

A Decision Support Framework for Airline Flight Cancellations and Delays

Author(s): AHMAD I. Z. JARRAH, GANG YU, NIRUP KRISHNAMURTHY and ANANDA RAKSHIT

Source: *Transportation Science*, Vol. 27, No. 3 (August 1993), pp. 266-280

Published by: INFORMS

Stable URL: <http://www.jstor.org/stable/25768594>

Accessed: 19-07-2017 03:12 UTC

REFERENCES

Linked references are available on JSTOR for this article:

http://www.jstor.org/stable/25768594?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://about.jstor.org/terms>



INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Transportation Science*

A Decision Support Framework for Airline Flight Cancellations and Delays

AHMAD I. Z. JARRAH and GANG YU

University of Texas at Austin, Austin, Texas 78712

NIRUP KRISHNAMURTHY and ANANDA RAKSHIT

United Airlines, Corporate Research and Development, Chicago, Illinois 60666

Aircraft shortages occasionally occur during day-to-day airline operation due to factors such as unfavorable weather conditions, mechanical problems, and delays in the schedule of incoming flights. Flight controllers need to respond to such shortages on a real-time basis by delaying or cancelling flights, swapping aircraft among scheduled flights, or requesting the usage of surplus aircraft. The choices undertaken aim at minimizing the losses incurred while retaining an operable flight schedule. In this paper, we represent two network models for aiding flight controllers in this complex decision environment. The models represent an attempt at conceptualizing this important and relatively unstudied problem, and form the basis for an evolving decision support system at United Airlines.

INTRODUCTION

The schedule of flights for an airline is built using various techniques that consider factors such as market demand, aircraft size and number, available crew resources, maintenance requirements, airport constraints, to name but a few. Due to its many dimensions, the schedule building process may span days, weeks, or months depending upon the size of the airline. A key feature of such a schedule is that resources are tightly coupled to it and even minor perturbations could have a severe impact from the standpoint of resource availability and utilization. This is unfortunate because perturbations resulting from unplanned events, like aircraft shortages, occur during actual operations. Traditionally, airlines have relied upon flight controllers with access to on-line operational data to manage the day-to-day situation as it unfolds. This entails analyzing the impact of aircraft shortages on the whole network of flights and providing appropriate real-time responses. These responses usually take the form of delaying or cancelling flights, swapping aircraft among flights, and

requesting unused (spare) aircraft. The procedure is very complex due to the large number of aircraft and flights that may need to be considered, the multitude of responses that can be undertaken, and the need to provide solutions in real-time. The various intricacies described emphasize the need for computerized decision support tools for consistently providing solutions with as little impact on the airline's operation as possible. Such tools are new to the airline industry where the current practice is to have the flight controllers resolve the problem based on their judgment and aided by a plethora of on-line information about the involved flights, aircraft, and crews.

The problem described has so far received little attention in the published literature. This contrasts with extensive research on airline crew scheduling (see [2] for a survey), and some research on airline maintenance scheduling (see [3, 7, 11]). TEODORVIĆ and GUBERINIĆ^[9] discuss the problem of minimizing overall passenger delays in the eventuality of a schedule perturbation. They attempt to find the least expensive set of aircraft routings using a branch and bound procedure. They present an

example with merely three aircraft but give no further numerical results. Given the vast number of possible routings for even moderately sized problems, it is doubtful that their approach would be practical for realistic problems. TEODOROVIĆ and STOJKOVIĆ^[10] present a greedy heuristic for solving the following goal programming problem: given some perturbations in the flight schedule, find the new set of aircraft routings that first minimizes the number of cancellations, and then minimizes the overall passenger delays. The heuristic algorithm processes the aircraft in sequence. For each aircraft an attempt is made to assign as many flights as possible; subsequently, the path with the least amount of delays that cover the same number of flights is found using a recursive delay function. Again, an example is shown but no computational results are provided.

In this paper, we start by giving an overview of the issues involved in flight cancellations and delays (Section 1), and by discussing a related but unpublished work^[4] that provides a useful conceptual starting point for our models (Section 2). Next, we cast some of the problems faced by flight controllers addressing aircraft shortages into minimum-cost network flow models. The first of these models (Section 3.1) chooses a set of flight delays that can absorb the shortages, while the second model (Section 3.2) chooses another set of flight cancellations that can achieve the same goal. These models can form the basis for building a decision-support system to assist flight controllers in finding good solutions in real-time. Benefits from such a real-time system could be large; assuming 1400 flights a day, an average of 100 passengers per affected flight, a 1% flight cancellation rate, savings can be in excess of $0.5x$ million dollars per year ($1400 \times 365 \times 100 \times 0.01 \times x > 0.5x$), where x is the dollars per passenger that a human-machine system could save on cancellation costs alone. Savings could be much larger if delay costs are also included. In addition to examples, very good results are presented for computational tests performed at United Airlines (UA) where decision support tools are under development using the framework presented in this paper (Section 4). We conclude by discussing the limitations of the suggested approaches and outlining possibilities for future research in the area.

1. FLIGHT CANCELLATIONS AND DELAYS: AN OVERVIEW

WHILE THE overwhelming majority of flights operate as scheduled, aircraft shortages do occur, resulting in flight delays or cancellations. The reason for

shortages at a certain airport at a certain point in time are various, ranging from weather conditions that make flying unacceptable, mechanical problems that call for immediate attention, or delays in the schedule of incoming flights. Such shortages are often managed through cancellation of flights. Flight controllers responsible for day-to-day operations may also resort to the option of delaying some of the flights in an attempt to avoid cancellations. To illustrate, let us consider the following scenario: on a particular day at 1:00 p.m., the flight controllers learn that a certain aircraft p_1 will need immediate maintenance which will keep it inoperable until 4:00 p.m.. Aircraft p_1 was supposed to take flight f_1 scheduled for 2:00 p.m., while two other aircraft p_2 and p_3 , arriving at 2:00 p.m. and 2:30 p.m., respectively, are scheduled to take flights f_2 and f_3 at 2:30 p.m. and 3:50 p.m., respectively. As an alternative to cancelling flights, the controllers may assign flight f_1 to aircraft p_2 for a delayed departure at, say, 2:15 p.m., flight f_2 to aircraft p_3 for a delayed departure at, say, 2:50 p.m., and flight f_3 to the fixed aircraft p_1 for a delayed departure at 4:00 p.m. Another technique which can be used by the controllers is that of requesting unscheduled surplus aircraft (spares) if such a request is deemed economically and operationally attractive. These surplus aircraft may be available either at the airports where the problem planes are or can be ferried in from nearby airports.

Flight controllers have to make real-time decisions as to the set of flights that need to be cancelled or delayed because of the aircraft shortage(s). Many complex factors have to be considered by the controllers. The chosen sets of cancellations and delays should preferably be the ones which cause the minimum, or close to minimum, loss in direct revenues (ticket refunds for passengers choosing other airlines, or hotel rates for passengers choosing to wait overnight) and indirect costs like customer goodwill. Ideally, this revenue loss should be analyzed for the immediately impacted flights as well as for all the subsequently affected flights to obtain an overall acceptable solution. The schedule should remain operable after the cancellations and delays; i.e., each scheduled flight should have an aircraft. In addition, the flight controllers have to identify how the chosen sets of cancellations and delays will affect the crews (pilots and flight attendants), and interactively obtain the crews' approval for any changes in their work schedules. Other considerations relate to the maintenance needs of the aircraft, for it would not be acceptable to have cancellations or delays with consequent changes in the flight schedule that prevent an aircraft from

arriving at one of the eligible airports for receiving scheduled maintenance. Given the above complex amalgam of constraints and options, it would not be reasonable to expect the flight controllers to provide solutions that are globally attractive. This is especially true because the controllers have to provide solutions in real-time to the problem at hand.

2. THE SUCCESSIVE SHORTEST PATH METHOD

A SUCCESSIVE shortest path method (SSPM), presented by GERSHKOFF,^[4] attempts to find a good set of flight cancellations to resolve aircraft shortages; however, it fails to consider several important features of the problem, like allowing delays and usage of spare aircraft. We choose to discuss SSPM because it serves as a conceptual starting point for the later developments in the paper. We present two models in Section 4 which are much wider in scope and address some of the shortcomings of SSPM.

Figure 1 can be used to describe SSPM. Time is placed on the vertical axis, while the horizontal axis lists various relevant stations (airports). The rectangular nodes are sequences of flights (termed "movement groups") during which there is one or more aircraft on the ground. In contrast, during the time between movement groups no aircraft are on the ground. The idea behind movement groups is that if a shortage exists in one of the groups, say *G*, then a cancellation should necessarily take place for one of the departures of group *G*, or any of the earlier groups at station PHX (Phoenix) such as group *F*; otherwise the number of departures would exceed the number of aircraft in group *G*, which is impossible. The arcs connecting the nodes represent flights between stations, and if an arc

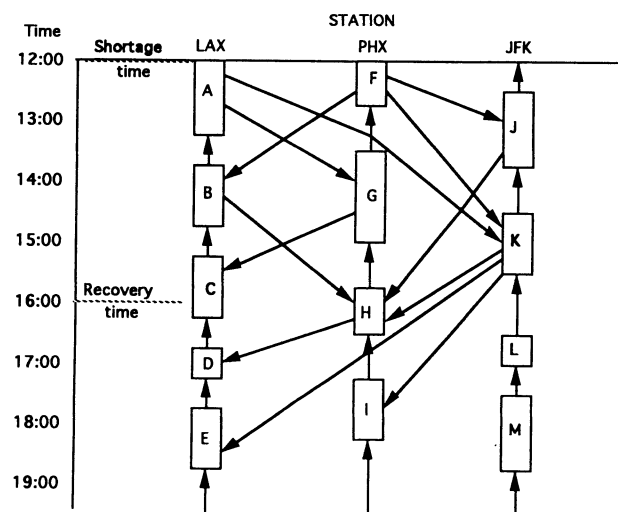


Fig. 1. Underlying graph for the SSPM.

(flight) connects two nodes *i* and *j*, then the arc (flight) will be called arc (flight) (*i*, *j*). In the scenario of Figure 1, an aircraft becomes unavailable at time 12:00 at station LAX (Los Angeles), and will become functional at time 16:00. Hence, a set of flight cancellations should originate from node *A* and terminate at station LAX at or after time 16:00, where the recovered aircraft will resume flying and "absorb" the cancellation. Movement groups at a certain station are connected by upward arcs to allow for the possibility of handling a shortage in a group at a certain station by cancelling a flight in one of the earlier groups. The cost associated with cancelling each flight is the loss in revenue to be incurred if the flight is cancelled. Hence, the shortest path is found between group *A*, and one of the several groups at LAX after time 16:00 (like *C*, *D*, and *E*). The shortest path would represent the least expensive set of cancellations. For example, if the shortest path is *AGFKHD*, then the flights (*A*, *G*), (*F*, *K*), (*K*, *H*), and (*H*, *D*) are to be cancelled. Note that the cancelled flight, (*F*, *K*), leaves an extra aircraft to be used in group *G*. Also the shortest path ends at node *D*, where no further cancellations are needed, because an extra aircraft is now available since the shortage aircraft is now fixed and available for usage.

Suppose, now, that there is more than one shortage (say two in our example, with one originating at *A* at time 12:00 and can be recovered at time 16:00, and the second originating at time 12:00 at *F* and can be recovered at time 17:00). Then, the method arbitrarily selects one of the shortages, say the one at *A*, and finds the shortest path (*AGFKHD*) as described above. Next, the arcs on the shortest path are reversed with their costs multiplied by -1 (see Figure 2). After that, the shortest path is found from node *F* to several of the nodes at PHX after time 17:00. The backward arcs with negative costs allow for the possibility of "uncancelling" a cancellation, i.e., using the flight. For example, suppose the shortest path starting at *F* is *FJHKI* (Figure 2). This means that flight (*K*, *H*) is not going to be cancelled after all. The final set of cancellations would become:

- (*A*, *G*), (*F*, *K*), and (*K*, *I*).
- (*F*, *J*), (*J*, *H*), and (*H*, *D*).

Hence, the shortage starting at LAX at time 12:00 ends at PHX after time 17:00, while the shortage starting at PHX at time 12:00 ends at LAX after time 16:00.

While this successive shortest path method is very interesting, there is obviously no guarantee of

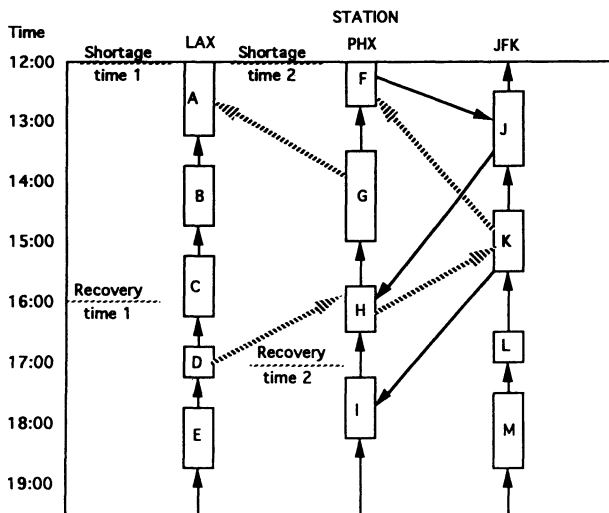


Fig. 2. Solution for the second shortage using SSPM.

optimality, since the final solution is dependent on the sequence of shortest paths chosen. For instance, in our example above, if the shortest path is found for the second shortage first, the shortest path to be reversed for the following run cannot be *FJHKI* (arc (H, K) would not exist) and a different final solution will necessarily follow. In addition, the model ignores many of the considerations described in Section 2, like surplus aircraft and allowing delays.

3. NEW NETWORK MODELS

WE PRESENT two new network flow models which will provide solutions in the form of a set of flight delays or a set of flight cancellations, while allowing for swapping aircraft among flights and for using spare aircraft. As in the SSPM, a network is used to model the flight schedule. However we do not use the idea of movement groups and actually represent individual aircraft and flights as nodes in the network. The main reason for this is to be able to differentiate between individual flight-to-aircraft assignments, since some of these assignments may not be feasible due to too short or too long ground times for some aircraft, or due to excessive delays for some flights.

The salient features of the models are:

- Multiple delays are considered.
- Multiple cancellations are considered.
- Swapping aircraft among flights is permitted.
- Usage of spare aircraft is allowed (both those available at the station where the problem aircraft is or those that can be ferried from other stations).

The models assume that a disutility can be assigned to each flight in order to reflect the value

lost if the flight is cancelled, and that the disutility of delaying each flight is assessable. Developing such disutility functions is by no means an easy or exact task and is complicated by flight connectivity considerations. However, these functions need not capture the exact disutilities of delays and cancellations for the network models presented in this paper to function properly as long as the numbers are correct in a relative sense. The factors used to generate the disutility of a flight delay or a cancellation include the number of passengers on the flight, number of passengers connecting when the flight arrives downline, possible downline delay, possible downline cancellations, lost crew time and disruption of aircraft maintenance.

To illustrate, consider the case where a flight f_i incurs a delay d_i at station s_i . The cost of the delay d_i can be related to the cost of the immediately following downline delay d_{i+1} of flight f_{i+1} at downline station s_{i+1} using, for example, the following recursive function:

$$DC(d_i, f_i, s_i) = MC(d_i, f_i, s_i) + CC(f_{i+1}, s_{i+1}) + DC(d_{i+1}, f_{i+1}, s_{i+1})$$

where,

$DC(d_i, f_i, s_i)$ = cost of a d_i minute delay of flight f_i at station s_i .

$MC(d_i, f_i, s_i)$ = loss of revenue for passengers leaving for a flight with another airline + illwill costs at station s_i due to d_i minute delay of flight f_i .

$CC(f_i, s_i)$ = cost of missed connections on flight f_i at downline station s_i + connecting passenger illwill at s_i .

One of several stopping rules can be used to terminate the recursive function, for example, end of a flying day. United Airlines has been able to quantify the elements of the disutility function with acceptable confidence through analysis of past data. We do not present, in this paper, a detailed analytic approach for developing these functions.

We choose the flows in our models to be shortages rather than aircraft. While this may sound unusual, it is a very natural choice for airline managers and flight controllers who are accustomed to thinking in terms of routing shortages through the already established flight schedules. This mode of thinking arises from the fact that it is easier to track the impact of several shortages on the flight schedule than to evaluate the possible routings of the hundreds of aircraft involved. At any rate, equivalent network formulations are attainable with aircraft rather than shortage flows.

Finally, the models are meant to be run independently for each of the fleets of an airline in order to

insure compatibility between the flights and the aircraft undertaking them.

3.1. The Delay Model

This model solves the problem of aircraft shortages at a station by delaying flights until the shortage aircraft is fixed. It allows aircraft swapping among flights as well as the use of spare aircraft available at the station or ferried from other stations. The model is a pure minimum-cost network with arcs bounded by a flow of unity.

Figure 3 depicts an example for the underlying network model for station LAX. Here we are assuming that the current time is 12:00 noon and that the user has just learned that a certain aircraft will not be available at time 13:15 as planned, but will need maintenance at the station and will be available once again at time 16:30. The vertical axis is a time axis depicting the hours of the day. The nodes on the left represent aircraft placed at the points of time when these aircraft are ready to fly. For example, node 1 represents an aircraft that can be made ready to fly at noon. The time at which an aircraft can fly is usually determined by its arrival time from its previous flight plus a certain amount of turnaround time which is needed for refueling, pre-flight checks, loading and unloading, etc. The nodes on the right represent scheduled departures of flights. For example node 2' represents a flight departure scheduled for time 13:00. An arc connecting a node on the left to another on the right represents an original flight-to-aircraft assignment at the station. For example, the flight of node 1' is originally assigned to the aircraft of node 1. In what follows, an aircraft represented by a node *n* will be referred to simply as aircraft *n*, and a flight departure represented by a node *n'* will be referred to as flight *n'*.

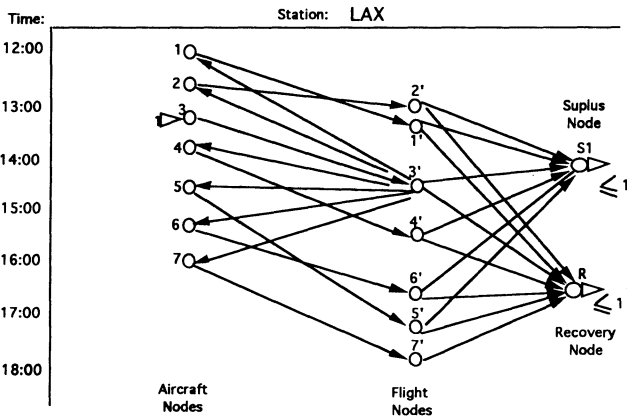


Fig. 3. Structure of the network for the delay model.

The fact that there will be a shortage at time 13:15 is represented by a supply of one at node 3 (with supply presented by $\triangleright \circ$). Each flight node is connected using backward arcs (arcs that point towards the left) to each of the aircraft nodes other than the node of the aircraft originally scheduled to take the flight and other than the node of the aircraft associated with the shortages. For instance, node 3' is connected to the aircraft nodes with the arcs: (3', 1), (3', 2), (3', 4), (3', 5), (3', 6), and (3', 7). Similar arcs emanate from the other flight nodes to the various aircraft nodes, but these are not shown in order to avoid crowding the figure with arcs. If a flow of one occurs on a backward arc, then the flight at the tail of the arc is assigned to the aircraft at its head. If a backward arc points upward, then no delay is involved because the flight departure is later than the time the aircraft is ready to fly, and, hence, the only cost on the arc is that of swapping aircraft among flights. For example, if the flow on arc (3', 4) is one, then flight 3' will be taken up by the aircraft 4 with no delay cost involved. The swapping of aircraft 4 from flight 4' to flight 3' would, however, involve some cost associated with such procedures as changing the flight gate for flight 3' from the gate where aircraft 3 is parked to the gate where aircraft 4 is parked, informing the crews and passengers of the changes, etc. On the other hand, backward arcs that point downward involve actual delays in the schedule of the departures. For example, if the flow on arc (3', 7) is one, then the departure of flight 3' is going to be delayed to 4:00 p.m. at which time aircraft 7 will be available. Arcs that involve delays, like (3', 7), have associated costs that reflect the costs of the delays.

A recovery node, *R*, is placed at time 16:30 with a demand (represented by $\circ \triangleright$) of one or less to indicate that the repaired aircraft can be used any time after time 16:30. Several flights beyond the repair time (16:30) should be considered in order to have a bigger pool of flights that can use the recovered aircraft; here, we choose to consider the three flights 6', 5', and 7'. Each of the flight nodes, 1' through 7', is connected to the recovery node to indicate that the repaired aircraft can be used for any of these flights. The cost on these arcs are, again, either swap costs if the arcs point upward, or delay costs, if the arcs point downward. Node S1 represents a surplus aircraft which can either be available at the station or be ferried from other stations. The position of the node indicates the time at which the aircraft will be available to fly from the station. Arcs connect the surplus node to the

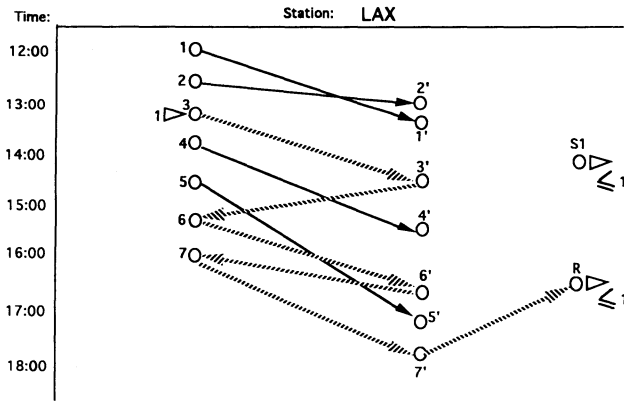


Fig. 4. One possible solution for the delay model.

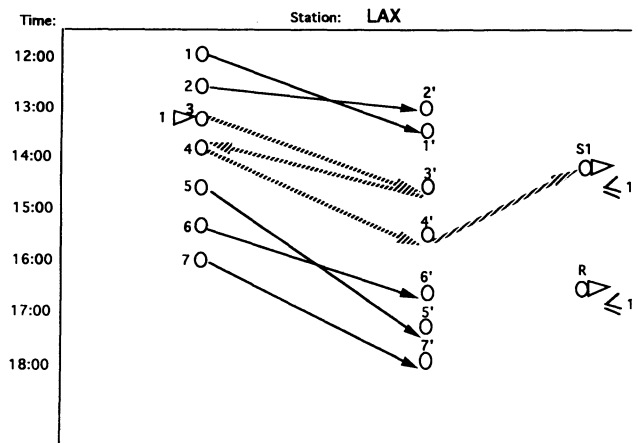


Fig. 5. A second possible solution for the delay model.

various flight nodes and the cost on these arcs include both the cost of securing the surplus aircraft and of any delay or swap costs involved. Any other available surplus aircraft can be modelled in the exact same way.

The model is to be solved as a minimum-cost network.^[5] The solution would be a sequence of arcs that starts at a supply node and terminates at a recovery node or a surplus node. For example, Figure 4 shows the solution to be: 33'66'77'R. This means that flight 3' will be delayed to be picked up by aircraft 6, flight 6' will be picked up by aircraft 7, and flight 7' will use the recovered aircraft. Hence, this solution involves only one delay. Figure 5, shows another solution, 33'44'S1, which involves no delays and makes use of the surplus aircraft. Here, flight 3' is picked up by aircraft 4, while flight 4' uses the surplus aircraft.

A mathematical formulation of the delay model will be given next. In what follows, we define the *candidate* flights for an aircraft as those most suited for reassignment to the aircraft given the flights' departure times and the aircraft availability time. While this set can conceivably include all the remaining flights (other than the originally assigned flight) of the day which do not violate the maintenance and crew restrictions, some of these flights do not represent interesting choices because of the associated excessive departure delays or aircraft layover time. Exclusion of such flights would reduce the input data collection and processing needed for the model. Similarly, we define the candidate aircraft for a flight as those most suited for undertaking the flight given the aircraft's availability times and the flights' departure times. In addition, the following terms are defined prior to the statement of the model:

A = set of aircraft originally scheduled for flights.

a = index for aircraft $a \in A$.

F = set of flights considered.

f = index for flight $f \in F$.

S = set of surplus aircraft considered.

s = index for surplus aircraft $s \in S$.

R = set of recovered aircraft considered.

r = index for recovered aircraft $r \in R$.

$\phi(a)$ = flight originally assigned to aircraft a .

$\alpha(f)$ = aircraft originally scheduled to undertake flight f . (Due to the one to one mapping of flights to aircraft, we have $\alpha = \phi^{-1}$).

F_a = subset of F consisting of candidate flights considered for aircraft a . If a is a shortage aircraft, F_a is set to empty.

A_f = subset of A consisting of candidate aircraft considered for flight f .

F_s = subset of F consisting of candidate flights considered for surplus aircraft s .

S_f = subset of S consisting of candidate surplus aircraft considered for flight f .

F_r = subset of F consisting of candidate flights considered for recovered aircraft r .

R_f = subset of R consisting of candidate recovered aircraft considered for flight f .

c_{fa} = the delay and/or swap costs involved in reassigning flight f to aircraft a .

c_{fs} = the total ferrying and delay and/or swap costs involved in reassigning flight f to surplus aircraft s .

c_{fr} = the delay and/or swap costs involved in reassigning flight f to recovered aircraft r .

$q_a = -1$ if there exists a shortage involving aircraft a ; 0, otherwise.

The decision variables involved are:

$$\begin{aligned}
 y_{\phi(a), a} \text{ (or } y_{f, \alpha(f)}) &= \begin{cases} 1 & \text{if the original assignment of} \\ & \text{flight } \phi(a) \text{ to aircraft } a \\ & \text{is discarded} \\ & \text{(or the original assignment} \\ & \text{of flight } f \text{ to aircraft} \\ & \alpha(f) \text{ is discarded)} \\ 0 & \text{otherwise} \end{cases} \\
 x_{fa} &= \begin{cases} 1 & \text{if the flight } f \text{ is reassigned} \\ & \text{to aircraft } a \\ 0 & \text{otherwise} \end{cases} \\
 x_{fs} &= \begin{cases} 1 & \text{if flight } f \text{ is assigned to} \\ & \text{surplus aircraft } s \\ 0 & \text{otherwise} \end{cases} \\
 x_{fr} &= \begin{cases} 1 & \text{if flight } f \text{ is assigned to} \\ & \text{recovered aircraft } r \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

The model can be expressed as:

$$\begin{aligned}
 \min \sum_{f \in F} & \left(\sum_{a \in A_f} c_{fa} x_{fa} + \sum_{s \in S_f} c_{fs} x_{fs} \right. \\
 & \left. + \sum_{r \in R_f} c_{fr} x_{fr} \right) \quad (1)
 \end{aligned}$$

subject to:

$$\begin{aligned}
 \sum_{f \in F_a \setminus \{\phi(a)\}} x_{fa} - y_{\phi(a), a} &= q_a \quad \forall a \in A \quad (2)
 \end{aligned}$$

$$\begin{aligned}
 \sum_{a \in A_f \setminus \{\alpha(f)\}} x_{fa} + \sum_{s \in S_f} x_{fs} + \sum_{r \in R_f} x_{fr} &= y_{f, \alpha(f)} \quad \forall f \in F \quad (3)
 \end{aligned}$$

$$\sum_{f \in F_s} x_{fs} \leq 1 \quad \forall s \in S \quad (4)$$

$$\sum_{f \in F_r} x_{fr} \leq 1 \quad \forall r \in R \quad (5)$$

$$\begin{aligned}
 y_{\phi(a), a}, y_{f, \alpha(f)}, x_{fa}, x_{fs}, x_{fr} &\in [0, 1] \quad \forall a, f, s, r \quad (6)
 \end{aligned}$$

The objective function (1) sums over all the flights the costs involved in rescheduling flights among the initially scheduled aircraft (term 1), in assigning flights to surplus aircraft (term 2), and in assigning flights to recovered aircraft (term 3). These costs include delay and/or swap costs in addition to possible ferrying costs for surplus aircraft. Equation 2 has a right-hand side of -1 when there exists a shortage involving aircraft a ; otherwise, the right-hand side of the equation is zero. In the first case, the set F_a would be empty, since no flights can be assigned to the shortage aircraft a , and the equation would simply enforce $y_{\phi(a), a}$ to be equal to one thus deleting the assignment of flight $\phi(a)$

to aircraft a . In the case where there is no shortage involving aircraft a , the equation simply forces the sum of assignments of flights to aircraft a (term 1) to equal 1 only if flight $\phi(a)$ is no longer to be assigned to aircraft a (i.e., when $y_{\phi(a), a}$ is 1). Similarly, Equation 3 enforces the sum of reassignments of flight f to the various aircraft (left-hand side) to be equal to 1 only if flight f is no longer to be assigned to aircraft $\alpha(f)$ (i.e., when right-hand side term is 1). Equation 4 ensures that at most one flight gets reassigned to each surplus aircraft s , and Equation 5 does the same for each recovered aircraft r . Finally, (6) enforces the flow to be bounded by 1. Of course, since the model is a pure network the decision variables will assume values of exactly 0 or 1 as desired.

3.2. The Cancellation Model

This minimum-cost network model solves the problem of aircraft shortages by providing an optimal solution consisting of a set of flight cancellations. All the flows on the arcs of the network are restricted to be less than or equal to one. The model can handle multiple cancellations, and makes use of aircraft swapping and surplus planes.

Figure 6 illustrates the structure of the proposed network. The nodes within each station are defined in the same way as for the Delay model. Unit supplies at nodes 3 and 9, indicate a shortage of one aircraft at each of the nodes. Node $R1$ represents the time at which the problem aircraft at LAX becomes available once again and has a demand less than or equal to one thus allowing, but not requiring, the use of the recovered aircraft. Because no delays are allowed in this model, only flights 6', 5', and 7' have connecting arcs into $R1$ since these flights can use the repaired aircraft without incurring any delays. Similarly, $R2$ represents the time at which the problem aircraft at PHX becomes operational and can receive arcs from flight 14' and any other later flights. The arcs connecting nodes across stations represent actual transfers across stations of aircraft performing flights. For example, the arc (4', 21) means that the aircraft performing flight 4' will be physically transferred to JFK where it will be ready to fly again at time 16:30. The costs on such a "transfer" arc represents the revenue which would be lost if the flight represented by the tail of the arc were to be cancelled. In order to avoid congesting the figure, only three such transfer arcs are drawn. The nodes $S1$ and $S2$, represent two surplus aircraft available in the system. The demand at each of these nodes is less than or equal to one thus allowing but not requiring the use of these aircraft.

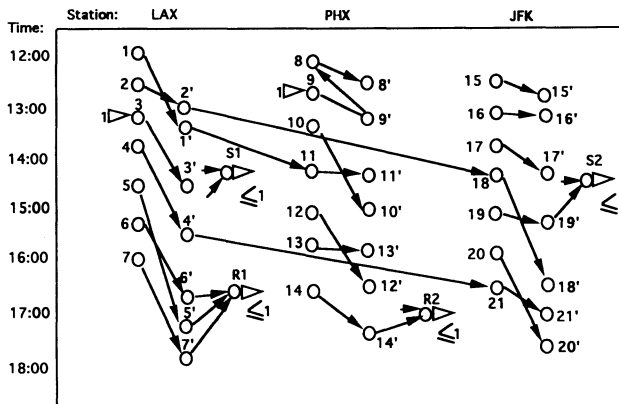


Fig. 6. Structure of the underlying network for the cancellation model.

Furthermore, an arc would connect a flight to one of these nodes only if it is possible to have the corresponding surplus aircraft available for the flight without incurring delays. For example, if an arc connects 19' to S2 then the surplus aircraft S2 can be made available for flight 19'. The cost on such arcs would be the cost incurred in making the surplus aircraft available for the desired flights. Note that the cancellation flow can now cancel either into a recovery node where a shortage aircraft is recovered, or into a surplus aircraft node.

The only backward arcs allowed at each station are those that involve no delays. For example, the only backward arc allowed out of node 9' would be arc (9', 8). Backward arcs for the other flights are not shown in the figure to avoid congestion. Each of these arcs is assigned a cost associated with swapping (see Section 3.1 above for more explanation).

It should be noted that an alternative model can be obtained if the aircraft and flight nodes are combined at all the stations, and aircraft supply nodes are defined only for aircraft that are unable to undertake their scheduled flights on time. In this network, an arc connects a flight node i to a flight node j , if the aircraft released after the completion of flight j can be used for flight i without incurring a delay and if the waiting time for the released aircraft is not deemed excessive. This would reduce the number of needed nodes in the model; however it leads to a complication. Several unit flows can now arrive at and leave from each of the flight nodes which is not acceptable since each flight can provide at most one aircraft for swapping. This situation will not occur when both aircraft and flight nodes are used (as in Figure 6), since the forward arcs connecting the two sets of nodes have an upper bound of one.

The mathematical statement for the cancellation model is similar to that of the delay model

except that, now, no flight-to-aircraft reassignments involving delays will be considered. The set A now includes the set of aircraft indices at all the stations. Each flight f has a corresponding aircraft index $\tau(f)$. Each flight f releases its aircraft $\tau(f)$ for further scheduling. Similarly, an aircraft with index a is said to be released for scheduling after performing the flight $\tau^{-1}(a)$. To illustrate, in Figure 6 flight 4' releases the aircraft with index 21. The following definitions are needed in addition to the definitions used in the delay model:

$\tau(f)$ = the aircraft released for scheduling by flight f .

$\tau^{-1}(a)$ = the flight that releases aircraft a for scheduling.

c_f = the cost of cancelling flight f .

z_f (or $z_{\tau^{-1}(a)}$)

$$= \begin{cases} 1 & \text{if flight } f \text{ (or } \tau^{-1}(a) \text{) is cancelled} \\ 0 & \text{otherwise} \end{cases}$$

The model is described as:

$$\min \sum_{f \in F} \left(\sum_{a \in A_f} c_{fa} x_{fa} + \sum_{s \in S_f} c_{fs} x_{fs} + \sum_{r \in R_f} c_{fr} x_{fr} + c_f z_f \right) \quad (7)$$

subject to:

$$\sum_{f \in F_a \setminus \{\phi(a)\}} x_{fa} - y_{\phi(a), a} + z_{\tau^{-1}(a)} = q_a \quad \forall a \in A \quad (8)$$

$$\sum_{a \in A_f \setminus \{\alpha(f)\}} x_{fa} + \sum_{s \in S_f} x_{fs} + \sum_{r \in R_f} x_{fr} + z_f = y_{f, \alpha(f)} \quad \forall f \in F \quad (9)$$

$$\sum_{f \in F_s} x_{fs} \leq 1 \quad \forall s \in S \quad (10)$$

$$\sum_{f \in F_r} x_{fr} \leq 1 \quad \forall r \in R \quad (11)$$

$$y_{a, \phi(a)}, y_{f, \alpha(f)}, z_f, z_{\tau^{-1}(a)}, x_{fa}, x_{fs}, x_{fr} \in [0, 1] \quad \forall a, f, s, r \quad (12)$$

The objective function (7) now has an additional term over objective function (1) for capturing cancellation costs. The interpretation of Equation (8) is identical to that of Equation (2) except that now a shortage can be initiated either due to an external cause (by having the right-hand side equal to -1) or by cancelling an incoming flight (i.e., if $z_{\tau^{-1}(a)}$ is 1). Similarly Equation (9) is identical to Equation (3) except that now if the right-hand side is 1 (indicating that the original assignment of flight f to $\phi(f)$ is discarded) the situation can be handled either by a swap (when x_{fa} is 1), a surplus or recovered aircraft (when x_{fs} or x_{fr} is 1, respec-

tively), or a flight cancellation (when z_f is 1).

4. COMPUTATIONAL EXPERIENCE

IN THIS section, we will present the results of the experiments conducted for the delay and cancellation models on representative sample data from UA. These models are general minimum-cost networks which involve multiple sources and sinks. However, we can combine the source nodes to one master source node using arcs with flows bounded by one. The supply at the master node would be equal to the total number of shortages impacting the system. Similarly, all the recovery and surplus nodes can be made to feed into one master sink node using arcs with flows bounded by one, and the demand at the master sink node would be set to the total supply into the system. Based on this transformation, we implemented the Busacker-Gowen's dual algorithm^[1, 8] for the minimum-cost flow network problem in which the shortest path is solved repeatedly to achieve the necessary flow. The shortest path procedure was modified to generate simple paths, since the costs on the edges could be negative and negative cycles could exist.^[6] The algorithm was coded in C and implemented on the DEC3100 workstation.

4.1. Results for Delay Model

Three stations, Chicago, San Francisco and Denver, were chosen for experimentation. These stations have a large number of departures per day and meaningful problems could be studied.

Inputs to run the model are:

- the flight data for the station in question.
- the disutility of delay for the outgoing flights at that station.

The flight data give us information about the incoming flights and the outgoing flights. A typical flight turn at, say Chicago, is represented as follows:

734 1433 GEG B 248 1514 ICT

This means that flight 734 lands in Chicago at 2:33 p.m. from Spokane (GEG) and turns to flight 248 leaving Chicago at 3:14 p.m. for Wichita (ICT). The aircraft used here is an equipment 737 type B.

Tables I, II, and III show the results of the runs for Chicago, San Francisco and Denver respectively. Column 1 represents the equipment types for which delay has occurred. Swaps can be effected within equipment type only. Columns 2 and 3 specify the time bank or interval within which the model is run. Only aircraft which either arrive during this time period or are already on the ground are considered for swaps. Column 4 indicates the number of delayed aircraft. This includes aircraft that were actually delayed and also those aircraft which are at the station but cannot leave due to mechanical problems. Column 5 indicates the number of spare aircraft available for swaps. These include actual spare aircraft if any, as well as aircraft which arrive at the station and leave the next day (overnight layovers). Column 6 speci-

TABLE I
Experimental Results for the Delay Model at CHICAGO

Equip	Time Begin	Bank End	No. Delayed /Mech.	No. of Spare	No. of Swaps	No. of Delays	Total Delay w/o Model (min)	Total Delay w/Model (min)	Disut. w/o Model (\$)	Disut. w/Model (\$)	Run Time (sec)
B, Q, N	500	900	2/0	0	4	2	70	321	3800	290	*
B, Q, N	500	900	2/1	0	6	3	146	80	4504	931	*
B, Q, N	500	1100	10/0	0	15	10	416	317	19,379	4632	*
B, J, P	1600	2359	28/2	33	35	5	503	238	57,289	12,503	30
B, J, P	1600	2359	28/2	33	35	5	503	208	57,289	11,132	28

TABLE II
Experimental Results for the Delay Model at SAN FRANCISCO

Equip	Time Begin	Bank End	No. Delayed /Mech.	No. of Spare	No. of Swaps	No. of Delays	Total Delay w/o Model (min)	Total Delay w/Model (min)	Disut. w/o Model (\$)	Disut. w/Model (\$)	Run Time (sec)
B, E	800	1100	3/0	0	4	2	127	114	10,622	8847	*
B, E, N	800	1100	5/0	0	10	3	187	167	15,308	3242	*
B, J, F	1500	2359	9/1	13	9	4	245	365	17,472	5410	*
B, J, F	1500	2359	13/2	13	17	7	423	488	22,602	6834	*
B, J, F, P	1500	2359	20/2	13	25	7	550	556	45,707	7071	16

TABLE III
Experimental Results for the Delay Model at DENVER

Equip	Time Begin	Bank End	No. Delayed /Mech.	No. of Spare	No. of Swaps	No. of Delays	Total Delay w/o Model (min)	Total Delay w/Model (min)	Disut. w/o Model (\$)	Disut. w/Model (\$)	Run Time (sec)
<i>E, K, B</i>	730	1100	2/0	0	6	1	121	90	4830	521	*
<i>E, K, B</i>	900	1200	4/1	0	9	2	112	60	12,186	1812	*
<i>E, K, D</i>	730	1200	7/3	0	13	7	244	324	17,326	4965	*
<i>B, E, D</i>	1500	2359	9/1	17	10	3	181	207	12,830	2868	*
<i>B, P, L</i>	1500	2359	13/2	17	18	5	542	455	41,837	9030	*

All runs made on DEC3100 workstation; *indicates CPU less than 10 seconds.

fies the number of swaps suggested by the model. Column 7 indicates the number of delays in the solution. Columns 8 and 9 give the total delay of all flights in minutes before and after running the model. The total delay before using the model is the cumulative delay of all flights if the aircraft took the original turns as specified in the flight data. Columns 10 and 11 give the disutility for the same two situations.

The following remarks pertain to the results of the runs:

- The problem situations tested included incidental delays which consist of a few every day delays, and mass delays as a result of inclement weather at a station. The model generated effective implementable solutions in reasonable time and is definitely amenable to real-time implementation.
- In some instances, the time of total delay in the proposed solution was greater than if the delayed flights took their scheduled turns. Upon examining the solution in more detail it was clear that in these cases a single flight was delayed for a very long time because this flight had a relatively flat disutility versus time curve. For example in Table II, run 3 had a delay of 245 minutes initially (using originally scheduled turns) and the proposed solution had a delay of 365 minutes. Table IV shows the scheduled turns for the flights for this problem situation. Table V shows the revised turns proposed by the model. The model recommends flight 1222 to Spokane (GEG) be delayed for 270 minutes which constitutes 78% of the total delay in the proposed solution. One can observe that, the flight had very little or no delay cost. From another viewpoint, we may say that this flight is a candidate for cancellation since not too many passengers will wait for 270 minutes.
- Runs 4 and 5 in Table I represent the mass delay situation due to inclement weather. A total of 30 planes have delays, 28 of them due to weather and 2 due to mechanical problems.

The number of spares is indicated as 33 all of which are in overnight layovers. In run 5 an actual spare was added to the model which resulted in a disutility reduction of 1471. The solutions obtained were very attractive for both runs as evident from the dramatic reduction in disutility.

- We also tested the model with stations which had under 20 departures per day. Overall, we concluded that the delay model proved effective in stations which had a high volume of flights. Stations with very few flights have very little operational flexibility for the model to be very effective although it finds obvious solutions.

4.2. Results for the Cancellation Model

For testing the model, three scenarios were considered by dividing the country into regions: the eastern region, the central-west region and the entire country as one region. The experiments were conducted for one UA sub-fleet, the 737Bs which are aircraft with a seating capacity of 128 and a range of about 4 hours of non-stop flying.

Tables VI, VII and VIII show the results of the runs for the cancellation model. It clearly validates the efficiency and applicability of the model for implementation in a real-time decision support system. Note that it is often the case that the number of cancellations exceed the number of aircraft shortages in the system. This is due to the fact that each of these aircraft are typically required to perform multiple flights during each day. Columns 4 through 7 display the number of cancellations, the number of swaps, the disutility, and the time taken to generate the solution.

The following remarks pertain to the results of the runs:

- The aircraft shortages included those which became available later on during the day (like a delayed arrival) as well as real shortages in which the aircraft is just not available for the rest of the day. In the former case, the model attempts to avoid cancellations by performing

TABLE IV
Initial Schedule Turns for Run 3 Table II

Incoming Flight No.	Arrival Time	Arriving from	Equip Type	Outgoing Flight No.	Departure Time	Destination
1712	1449	LAX	N	1719	1530	LAX
1759	1427	SEA	B	1514	1610	EUG
1112	1422	LAX	B	1758	1530	SEA
0535	1516	ORD	B	535	1615	BUR
1114	1519	LAX	B	1247	1601	SAN
1037	1519	PDX	N	1222	1600	GEG
0250	1523	SAN	B	250	1605	ONT
1194	1525	ONT	D	1194	1605	PDX
1294	1526	LAS	B	1063	1610	LGB
1121	1527	SEA	N	1121	1600	LAX
0675	1530	GEG	E	675	1605	MRY
1042	1530	MRY	N	1042	1610	MFR
0973	1531	MCI	D	514	1615	DEN
1770	1534	BUR	B	1770	1630	SEA
1242	1535	PHX	E	818	1630	ORD
1714	1549	LAX	N	1721	1630	LAX
0897	1615	DEN	B	567	1700	SNA
1116	1619	LAX	N	1123	1700	LAX
1769	1622	SEA	K	1772	1730	SEA
1716	1647	LAX	C	1723	1730	LAX
1118	1715	LAX	J	1125	1800	LAX
1170	1745	LAS	E	1458	1845	PHX
1499	1746	MFR	B	1499	1835	LAS
1718	1749	LAX	N	1718	1830	SEA
1127	1749	EUG	N	1127	1900	LAX
1071	1750	PDX	F	1593	2055	SNA
0346	1751	BUR	B	995	1835	BUR
1228	1751	SBA	N	1725	1830	LAX
1283	1753	SLC	J	682	1830	SLC
0571	1753	MSY	B	1275	1835	LGB
1634	1755	LGB	B	1074	1845	GEG
1178	1756	SNA	F	1178	1840	PDX
1439	1758	BOI	B	1439	1840	SAN
1068	1802	ONT	E	490	1840	DEN
1186	1802	MRY	E	1186	1840	BOI
1473	1805	PHX	E	1473	1845	ONT
0961	1805	GEG	K	1502	0720	PDX
1120	1819	LAX	B	1778	2110	SEA
1073	1826	SEA	B	1073	1905	SBA
0347	1907	DEN	F	38	2210	EWR
1122	1920	LAX	J	1131	2100	LAX
0175	1921	ORD	E	1777	2100	PHX
1033	1926	SEA	B	1033	2100	BUR
1176	1940	BUR	B	1173	2055	SBA
1722	1947	LAX	J	1421	0905	SAN
1171	2002	MFR	N	1101	0700	LAX
1435	2014	EUG	B	1476	2115	EUG
1260	2015	LGB	B	582	2115	DEN
1124	2019	LAX	N	1245	2100	ONT
0517	2022	ORD	M	318	0630	DEN
1132	2025	SBA	D	1273	0850	SBA
1234	2026	SNA	B	1234	2110	MFR
1177	2028	SEA	B	1177	2140	SAN
1026	2030	LAS	B	1026	2125	MRY
1274	2040	ONT	E	1051	0915	ONT
1016	2040	BUR	B	1701	0730	LAX
1537	2043	SAN	E	1750	0630	SEA
0439	2045	PHX	E	882	0830	DEN
0019	2052	EWR	M	24	2215	JFK
0093	2112	BOS	F	16	2210	IAD
1126	2121	LAX	J	1752	0730	SEA
1128	2217	LAX	N	1703	0830	LAX
1781	2225	SEA	B	1257	0645	BUR
0027	2320	IAD	F	92	0800	BOS

TABLE V
Turns Proposed by the Delay Model for Run 3 of Table II

Delay (min)	Action Code*	Incoming Flight No.	Arrival Time	Arriving From	Equip Type	Outgoing Flight No.	Departure Time	Destination
0		1712	1449	LAX	N	1719	1530	LAX
0		1759	1427	SEA	B	1514	1610	EUG
0		1112	1422	LAX	B	1758	1530	SEA
0		0535	1516	ORD	B	535	1615	BUR
0		1114	1519	LAX	B	1247	1601	SAN
0	S	1037	1519	PDX	N	1499	1835	LAS
0		0250	1523	SAN	B	250	1605	ONT
0		1194	1525	ONT	D	1194	1605	PDX
0		1294	1526	LAS	B	1063	1610	LGB
0		1121	1527	SEA	N	1121	1600	LAX
0		0675	1530	GEG	E	675	1605	MRY
0		1042	1530	MRY	N	1042	1610	MFR
0		0973	1531	MCI	D	514	1615	DEN
0		1770	1534	BUR	B	1770	1630	SEA
0		1242	1535	PHX	E	818	1630	ORD
0		1714	1549	LAX	N	1721	1630	LAX
0		0897	1615	DEN	B	567	1700	SNA
0		1116	1619	LAX	N	1123	1700	LAX
0		1769	1622	SEA	K	1772	1730	SEA
0		1716	1647	LAX	C	1723	1730	LAX
0	S	1118	1800	LAX	J	682	1830	SLC
15	D	1170	1830	LAS	E	1458	1900	PHX
0	S	1499	1900	MFR	B	1778	2110	SEA
0		1718	1749	LAX	N	1718	1830	SEA
0		1127	1749	EUG	N	1127	1900	LAX
0		1071	1750	PDX	F	1593	2055	SNA
0		0346	1751	BUR	B	995	1835	BUR
0		1228	1751	SBA	N	1725	1830	LAX
0	S	1283	1810	SLC	J	1186	1840	BOI
0		0571	1753	MSY	B	1275	1835	LGB
0		1634	1755	LGB	B	1074	1845	GEG
0		1178	1756	SNA	F	1178	1840	PDX
0		1439	1758	BOI	B	1439	1840	SAN
0		1068	1802	ONT	E	490	1840	DEN
65	SD	1186	1835	MRY	E	1125	1905	LAX
0	S	1473	1905	PHX	E	1131	2100	LAX
15	SD	0961	1830	GEG	K	1473	1900	ONT
270	SD	1120	2000	LAX	B	1222	2030	GEG
0		1073	1826	SEA	B	1073	1905	SBA
0		0347	1907	DEN	F	38	2210	EWR
0	S	1122	2130	LAX	J	1502	0720	PDX
0		0175	1921	ORD	E	1777	2100	PHX
0		1033	1926	SEA	B	1033	2100	BUR
0		1176	1940	BUR	B	1173	2055	SBA
0		1722	1947	LAX	J	1421	0905	SAN
0		1171	2002	MFR	N	1101	0700	LAX
0		1435	2014	EUG	B	1476	2115	EUG
0		1260	2015	LGB	B	582	2115	DEN
0		1124	2019	LAX	N	1245	2100	ONT
0		0517	2022	ORD	M	318	0630	DEN
0		1132	2025	SBA	D	1273	0850	SBA
0		1234	2026	SNA	B	1234	2110	MFR
0		1177	2028	SEA	B	1177	2140	SAN
0		1026	2030	LAS	B	1026	2125	MRY
0		1274	2040	ONT	E	1051	0915	ONT
0		1016	2040	BUR	B	1701	0730	LAX
0		1537	2043	SAN	E	1750	0630	SEA
0		0439	2045	PHX	E	882	0830	DEN
0		0019	2052	EWR	M	24	2215	JFK
0		0093	2112	BOS	F	16	2210	IAD
0		1126	2121	LAX	J	1752	0730	SEA
0		1128	2217	LAX	N	1703	0830	LAX
0		1781	2225	SEA	B	1257	0645	BUR
0		0027	2320	IAD	F	92	0800	BOS

*Indicates whether outgoing flight has been swapped (S), delayed (D) or both (SD).

TABLE VI
Results for the Cancellation Model for EASTERN Region

Equip.	Start Time	Aircraft Shortage	No. Cancelled	No. of Swaps	Total Disutility (\$1000)	Total CPU (sec)
737s	500	1	2	3	72	3.08
737s	500	2	4	5	258	3.09
737s	1200	3	2	6	54	2.50
737s	500	4	4	9	111	3.20
737s	500	5	6	11	187	3.33

TABLE VII
Results for the Cancellation Model for CENTRAL WEST Region

Equip	Start Time	Aircraft Shortage	No. Cancelled	No. of Swaps	Total Disutility (\$1000)	Total CPU (sec)
737s	500	1	2	7	50	4.90
737s	500	2	5	7	127	5.11
737s	1200	3	7	10	177	5.20
737s	500	4	9	17	266	5.33
737s	500	5	14	18	381	5.45

TABLE VIII
Results for the Cancellation Model for Entire Airline

Equip	Start Time	Aircraft Shortage	No. Cancelled	No. of Swaps	Total Disutility (\$1000)	Total CPU (sec)
737s	500	1	2	3	51	4.00
737s	500	2	4	6	61	4.22
737s	500	3	7	10	158	4.33
737s	500	4	10	10	185	4.46
737s	000	5	12	18	222	6.50

swaps whereas in the latter case it is forced to cancel flights.

- The flight turns information used contained the turns for a 24-hour period and it was sufficient to balance cancellations for the 737s because of the large size and high flight frequency of the fleet per day. Balanced cancellation sequences or loops can be found within 24 hours of flying. For fleets which have a high average flying time per flight, turns over 2 or 3 days may be required to balance cancellations.

5. EXTENSIONS, LIMITATIONS AND RESEARCH PROSPECTS

THE AUTHORS view the models and discussions presented in this paper as an initial effort in addressing the problem of flight cancellations and delays in the airline industry. The insights gained from the analysis formed the basis for an evolving decision support system at UA. However, many complex issues exist besides and beyond the immediate decisions of swapping aircraft and delaying or cancelling flights. It is important to discuss and under-

stand these issues and to identify possibilities for future research for improved approaches. Throughout the discussion, one should bear in mind that the problem at hand is a real-time one, and, hence, future related efforts should focus on good and quick solutions for realistic models.

5.1. Crew Considerations

Crew scheduling is the important problem of assigning flight crews to tours of duty (called bids) extending several weeks (see Introduction). The problem is complicated by a host of constraints that define a feasible bid, like the maximum allowed flight time, and time away from home, maximum allowed layover time between assignments, etc. Cancelling or delaying flights often necessitates changes in the scheduled crew assignments. Because of the enormous complexity involved in combining the crew scheduling problem with the cancellations/delays problem, we suggest an approach similar to what is done in practice: once a solution is identified, the resultant extensions or modifications in the tours of duty are negotiated with the affected crews to obtain their approval. If a solution that involves an extension in the duty period of a certain crew gets rejected by that crew, then the backward arc causing the unacceptable extension is deleted, and the model is rerun for an alternative solution which is acceptable to the crew. For example, if choosing arc (3', 7) (see Figure 3) results in an unacceptable extension of the duty period of the crew for flight 3', arc (3', 7) is then removed for rerunning the model. Similarly if a cancellation results in unacceptable changes in the schedule of the affected crew, the corresponding cancellation arc is deleted in the next search for a feasible solution.

5.2. Aircraft Maintenance Considerations

The scheduled assignment of aircraft to flights is done in a way which insures that each of the aircraft receives all its various types of maintenance checks which are required after certain prespecified numbers of flying hours. This, in itself, is a complex problem, which requires extensive effort (see Introduction). When aircraft are rescheduled (due to delays, cancellations, or swaps) it is important to ensure that all affected aircraft will still receive their scheduled maintenance. To check if this condition is met, the network of flights can be searched exhaustively to see if the aircraft undertaking the flight can reach an appropriate maintenance airport in time. So again, the approach is to attempt and recapture feasibility after the decision to cancel/ delay flights or swap aircraft is

made. In some severe situations, the problem cannot be resolved satisfactorily without the usage of surplus aircraft beyond those suggested by the models.

5.3. Combining Cancellations and Delays

Flight controllers typically attempt to remedy the problem of aircraft shortages first by considering the possible usage of delays. If needed delays are deemed excessive, the controllers would consider cancelling flights to absorb the shortage. The models we presented in Sections 3.1 and 3.2 are shown to be effective in aiding the controllers in these efforts. In addition, it would be desirable to investigate the possibility of finding even better solutions through combining delays and cancellations. This can be modeled using the network in Figure 6 except that we now allow both upward and downward backward arcs. For example, flight 3' would now have backward arcs emanating from it to nodes 1, 2, 4, 5, 6 and 7. Similar backward arcs emanate from the various flight nodes at a given station to all the surplus and recovery nodes at the station.

The above model, if solved as minimum-cost network, will not, necessarily, give a correct solution, because whenever a delay arc is chosen, the position of some of the network nodes will be impacted and the costs on some of the backward arcs will no longer be valid. To illustrate, suppose the optimal solution to the network of Figure 6, once solved as a minimum-cost network, involves a flow of one on the backward arc (2', 6). This means that the departure of flight 2' is to be delayed to 15:15, the time at which aircraft 6 is ready to fly. This in turn means that node 18 will have to be shifted downward to reflect the subsequent delay in the arrival of flight 2', and this will impact the delay costs on the backward arcs which connect into node 18 at JFK. This problem represents an opportunity for future research for approaches that can capture this delay/cancellation interaction. The salient feature of any such approach would be the treatment of the departure time of the flights as a variable that is updated depending on the delays chosen in the solution. The authors are currently involved in developing and testing a model along these lines.

5.4. Multicommodity Approaches

The exposition in this paper is based on the assumption that shortage problems involving aircraft of a certain fleet are to be resolved using cancellations, delays, and swaps within that same fleet. This was done to insure compatibility of the aircraft with the assigned flight segments. Another approach is to consider all the fleets simul-

taneously while allowing assignment of flights to compatible fleets only, which results in a multicommodity formulation. The obvious shortcoming of such an approach is the added complexity, while the advantage is the opportunity of obtaining better solutions because of the increased number of aircraft to choose from.

Another relevant aspect is the concept of hub and spokes which is prevalent in the modern airline industry. Here, a bank of incoming flights feed passengers from various airports (spokes) into a major airport (hub), where the passengers redistribute among the next bank of departing flights. The system has proven advantages from both operational and revenue considerations. However, the system makes the development of disutility functions for delays and cancellations a difficult task and the assessment of the impact on connecting passengers only approximate. As an alternative, one can visualize an additional commodity representing the flow of the passengers in the system, and relate the commodity to the aircraft commodities and flows. Needless to say, the resultant problem is rather complex.

In conclusion, we have presented two network models which can assist in choosing which flights to delay or cancel in the event of unexpected shortages of aircraft due to situations that may arise during the operation of an airline. Computational experiments reveal the amenability of the models to real-time interactive usage. A decision support system is currently under development at United Airlines based on this study. We have also identified various avenues for further research in this important airline scheduling problem.

ACKNOWLEDGMENTS

WE WOULD like to thank S. G. Regulinski of United Airlines for his helpful comments. In addition we would like to thank the Editor for his insightful remarks, and the anonymous referees for their suggestions regarding the paper's conceptual development and the model formulations.

REFERENCES

1. R. G. BUSACKER AND P. G. GOWEN, "A Procedure for Determining a Family of Minimum-Cost Flow Patterns," Operations Research Office Technical Report 15, Johns Hopkins University, 1961.
2. M. M. ETSCHMAIER AND D. F. X. MATHAISEL, "Airline Scheduling: An Overview," *Trans. Sci.* **2**, 127-138 (1985).
3. T. A. FEO AND J. F. BARD, "Flight Scheduling and Maintenance Base Planning," *Mgmt. Sci.* **35**, 12, 1415-1432 (1989).

4. I. GERSHKOFF, "Aircraft Shortage Evaluator," presented at the ORSA/TIMS Joint National Meeting, St. Louis, MO (October), 1987.
5. F. GLOVER, D. KLINGMAN AND N. V. PHILLIPS, *Network Models in Optimization and Their Applications in Practice*, John Wiley & Sons, New York, 1992.
6. R. V. HELGASON AND J. L. KENNINGTON, *Algorithms for Network Programming*, John Wiley & Sons, New York, 1980.
7. O. HOLST AND B. SORENSEN, "Combined Scheduling and Maintenance Planning for an Aircraft Fleet," *Operational Research '84*, 735-747 (1984).
8. M. M. SYSLO, N. DEO AND J. S. KOWALIK, *Discrete Optimization Algorithms*, Prentice Hall, Englewood Cliffs, NJ, 1983.
9. D. TEODOROVIĆ AND S. GUBERINIĆ, "Optimal Dispatching Strategy on an Airline Network after a Schedule Perturbation," *Eur. J. Opnl. Res.* **15**, 178-182 (1984).
10. D. TEODOROVIĆ AND G. STOJKOVIĆ, "Model for Operational Daily Airline Scheduling," *Trans. Plan. Technol.* **14**, 273-285 (1990).
11. G. YU AND A. I. Z. JARRAH, "A Mathematical Programming Model for Airline Maintenance Scheduling," Department of Management Science and Information Systems Working Paper 89/90-3-3, University of Texas at Austin, 1990. Also, presented at the ORSA/TIMS Joint National Meeting, Philadelphia, PA (October), 1990.

(Received, July 1991; revisions received March 1992, October 1992; accepted January 1993)