcs?** How can we use bounds to enhance solvability? Can we generate a lower bound on the optimal solution?

Feasibility problem In order to iterate the master problem, it may be advantageous to determine if the current master solution is feasible. That is, we wish to determine if there is a feasible solution using only flight arcs associated with the super arcs chosen by the master problem. There are a number of ways to approach this. We discuss this issue below.

Note that whenever we find a feasible solution that decreases the upper bound, we may be able to decrease the length of some time windows, given a non-linear penalty function. That is, the penalty of a significant delay may outweigh the gap to the upper bound.

- Option one – expand the chosen super arcs to true arcs and solve the integer program. If the integer program finds an optimal solution, this provides us with an upper bound. If no integer solution is found, then the expansion of the super arcs should be altered to reflect the fact that new arcs should be incorporated. Otherwise, several iterations might occur before new arcs enter the master.

- Option two – solve the feasiblity model using dynamic programming. By avoiding total enumeration of the chosen super arcs, the feasiblity problem may become much more tractable
- Option three – smart enumeration of strings. The nature of airline networks, with high utilization of resources, restricts the number of feasible sequences of work for each aircraft, allowing for a tractable enumerative approach. In the next section, we discuss this option further.

Smart Enumeration of Strings We are given the solution to an augmented fleeting problem. This is a set of super arcs all assigned to a specified fleet. Associated with each super arc are earliest and latest departure times and the true arc length (including turn time). We are also given the start and end locations for each aircraft in that fleet. The problem is to determine whether there is a feasible set of actual start times for each flight to ensure cover and connectivity.

Clearly, one approach is to build minute by minute copies of each super arc and solve the associated IP. However, with large time windows this may become very difficult to solve. We seek to exploit the structure of the problem to find an alternative approach.

Elements of the Problem

There are some key elements to the problem which can be exploited in solving the feasibility question. First, all flights must be covered. Second, it is unusual that aircraft will sit idle for prolonged periods, especially under delay conditions. Third, there are temporal and spacial conditions limiting the ordering of flights (for example, flight B cannot follow immediately after flight A unless the origin of B is the destination of A and the latest departure of B is after the earliest arrival of A). Fourth, we are only concerned with feasible solutions rather than optimal solutions.

All of these factors suggest that an intelligent enumerative approach may be an improvement over the IP approach.

Single Aircraft

We begin by considering the single aircraft case. We will denote the number of flights in the problem by *n*. Clearly, the maximum number of possible paths is bounded from above by *n!*. Assuming that the minimum length of a flight is 90 minutes (including turn time), in a 24 hour period *n* is at most 16, so the number of possible paths has a very weak upper bound of 2 x 10^13.

However, due to temporal and spacial issues this is a vast exaggeration. At any decision point, the number of options remaining is not the number of remaining flights to be covered, but the number of remaining flights out of the current city whose latest departure time exceeds the current time. This significantly decreases the number of possible paths considered by an enumerative algorithm.

Let us consider the case which allows for the maximal number of potential paths (by potential path, we refer to a path that the algorithm might follow, at least in part, until infeasibility is determined - these could be viewed as unique branches in a tree). If we only have two airports, with the minimum length flight between them, and infinite time windows, then in a day we could have 8 flights from A to B and 8 flights from B to A. Assuming our aircraft is to begin and end at airport A, we have 8 options to choose from for our first flight of the day. We then have 8 return flights out of B. Continuing, we next have 7 options followed by 7 options, then 6 followed by 6, and so forth. Thus, we have a total of (8!) * (8!), which is approximately 80,000 potential paths. Note, however, that all of these potential paths are in fact feasible paths and thus a feasible solution to this problem would be found immediately, assuming an intelligent search mechanism.

Next consider the case where we have 9 flights from A to B and from B to A. In this case, we have (9!) * (9!) potential paths, approximately 6,400,000. However, none of these are feasible. Clearly, enumerating all of them to prove infeasibility would be highly inefficient. On the other hand, all flights from A to B have identical properties, as do all flights from B to A. Therefore, there is no difference between having flight A1 follow or preceed flight A2 in time. Thus we can assign an ordering to identical flights without loss of information. If we number flights A1, A2, , A9 and flights B1, B2, , B9, and assume that Ai has an earlier departure than Aj if and only if i < j then the number of potential paths reduces to 1, which is immediately found to be infeasible.

Let us now consider the more realistic case. We begin at airport A. The number of possible first flights is first and foremost limited to those flights originating from airport A. Further their earliest start time must not exceeded the first latest departure time out of that airport. For example, if we have a flight out of A whose time window runs from 9 until 9:45, then no flight can be assigned prior to it unless that flight has an earliest departure time preceeding 9:45. We can further restrict this set by noting that if a flight other than A is chosen as the first flight, the same aircraft must return to A before 9:45. Thus, it not only needs to depart by 9:45, but to depart by 9:45 - (2 * flight length). This will dramatically limit the number of choices available at each decision point, especially with smaller time windows.

Multiple Aircraft

What happens when we expand the problem to more than one aircraft? Clearly, the number of options increase. What methods do we have for restricting these options? We can no longer restrict the choices for the first flight out of airport A as tightly as in the single aircraft case. This is because we no longer need to rely on the same aircraft returning to cover other flights available around the same time - an aircraft originating in another city may be used, perhaps only minutes after the first flight's departure.

We can, however, still take advantage of the fact that, given two flights with the same origin and departure cities, one flight can always be assumed to be first as long as its time window is not fully imbedded within the other flight's window.

We can also take advantage of the fact that, given two aircraft on the ground in the same city at the same time, the aircraft are considered identical and thus one can be chosen arbitrarily to make the next departure.

These facts will help to decrease the number of options at each decision point. However, greater insight into the size and structure of multiple aircraft problems is needed to ascertain if an intelligent enumerative approach will out-perform the traditional IP formulation.

Longer Time Horizons

Similarly, expanding the time horizon beyond one day may significantly impact computational performance.

5.1.2 An Alternate Approach to Solving the Master Problem

In this section, we present an alternate approach to solving the master problem.

Recall that our master problem contains a combination of true and super arcs. In the standard formulation, we have a set of cover constraints that ensure that either exactly one arc representing a flight is chosen (which determines the delay time and associated penalty), or that the flight is cancelled. In addition, we have connectivity constraints that ensure that the value of arcs entering a node equals the value arcs leaving that node. Unlike basic FAM, we do not need count constraints. Instead, inputs at the beginning and end of the day dictate the number of aircraft used. Note also that for a given master problem, only one fleet is considered.

One potential problem in solving our master problem is that it is an IP with a potentially large number of binary variables. This master problem may need to be solved several times in iteration to reach an optimal solution and therefore performance time may be critical. Thus, we consider an alternative approach to solving the master problem.

Consider the problem in which we relax the cover constraints. We furthermore adjust the cost on each flight arc to be the penalty associated with that arc minus the cost of cancelling the associated flight. This new problem is simply a minimum cost flow problem, with sources and sinks associated with the boundary conditions of the aircraft for the day.

If we solve the LP relaxation of this problem, we will get an integer solution. If this solution also covers no flight more than once, then it is optimal for the master problem. [Note that the objective value is assumed to include a constant representing the cost of cancelling all flights; when we cover a flight, we add in the cost of its delay and subtract out the cost of its cancellation.] If it covers some flight more than once, than the violated constraint must be incorporated in some way.

By adding a constraint to represent the violated cover constraint, we may destroy the integral structure of the LP polyhedron. Instead, we create sub-problems to address the various possibilities. For example, of there are 5 flight arcs covering a given flight and two are set to one in the current solution (say, x1 and x2), we branch to form three sub-problems. In the first, we set x1 to be one and the remaining 4 to be zero. In the second, we set x2 to be one and the remaining 4 to be zero. In the third, we set x1 and x2 to be zero and do not fix the remaining arcs.

What are the trade-offs and potential pit falls of this approach? The primary issue is that if we frequently choose many copies of a flight arc, we will branch exponentially and solve and unmanageable number of sub-problems. Thus, convergence is a concern.

On the other hand, the normal constraints on aircraft availability suggest that there may not be enough slack in the system (especially given delays) to fly a significant number of extra flights.

Even if the number of flights flown remains the same, we may run into difficulties where flight A is flown twice and flight B not flown at all. This may be manageable because each subsequent copy of a flight has increasing delay penalties. However, if flight B has a substantial initial delay, this may be a problem.

Note also that the network can be topologically ordered, which may decrease the number of violated cover constraints.

An additional question is whether we can easily solve the sub-problems using the parent basis as a starting solution.

Note also that we may be able to use branch-and-bound-like pruning techniques to reduce the number of sub-problems considered. For example, given an optimal solution at a node with some flights covered multiple times, we have a lower bound. Given a feasible solution, i.e. an optimal solution at a node in which all flights are covered at most once, we have an upper bound.

Key questions include: Is the IP approach fast enough? How big is the network? How many constraints will be violated?

Other Notes

- Using a connection arc network will not fully capture passenger connections. A connection arc only tells us that flight A is followed by flight B, allowing us to take into account the fact that A passengers can connect to B, but not telling us anything about any other flights that A passengers could connect to (and, perhaps more importantly, which flights they can't connect to).

- To incorporate passenger connection costs, we could add an extra set of variables. Let y(ij) represent passengers on flight arc i being able to connect to flight arc j. Such a variable exists for any appropriate flight pair. The constraint $xi + xj >= 2yij$ will capture this. The value of arc yij will be the negative of the penalty associated with the missed connection. Note that this will not address the possibility of passengers taking another flight to their destination.

- Using the maximum penalty on super arcs (rather than the minimum) does not necessarily yield an upper bound, since it may not have a feasible solution of corresponding true arcs.

5.1.3 Differences from FAMTW

In this section, we highlights some of the ways in which this problem differs from the FAMTW approach.

Recall that in the Rexing model, individual arc copies are assumed at each minute within the feasible time window. All copies are assumed to be the same in cost and flight duration. Then, flight copies are clustered together into super arcs and expanded as needed.

In our approach, the costs of flight arc copies vary because of the delay penalty structure. Furthermore, the time period for valid departures is much larger (perhaps for the remainder of the day), making minute-by-minute copies intractable. Instead, we create frequent flight copies during "likely" time periods. For other periods, we have "surrogate" arcs, similar to the super arcs in the Rexing model. These depart at the end of the time period, arrive as though they had departed at the beginning of the time period, and assume the delay cost associated with the beginning of the time period. Should such an arc be chosen, we then need to check feasibility and optimality. We use an approach similar to Rexing's model to determine feasibility. Optimality is a new issue, since we no longer assume that all flight copies have the same cost, independent of time. Thus, we need an additional stage in which the actual flight time chosen is given a new flight arc with its true delay cost. In addition, we might want to add further flight arcs around that time. We need to iterate within the model until all flights are assigned to an arc with true associated delay costs.

Questions associated with this approach are

- What kinds of arcs should be created?
- How is the feasibility of a solution checked?
- How is the optimality of a solution checked?

Arcs

In Rexing's work, there is a valid time window around some (possibly exhaustive) subset of flights. Ideally, he would like to have copies of a flight arc at one minute intervals throughout its time window. However, this may make the network intractible. Therefore, he replaces each flight's set of time window copies with a single "super arc." This super arc starts at the end of the departure window and ends at the beginning of the arrival window.

Thus, it captures all possible connections - and then some. When this model has been solved, he runs a feasibility check to see if there are valid actual flight arcs of true flight length that make the model valid. He does so by decomposing the solution by fleet.

For each fleet, he builds a true time window network. In addition, he adds backwards arcs in order to ensure a feasible network flow. The cost on flight arcs is zero; the backwards arcs carry a penalty. If there is a solution to this problem with cost zero, then the original solution was feasible. If not, problem flights (i.e. chosen flights with no possible connection given true flight lengths) are identified. For such problem flights, the associated super arcs are replaced by a set of time window copies and the process is iterated.

Our problem is different in that, instead of having relatively small time windows, we can allow for flights to be delayed as much as needed throughout the day. Thus, the time windows are extremely long. A variation of Rexing's approach for us would be to divide the day into time windows and use one super arc for each flight for each window.

Suppose, for example, that there is a flight scheduled to depart at 1:35 pm which has a delayed earliest departure of 2:15 pm. We might create actual flight copies at 1 minute intervals from 2:15 to 3:15. Then, from 3:15 to 4:15, we might have one "super arc" every 5 minutes. From 4:15 to 6:15, we might have a super arc every 15 minutes. From 6:15 to 8:15 we might have a super arc every half hour. Finally, one remaining super arc would cover the interval from 8:15 until the end of the day's time horizon.

Thus, unlike the Rexing model in which the master problem assigns fleets to super arcs, primarily with the goal of assigning fleets to flights, our master problem is making the decision to cancel a flight or to assign it to a time window..

An additional issue in our model not found in the Rexing model is the question of what cost to put on each super arc, since all flight copies are no longer the same because they incorporate delay costs. In the same way that we are overly optimistic in the flight length of the super arcs, we should also use the delay cost associated with the earliest arc in the interval.

Feasibility

In the Rexing model, the master problem considers all flights and all fleets, and the sub-problems are decomposed by fleet. Given that the super arcs are shorter in flight time than true arcs, it is therefore possible to have a sub-problem in which the given set of flights cannot all be covered by the associated fleet. Furthermore, the actual schedule of the flights is not known from the super arcs. Therefore, a sub-problem is solved to determine the feasibility of the model and, if feasible, the actual schedule.

In our problem, the master problem is already decomposed by fleet, so replacing all super arcs with true flight copies simply results in the original approach 1 formulation. Another approach would be to take only those arcs included in the solution to the master problem and place them in the sub-problem, similarly expanding super arcs into time window copies. Backwards arcs would again be needed for feasibility checking, and super arcs identified as problem arcs would be replaced in the master problem with true flight copies.

However, the Rexing model has the benefit of each flight copy in a time window having the same cost. Thus, when feasibility is reached, optimality is ensured. In our model, the cost of each arc varies according to the delay cost function. Thus, in addition to checking feasibility, we also have to establish optimality. We may be able to do both in the following step.

Optimality

Since each super arc is optimistic in both its length and delay cost, we can only be sure that a solution is optimal if it uses true flight copies exclusively. Thus, we could replace Rexing's altering sequence of master and sub-problems with a sequence of ever-expanding master problems.

We begin with the initial set of arcs, some of which are true flight copies, but most of which are super arcs. We solve the optimization problem. If only true flight copies are chosen, then the solution is both feasible optimal. If not, some set of super arcs have been chosen. If such an arc represents a small time window, we replace it with an exhuastive set of true flight copies. If it represents a large time window, we divide that time window into smaller increments and assign to each a new super arc. We then re-solve the augmented master problem and repeat. We no longer need the intermediate sub-problem step with backwards arcs since we do not decompose the problem.

5.2 RFAM with fleet swapping

From a formulation stand point, we can simply add an additional index on all flow variables representing the associated fleet type. However, significant algorithmic changes may be warranted due to the explosion in problem size.

5.2.1 Formulation

$$x^t_{fk} = \begin{cases} 1 \text{ if flight } f \text{ departs at time } t \text{ on fleet } k \\ 0 \text{ else} \end{cases}$$

$$z_f = \begin{cases} 1 \text{ if flight } f \text{ is cancelled} \\ 0 \text{ else} \end{cases}$$

y^{t+}_{jk} = flow of fleet k on ground arc out of time t at airport j

y^{t-}_{jk} = flow of fleet k on ground arc into time t at airport j

$A \equiv$ set of airports
$F \equiv$ set of flights
$K \equiv$ set of fleets
$F^t_{ojk} \equiv$ set of flight copies originating at time t at airport j for fleet k
$F^t_{djk} \equiv$ set of flight copies terminating at time t at airport j for fleet k
$T^k_f \equiv$ set of times at which flight f can depart or arrive for fleet k
$T \equiv \cup_{k \in K} \left(\cup_{f \in F} T^k_f \right)$
$p^t_f \equiv$ penalty cost of flight f at time t
$c_f \equiv$ cancellation cost of flight f
$-$ $\equiv$ end of time horizon
$j_{0k} \equiv$ source of aircraft of type k at j at time 0
$j_{-k} \equiv$ demand for aircraft of type k at j at time $-$

$\min \sum_{k \in K} \sum_{f \in F} \sum_{t \in T_f} p^t_f x^t_{fk} + \sum_{f \in F} c_f z_f$
st
$\sum_{k \in K} \sum_{t \in T_f} x^t_{fk} + z_f = 1 \qquad \forall f \in F$

$\sum_{f \in F^t_{dj}} x^t_{fk} + y^{t-}_{jk} = \sum_{f \in F^t_{oj}} x^t_{fk} + y^{t+}_{jk} \qquad \forall j \in A, \forall t \in T/\{0,-\}, \forall k \in K$

$\sum_{f \in F^0_{ojk}} x^0_{fk} + y^{0+}_{jk} = j_{0k} \qquad \forall j \in A, \forall k \in K$

$\sum_{f \in F^{-}_{djk}} x_{fk} + y^{-}_{jk} = j_{-k} \qquad \forall j \in A, \forall k \in K$

$x^t_{fk} \in \{0,1\} \qquad \forall t \in T, \forall f \in F, \forall k \in K$
$z_f \in \{0,1\} \qquad \forall f \in F$
$y^{t*}_{jk} \geq 0 \qquad \forall j \in A, \forall t \in T, \forall k \in K$

The first set of constraints ensures that each flight is either assigned to exactly one time or is cancelled. The second set of constraints ensures balance of aircraft in time and space. The third and fourth sets of constraints establish end conditions. The objective minimizes delay and cancellation costs.

5.3 RFAM followed by passenger mix

Given a solution to the basic RFAM, we have a new schedule for the day. Assume as input to our next problem the following:

- Flights with new departure times and capacities (these may have changed if fleet swapping was permitted)
- Passengers with sets of feasible itineraries (including the null itinerary)
- Each itinerary cost may incorporate delay or cancellation penalties, penalty for changing flights, penalty for adding connections, etc. and may vary by fare class, etc.
- Itineraries may be added through column generation

5.3.1 Formulation

$$x_{pi} = \begin{cases} 1 \text{ if passenger } p \text{ is assigned to itinerary } i \\ 0 \text{ else} \end{cases}$$

$$\delta_{fi} \equiv \begin{cases} 1 \text{ if flight } f \text{ is part of itinerary } i \\ 0 \text{ else} \end{cases}$$

$c_{pi} \equiv$ the cost of assigning passenger p to itinerary i

$P \equiv$ the set of all passengers

$C_f \equiv$ the capacity of flight f

$I_p \equiv$ the set of all itineraries for passenger p

$I \equiv \cup_{p \in P} I_p$

$F \equiv$ the set of all flights

$\min \sum_{p \in P} \sum_{i \in I_p} c_{pi} x_{pi}$
st
$\sum_{i \in I_p} x_{pi} = 1 \qquad \forall p \in P$

$\sum_{p \in P} \sum_{i \in I_p} \delta_{fi} x_{pi} \square C_f \qquad \forall f \in F$

$x_{pi} \in \{0, 1\} \qquad \forall p \in P, \forall i \in I_p$

The first set of constraints ensure that each passenger is assigned to exactly one itinerary (including the null itinerary). The second set of constraints ensure that no flight exceeds its capacity.

5.4 RFAM concurrent with passenger mix

In order to simultaneously adjust flight times (without swapping fleet types) and reassign passengers, we combine the above models.

5.4.1 Formulation

$$x_f^t = \begin{cases} 1 \text{ if flight } f \text{ departs at time } t \\ 0 \text{ else} \end{cases}$$

$$z_f = \begin{cases} 1 \text{ if flight } f \text{ is cancelled} \\ 0 \text{ else} \end{cases}$$

y_j^{t+} = flow on ground arc out of time t at airport j

y_j^{t-} = flow on ground arc into time t at airport j

$$q_p^i = \begin{cases} 1 \text{ if passenger } p \text{ is assigned to itinerary } i \\ 0 \text{ else} \end{cases}$$

$A \equiv$ set of airports
$F \equiv$ set of flights
$P \equiv$ set of passengers
$I_p \equiv$ set of itineraries for passenger p
$I \equiv \cup_{p \in P} I_p$
$F_{oj}^t \equiv$ set of flight copies originating at time t at airport j
$F_{dj}^t \equiv$ set of flight copies terminating at time t at airport j

$F_i \equiv$ set of flight copies included in itinerary i
$T_f \equiv$ set of times at which flight f can depart or arrive
$T \equiv \cup_{f \in F} T_f$
$d_f^t \equiv$ delay cost of flight f at time t
$c_f \equiv$ cancellation cost of flight f
$b_p^i \equiv$ cost of moving passenger p to itinerary i

$C_f \equiv$ the capacity of flight f

$- \equiv$ end of time horizon
$j_0 \equiv$ source of aircraft at j at time 0
$j_- \equiv$ demand for aircraft at j at time -

$\min \sum_{f \in F} \sum_{t \in T_f} d_f^t x_f^t + \sum_{f \in F} c_f z_f + \sum_{p \in P} \sum_{i \in I_p} b_p^i q_p^i$
st
$\sum_{t \in T_f} x_f^t + z_f = 1 \quad \forall f \in F$

$\sum_{f \in F_{dj}^t} x_f^t + y_j^{t^-} = \sum_{f \in F_{oj}^t} x_f^t + y_j^{t^+} \qquad \forall j \in A, \forall t \in T/\{0, -\}$

$\sum_{f \in F_{oj}^0} x_f^0 + y_j^{0^+} = j_0 \qquad \forall j \in A$

$\sum_{f \in F_{dj}^-} x_f^- + y_j^{-^-} = j_- \qquad \forall j \in A$

$\sum_{i \in I_p} q_p^i = 1 \qquad \forall p \in P$

$\sum_{p \in P} \sum_{i \in I_p} \delta_{fi} x_{pi} \square C_f \qquad \forall f \in F$

$q_p^i \square x_f^t \qquad \forall p \in P, \forall i \in I_p, \forall f, t \in F_i$

$x_f^t \in \{0, 1\} \qquad \forall f \in F, \forall t \in T_f$
$z_f \in \{0, 1\} \qquad \forall f \in F$
$y_j^{t^*} \geq 0 \qquad \forall j \in A, \forall t \in T$
$q_p^i \in \{0, 1\} \qquad \forall p \in P, \forall i \in I_p$

The first set of constraints ensure that each flight is either cancelled or assigned to exactly one flight time. The second, third, and fourth sets ensure network balance of aircraft. The fifth set ensure that each passenger is assigned to exactly one itinerary. The six set ensure that aircraft capacity is not exceeded. The final set ensure that a passenger is not assigned to an itinerary that is not made up entirely of valid flight times.

6 Solution Analysis–The Scenario Matrix

A number of different approaches may be taken to developing a more robust fleet assignment. For example, consider the following questions:

- **How does the choice of flight times impact the cost and reliability of the FAM model?** We might consider different inputs to the FAM model with time windows (that is, flight lengths based on different percentiles). For each model, solve to find the optimal solution and the associated cost. Then determine the robustness of the new FAM and compare the trade-off of cost versus reliability.
- **Can less reliable turns be discouraged without unacceptably impacting cost?** We might begin by solving the basic FAM model with time windows, using the original flight data, to get a lower bound on cost and an upper bound on robustness. Then, we develop a new model that includes both time windows and arcs of multiple length for each flight. Using a connection arcs network, discourage those turns with significant probability of failure. Optimize some measure of robustness subject to the constraint that cost not exceed the initial lower bound by more than an input percent.
- **Can passenger reliability be encouraged without unacceptably impacting cost?** We might build on the preceding approach by adding in passenger flows and having rewards or penalties on connection arcs to acknowledge passenger connections.

In all of these cases, we need to have a way to compare the robustness of different fleetings. In the preceding section, we discussed a recovery model that allows us to compute a lower bound on the daily operating costs under specified disruptions. In this section, we present a matrix framework for comparing different fleet assignments under a range of operating conditions, or scenarios.

6.1 Matrix Description

When airlines solve the fleet assignment problem (**FAM**), they typically do so assuming deterministic schedules, flight times, aircraft availability, and so forth. The objective of the problem is to minimize deterministic estimates of operating and spill costs.

Given a fleeting, as well as several other subsequent decisions such as routing and crew scheduling, the plan is then put into operational use, usually repeated daily for a period of several weeks. During each day, variability arises, often making the original plan infeasible. For example, an aircraft may be unable to depart from a given airport at the scheduled time if there was a delay associated with the prior flight flown by that aircraft. Thus, daily adjustments must be made to the original plan to regain feasibility and minimize disruptions.

We seek to learn more about the quality of the original FAM solution under operational use. For example, a very efficient FAM solution might be optimal in a deterministic environment, but highly unreliable (and thus sub-optimal by some criteria) when implemented in a daily operational environment. In this portion of our research, we will examine how different scenarios, or daily realizations of operational conditions, perform under different fleetings. We hope that this in turn will guide us in developing an enhanced fleet assignment model which more accurately captures system variability.

The preceding diagram enables us to examine how minor variations to the fleeting input data affect many aspects of the fleet assignment in operation and, in particular, their deviation from optimality. Each section is described in detail below.

What is a scenario?

A scenario is a specific realization of the network data. It contains the original flight schedule, new earliest permissible departure times (associated with delays intrinsic to that specific flight such as maintenance problems or ground holds), new flying times, and passenger demand data.

What is a result vector?

A result vector is a vector of metrics associated with a given scenario and a given fleeting/schedule adaptation.

Specifically, for each scenario, we will solve some variation of the fleet assignment problem, which may also permit flight delays and cancellations. This may or may not incorporate time windows, passenger delays, etc. For each scenario's "optimal" fleeting, we will then compute a vector of parameters, associated with the different metrics displayed at the bottom of the diagram. Computing this vector may include solving a version of the passenger mix problem in order to estimate passenger delays and missed connections. Metrics may include cost, number of canceled flights, number of passenger

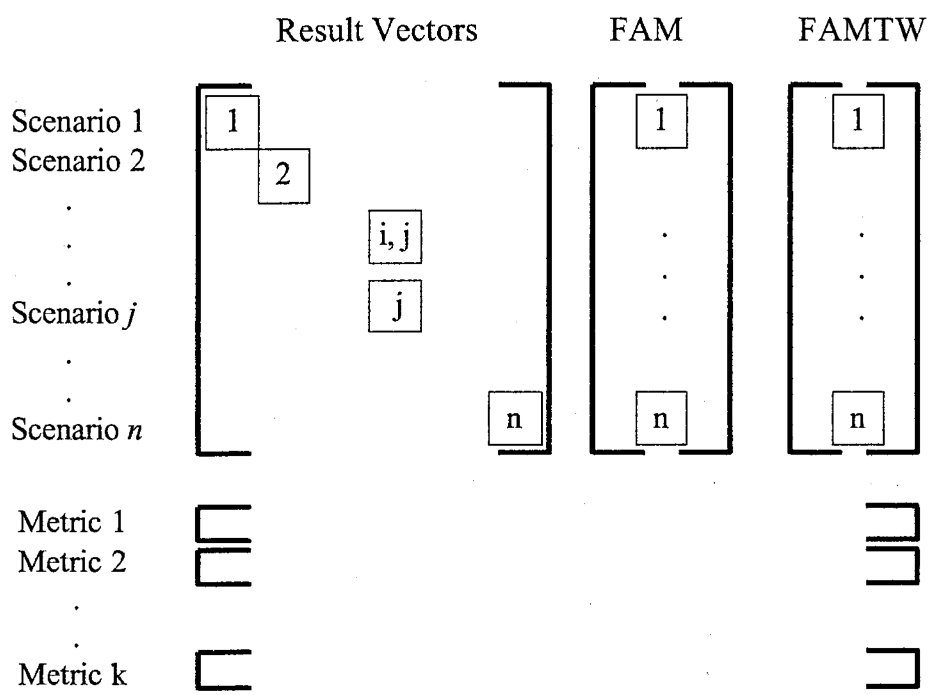
Result Vectors
FAM
FAMTW
Scenario 1
Scenario 2
Scenario j
Scenario n
1
2
i, j
j
n
1
n
1
n
Metric 1
Metric 2
Metric k

Figure 1:

delay minutes, number of missed passenger connections, etc.

Each element along the diagonal of the matrix will thus be a vector of metrics associated with a scenario and its corresponding "optimal" fleeting/schedule.

For a given column j we will then compute a similar vector for each of the other scenarios, given the current fleeting for scenario j and perhaps solving a restricted delay/cancel model in order to achieve feasibility. We will compute the same vector of metrics for each scenario, ultimately resulting in an n x n matrix of parameter vectors.

What is the FAM column?

In the FAM column, we similarly solve delay/cancellation models to compute a result vector for each scenario, using as a fleeting the optimal solution to the basic FAM model for the original deterministic inputs.

What is the FAMTW column?

In the FAM TW column, we solve delay/cancellation models to compute a result vector for each scenario, using as a fleeting the optimal solution to the FAM model with time windows for the original deterministic inputs.

What is a metric row?

The result vectors are made up of elements associated with many different metrics - for example, cost, flight delay minutes, passenger delay minutes, and so forth. For each of these metrics, we compute the column averages (as well as minimums, maximums, and standard deviations?). This allows us to see how the fleet assignment associated with the column would perform when implemented over several days with variability.

Models

What model should be solved to find the fleeting associated with each column? Clearly, the FAM and FAMTW columns will be based on the FAM and FAMTW models currently in use. For the main matrix, there are a number of options of increasing complexity. We suggest three possible models below.

Basic FAM Model

In the simplest case, we can simply begin by solving the basic FAM model with minor variations. Given variability in flight available times and flying times, the schedule input to the original FAM model may no longer be valid, or may result in an infeasible model. Thus, we use an approach similar to that used in the time windows model in order to allow flights to be shifted in order to gain feasibility. Unlike the time windows model, however, shifting a flight's departure time will result in a different cost along that flight copy.

This cost reflects some estimate of the delay-per-minute cost. In addition, we will allow flights to be canceled, making the model a set packing problem rather than a set partitioning problem. Associated with each canceled flight will also be a cancellation penalty.

FAMTW Model

We can make the model more sophisticated by allowing for time window variations to the schedule. Just as in the basic FAM version, significant deviations to the schedule would result in increased costs on the flight copy. However, changes to the flight schedule within the range defined by the time window parameters would not have any cost penalty. Note that the time windows would be constrained to the earliest available time for the flight under the given scenario - in some cases this might make shifts to an earlier time within the window infeasible. As in the basic FAM model, flight cancellations must also be permitted.

ODFAMTW Model

A more complicated version of the model would also take into account passenger flow. By using a variation of the ODFAM model with time windows, we could not only better capture spill costs, but could more accurately capture passenger delays as well. These passenger delays could then be incorporated into the objective function.

Column Models

Given the fleeting and schedule for a given column associated with the optimized scenario, we may need to make further changes to that plan to gain feasibility for the other scenarios. This may be a delay/cancellation model. For each fleet, we will solve a network model that takes into account delay and cancellation costs to find the best possible recovery given the specified plan and the scenario's realized data.

Passenger Delay Estimate Model

In order to estimate the cost of passenger delays, we need to first determine the set of passenger itineraries being flown. This can be done for the scenario being optimized by using a variation of the passenger spill model that also includes delay costs (we don't need to include cancellation or missed connection costs since at this pointa the scenario parameters are assumed to be fixed). For the other scenarios in the column, we will assume that the passenger itineraries are known. A flow model will be solved to determine the optimal routing of these passengers given disruptions to the schedule.

In a more sophisticated model, we might simultaneously solve the passenger flow and delay/cancellation models used to develop non-diagonal result vectors.

Off-Diagonal Recovery Models

We reiterate here some of the extensions to the operational recovery model.

Rolling Horizon

We might continue to assume daily re-planning, but solve over an n-day rolling horizon. We would assume an initial forecast of the n days, assuming the forecasts for earlier days are more reliable than for later ones. We would have initial conditions at the beginning of day 1, with each subsequent day beginning where the previous day left off. The last day would be constrained to end with aircraft in place to be back on schedule the next day. Then, we would simulate updated forecasts to continue solving the problem each day, where a scenario's actual cost would be based on its start conditions, its most recent (i.e. that day's) forecast, and the output from the first day of the solution to that scenario's n-day planning horizon.

One difficulty with this approach might be the tractability of a longer time horizon. This might no longer be solvable using one of the approaches suggested for the basic problem, but instead require some form of decomposition.

Another difficulty might be in acquing the distributions for the simulation. Instead of one set of simulated input per scenario, for each scenario we will now need a forecast for all remaining scenarios in the column. In addition, there must be some correlation from scenario to scenario. That is, if today I am solving over a 3-day horizon, the third day of this horizon will be the second day of tomorrow's horizon and the first day of the subsequent horizon. The simulated data for this must change across these three days, but must be somewhat dependent on the prior day's data. How difficult will it be to capture this, for example through conditional distributions?

Passenger Flows

A second extension would be to incorporate passenger costs directly into the sub-problem, allowing it to affect the decision-making process, rather than simply computing the passenger costs of the outcome. We might assume that the day's passenger itineraries are known (and, presumably, do not exceed the capacity of the original fleeted schedule). We could incorporate costs associated with mixed connections, delays to final destination, cost of transfering to another airline, and cost of not reaching destination. Again, tractability of the original model will be critical in considering this extension.

Allow Fleet Swaps

Another extenstion would be to allow for fleet swaps to improve the recovery effort. Note that this will make the network much larger, suggesting that approach 1 might be intractible even if feasible for the initial, simpler problem.

We would need to incorporate some penalty cost for swapping to cover issues not explicitly addressed, such as crew issues, the impact of decreasing flight capacity, and the time impact associated with potential gate changes, re-issuing of boarding passes, passenger confusion, etc.

6.2 Further Enhancements

It would be advantageous to take into account crew impact within robustness modelling and recovery. Below, we note a couple of the complicating factors associated with such an enhancement.

- The cost structure is non-linear and sometimes non-additive as well.
- It will be difficult to contain the problem in a one-day horizon. The recovery scope for crews is naturally longer than for the fleeting/scheduling problem.
- Perhaps we could in some way restrict the scheduling and fleeting changes to those within a window that wouldn't impact the crew. Is this too restrictive? Which costs dominate - cancellations, delays, etc., or crew costs? Does such a restriction always yield a feasible solution?
- Could we use an iterative process? Perhaps we could solve the fleeting/scheduling problem with crews, then compute the crew cost then, if "too high" or failing to meet certain criteria, perhaps find a way to constrain the fleeting/scheduling problem further and resolve. Again, what horizon do we need to solve the crew problem?
- Maybe we can approximate crew costs by capturing within fleeting/scheduling the things that cause greatest cost increases (e.g. imbalance of planes at the end of the day might translate to crews in the wrong cities; delay minutes consume allowed work time for crews, etc.).
- How does allowing for fleet swaps impact the ability to handle crews? This will complicate the model significantly. Which issues are more important to address explicitly?