Maintenance and Crew Considerations in Fleet Assignment

L. W. CLARKE
Georgia Institute of Technology, Atlanta, Georgia 30332

C. A. HANE
CAPS Logistics, Atlanta, Georgia 30339

E. L. JOHNSON
Georgia Institute of Technology, Atlanta, Georgia 30332

G. L. NEMHAUSER
Georgia Institute of Technology, Atlanta, Georgia 30332

Given a flight schedule, which is a set of flight segments with specified departure and arrival times, and a set of aircraft, the fleet assignment problem is to determine which aircraft type should fly each flight segment. The objective is to maximize revenue minus operating costs. In the basic fleet assignment problem considered by Hane et al. (1995) a daily, domestic fleet assignment problem is modeled and solved with up to eleven fleets and 2,500 flight legs. This paper provides modeling devices for including maintenance and crew considerations into the basic model while retaining its solvability.

Accepted by Mark S. Daskin.

ABARA (1989) and DASKIN and PANAYOTPOULOS (1989) and the references cited therein.

In the basic fleet assignment problem considered by HANE ET AL. (1995), a daily, domestic fleet assignment problem is modeled and solved with up to eleven fleets and 2,500 flight legs. Daily means that every flight is flown every day of the week. This is a simplifying assumption that keeps the model from growing too large but still yields reasonable solutions. The network is a hub and spoke system of domestic routes. Fleet assignment problems derived from international schedules are very different.

This paper is a sequel to HANE ET AL. (1995) in which we generalize the fleet assignment model to capture certain aspects of maintenance and crew scheduling. Solutions that do not satisfy maintenance constraints are not feasible and modifying them after optimization can significantly reduce the objective function. By ignoring crew scheduling al-
together we may create a situation where the savings in fleet assignment is needlessly offset by excessive crew costs. But it is not obvious how to incorporate maintenance constraints or any aspects of crew scheduling into the fleet assignment model without complicating the model to the extent that it is no longer solvable. The goal of this work is to provide modeling devices for achieving these modifications while retaining its solvability.

**Maintenance**

Airlines must meet FAA (Federal Aviation Administration) standards for maintenance of their aircraft. However, the airlines generally impose much stricter standards upon themselves and the practices differ among airlines to some extent. Although our study is largely based on the practice of one major airline, the ideas should be applicable to all domestic carriers.

Maintenance requirements range from pilot inspections after each flight to heavy C checks that rebuild the aircraft from scratch every few years. The time necessary to perform these maintenance checks varies from minutes to over a month. Since the time frame of this fleet assignment problem is daily, we will be concerned only with those maintenance checks that require less than 24 hours. The heavy C checks are accounted for by reducing the number of available aircraft in the model.

On the daily time scale there are three significant maintenance checks; two types of A checks, denoted M and Av, and B checks. The A checks are the shortest in duration and are on the order of four hours. The M check is an inspection performed every 2 to 3 days on each aircraft, and the Av check includes the M check in addition to an avionics inspection and is performed every 4 to 5 days. Not all stations (airports) are equipped to do maintenance and those stations that do maintenance may do M checks or Av checks for one, several or all fleets. The general rule from maintenance departments concerning the number of A checks is 'more is better.' Thus, the maintenance planners would like to see many A check opportunities (an aircraft on the ground at a maintenance base for a sufficient length of time) in the fleeted schedule.

The B checks require a 10 to 15 hour long stay at the maintenance hangar. These B checks are scheduled by the maintenance department, which informs the fleet planners where aircraft are needed each day and provides a maintenance time window. For example, the B check plan could call for one 747 in Phoenix for 10 hours between 7 pm and 8 am, a 737 and a 757 in Chicago for 12 hours between 5 pm and 7 am, etc. Thus, the fleet planners have to search the maintenance time window for an opportunity to remove an aircraft from service at that maintenance base.

We will refer to A checks as 'short' and B checks as 'long' maintenance. The motivation for this distinction is that the duration of the maintenance visit plays an important role in its modeling. The difficult part is guaranteeing the time window which is essential to do for long maintenance. However, for short maintenance it may not be necessary to guarantee time windows explicitly, in which case modeling is easy.

Fleet assignment can only provide an adequate number of maintenance opportunities. It is not capable of guaranteeing that the intervals between maintenance is sufficiently short or that the maintenance visits are appropriately spaced. Guaranteeing a proper maintenance schedule depends on the routing of the individual airplanes, which is called rotation, (FEo and Bard (1989), GOPALAN and Talloria, ZHU (1994), ZHU et al.) and is solved after fleet assignment. The two major goals of the rotation problem are to link up pairs of segments that have been assigned to the same fleet to create attractive one stop flights and to achieve appropriate spacing between maintenance visits. Typically, the airlines require that each plane in a fleet fly an identical route consisting of all the legs assigned to that fleet, which implies that the route must be a cycle. Note that if fleeting and rotation were combined into one problem, maintenance scheduling could be handled in one step, but a model that included both would be computationally intractable.

**Crew Issues**

For computational reasons, it is not possible to solve a model that does fleet assignment and crew scheduling simultaneously. When the two problems are solved independently, fleet assignment must be done first because the fleet type determines the number and type of crew members needed.

A pairing is a sequence of flights assigned to a crew, beginning and ending at the crew base. A pairing can be from one day long to 3 or 5 days depending on the specific airline operating procedures as long as it satisfies legal and contractual requirements. The crew scheduling problem is to find a set of pairings such that each flight leg belongs to exactly one pairing and that minimizes the cost for time spent not flying and time away from base.

We focus mainly on crew issues for pilots. Pilots are qualified for particular cockpits. The fleet types can be partitioned into crew compatible fleet groups, i.e. two fleet types are crew compatible if pilots qual-
ified for one are qualified for the other. An example of a crew compatible fleet group is the 757, 767 and 767 stretch. They all have the same cockpit but are considered different fleet types because of seating capacity. For purposes of crew modeling they are in the same group.

It is easy to include some crew constraints in fleet assignment. For example, lower bounds on the number of departures of a fleet from crew bases for that fleet and upper bounds on number of flying hours for the fleet are easy to model. But opportunities for keeping down crew costs are not so easy to model in a way that will keep the model computationally tractable. Perhaps, the most important of these is the avoidance of lonely overnights.

A lonely overnight occurs when a crew arrives late at night at a station that is not its base, and the aircraft it arrived on leaves before the crew has had sufficient (usually 11 hours) rest (nonflying time) and there is no other departure of that fleet that day. For example, the only activity at Dullsville for DC9's is an arrival at 11 pm and a 7 am departure. A crew that arrives at 11 pm on Monday cannot crew a departing flight until 7 am Wednesday, a 32 hour layover. Such a long layover in the middle of a crew pairing can increase the time-away-from-base penalty which is part of the formula for determining the crew's wages. The airline must balance the expected cost of this penalty versus changing the fleeting at Dullsville or having the crew fly out as passengers on another flight (deadheading). This problem would be eliminated if the crew remained with its aircraft for an entire pairing (the time a crew is away from its home base). It would also eliminate many problems of tactical crew scheduling due to mechanical breakdowns, bad weather and other delays. However, the aircraft has no limit on the amount of time per day that it flies whereas the crew does. Thus, to achieve high aircraft utilization it is necessary to separate the crews from the aircraft.

Another complicating factor in handling lonely overnights is the use of ground transportation to relocate crews between nearby stations. For example, if Ennuitown is near Dullsville, then a DC9 crew that arrives at Ennuitown at 7 pm could be driven to Dullsville where they would receive sufficient rest and fly the 7 am departure. The crew arriving at Dullsville would be driven to Ennuitown and fly out a late morning flight. In this case Dullsville and Ennuitown are called co-terminals.

Finally, the modeling is further complicated by the use of crew compatible fleet groups to avoid lonely overnights. For example, a crew that flies a 757 into a station may fly a 767 out to avoid a lonely overnight.

Outline of the Paper

In this paper we show that maintenance time windows and the avoidance of lonely overnights can be included in the fleet assignment model without destroying its solvability. Section 1 reviews the basic model of the fleet assignment problem and presents some techniques to reduce the model size. Section 2 details the changes to the basic model to capture the maintenance constraints and Section 3 does the same for crew considerations. Section 4 presents the computational results to show that the more realistic model remains tractable.

1. BASIC MODEL

THE FLEET ASSIGNMENT solution must satisfy balance constraints that force the aircraft to circulate through the network of flights. The balance constraints are enforced by modeling the activity at each station with a time line for each fleet, see Figure 1. This time line has entries designating the arrivals and departures from the station for each fleet. Each departure (arrival) from the station splits an edge and adds a node to the time line at the departure (arrival + refueling/baggage handling) time. The nodes created at the arrival station and departure station are connected by the decision variable representing the assignment of that fleet to that flight. Thus, aircraft balance is enforced by the conservation of flow equations for a time-expanded multicommodity network. We make the time line a cycle, which forces the solution to be a circulation through the network. The circulation arises from the balance constraints and the lack of source or sink nodes in the network. Since the network has a time span of 24 hours, the circulation defined by the solution defines a daily schedule.

The time line’s purpose is to preserve aircraft balance, but it also must allow the proper aircraft flight connections. Therefore, the placement of the arrival end of the flight arc must coincide with the time when the aircraft is ready to takeoff. Any earlier placement could violate the feasibility of having a single aircraft fly two consecutive flights. We use the term ‘ready time’ to indicate the time at which
the arriving flight is ready to takeoff. The ready time
is the arrival time plus the needed 'turn time.' The
turn time is the time needed to get the plane ready
for another flight. This time includes, passenger de-
barkment, refueling, cleaning and other similar
needs. Turn time is fleet and station dependent be-
cause larger aircraft and busier stations require
more time.

1.1. Mathematical Model

The set of stations serviced by the schedule is
denoted by \( C \), the set of available fleets by \( F \)
and the number of aircraft in each fleet is \( S(f) \) for \( f \in F \). The
set of flights in the schedule is denoted by \( L \), with elements \( (o, t) \), or \( (o, d, t) \), with \( o, d \in C \) and \( t \) the depar-
ture time. The set of flight arcs \( O(f) \) for \( f \in F \) denotes
the arcs whose time span contains 3 am EST. The
actual time is arbitrary but picking an early morn-
ing time reduces the size of \( O(f) \). The
nodes in the network, \( N \), are enumerated by \( [o, t] \), with \( f \in F \), \( o \in C \) and \( t \) a
time. We use \( t^- \) to denote the time preced-
ing \( t \) and \( t^+ \) the following time. When arrivals and
departures occur simultaneously, arrivals precede
departures in the time line. This is done so that
allowable connections are always later than the ar-
riving flight. The last node in a time line is \([fot_a]\), the
node that precedes 3 am EST; its successor is \([fot_{1}]\).

The decision variable \( x_{foil,t} \), also written as \( x_{it} \), has
value 1 if fleet \( f \) flies the flight leg from \( o \) to \( d \)
departing at time \( t \), and 0 otherwise. The other vari-
ables that appear in the model are 'ground arcs'
which count the number of aircraft on the ground
on each station at every point in time for each fleet.
These ground arcs are \( y_{foil,t} \) with \( f \in F \), \( o \in C \), and \([t, t^+] \) the time interval covered by the arc. The basic
model is shown below.

Model 1. Integer Programming Formulation of Basic
Fleet Assignment Model.

\[
\min \sum_{i \in L} \sum_{f \in F} c_i x_{it} \\
\sum_{i \in L} x_{it} = 1 \forall i \in L \\
\sum_{d \in O(f)} x_{foil,t} + y_{foil,t^-} - \sum_{d \in O(f)} x_{foil,t^+} - y_{foil,t^+} = 0 \forall \{fot\} \in N \\
\sum_{i \in C} x_{it} + \sum_{o \in C} y_{foil,ot} = S(f) \forall f \in F \\
y_{foil,t} \geq 0 \forall \{fot\} \in N \\
x_{it} \in \{0, 1\} \forall i \in L \text{ and } f \in F
\]

The objective coefficient, \( c_{it} \), represents the cost of
assigning fleet type \( f \) to flight \( i \). It is a composite of
the opportunity cost of choosing too small an aircraft
for the passenger demand and operating costs such as
fuel, in flight crew costs (the number of crew
members multiplied by the flying time and pay
rate), and others. There are three main sets of con-
straints in the basic model. The first set is the cover
rows, forcing each flight leg to be flown by exactly
one fleet. Thus, the solution cannot eliminate un-
profitable flight legs or relocate aircraft on non-
scheduled flights. The second set of constraints is
the 'balance' constraints. These are the flow conser-
vation equations for the nodes of each fleet that force
the flow to be a circulation. The final set of con-
straints in the basic model is the fleet size con-
straints which count the number of aircraft of each
fleet used by the solution. These constraints 'slice'
each fleet network at 3 am EST and count the planes
in each fleet. The first term counts planes that are
flying at 3 pm and the second term counts the planes
that are on the ground.

1.2. Simple Extensions

The basic model can be made more realistic by the
addition of several sets of constraints that are easy
to model relating to gate availability, noise limita-
tions, crew considerations and maintenance. We
only mention the crew and maintenance constraints
here.

For maintenance that occurs every \( x \) days, we can
require 100\% of the planes to overnight at stations
that can do this maintenance. These constraints are
of the form

\[
\sum_{o \in CM(f)} y_{foil,ot} \geq NM(f) \forall f \in F
\]

where \( NM(f) \) is the number of planes of fleet \( f \) that,
on the average, need maintenance on a daily basis,
and \( CM(f) \) is the set of stations that can perform
the specified maintenance. However, (6) does not guar-
antee the existence of the time windows to actually
do the maintenance, which is a much greater prob-
lem for long maintenance than for short mainte-
nance.

For crews, we can place a lower bound on the
number of departures from crew bases for each fleet.
These bounds are given as

\[
\sum_{i \in MT} x_{foil,ot} \geq RD(o, f) \forall (o, f) \in CB(o, f)
\]

where \( CB(o, f) \) is the set of station-fleet pairs that
are crew bases and the minimum number of re-
quired departures for each member of the set is
\( RD(o, f) \). The summation over \( MT \) collects morning
flights, which is the usual time for crews to begin their work assignment. There is also an upper bound on the number of flying hours for each fleet group \( k \) given by

\[
\sum_{f \in F(k)} \sum_{i \in L} \tau_{f_i} x_{f_i} \leq PH(k) \forall k,
\]

where the flying time for each flight leg is represented by \( \tau_{f_i} \) and the maximum pilot-hours available for fleets in the set \( F(k) \) is \( PH(k) \) for each \( k \).

### 1.3. Preprocessing

Several preprocessing steps must be taken in order to make the basic model computationally tractable. Indeed, without them the basic model cannot be solved in days of CPU time on a RISC workstation. Only an overview is given here, the details can be found in HANE ET AL. (1995). First among these is 'node consolidation'. The balance constraints (2) contain a separate node for each arrival and departure at every station. However, their purpose can be achieved with many fewer constraints. There is no benefit to placing an arrival node in the network earlier than the first departure to which it can connect. Similarly, departure nodes can be moved earlier to coincide with the latest arrival that can make the connection. Thus, a node represents a time interval in which there is a sequence of consecutive arrivals followed by a sequence of departures.

The other modifications rely on the hub-and-spoke topology of the network. At the hubs, arrivals and departures occur in groups called complexes. A complex is a series of arrivals followed by a series of departures which allow the passengers to make connections at the hub. At the spokes there is often just a single arrival followed by one departure, or at least long periods with no aircraft on the ground. If we can determine the periods when there are no aircraft on the ground at a station in advance of fleeting, then we can eliminate the corresponding ground arcs from the model. This preprocessing is called making 'islands', since the time line at a spoke becomes disconnected intervals representing times when there are aircraft on the ground.

Once the ground arcs have been eliminated, certain complexes at spokes will have only one flight leg arriving and one departing. These isolated pairs of flights must be flown by the same fleet, or an aircraft would be on the ground where we assumed there are none. The \( x \) variable for one of these flights can be eliminated in the same fashion when the schedulers insist that the plane arriving on one flight leg be the same plane departing on another, even if there are alternate choices. This is called a forced hookup and is done to market the attractiveness of a one-stop flight.

During island processing, it becomes apparent that fleets with longer turn times cannot make connections that the flight schedulers intended. For example, if a pair of flights that define an island are only 25 minutes apart, then it is unreasonable to assign an L-1011 with a 60 minute turn time to the inbound flight. If an L-1011 does fly the inbound flight, it violates the island structure, and some other aircraft must violate the island structure to connect with the outbound flight. This analysis shows that if the wrong aircraft type is assigned to either of these flights an extra aircraft must overnight at this station. By judiciously choosing where extra aircraft are allowed to overnight, we can eliminate the \( x \) variables that correspond to missed connections. In other words, we do not allow an L-1011 to fly the flight legs leading to an extra plane on the ground.

All of the island preprocessing results hinge on determining the number of aircraft on the ground at each spoke station throughout the day. However, the schedule does not completely determine this quantity because there generally is flexibility in choosing the location of the aircraft and, in addition, it is possible, but unlikely, that there are more than enough aircraft to fly the schedule. Thus, although island processing is heuristic, when applied appropriately, it is very effective at significantly reducing the problem size without affecting the objective value. The reason for not affecting the objective is that the restriction that no extra plane be on the ground is not applied at hubs or maintenance stations, which is usually where they would have to be in any case in order to take advantage of them for back-ups, pilot training, or maintenance opportunities.

### 2. MAINTENANCE CONSTRAINTS

This section describes the constraints for long and short maintenance opportunities. Our distinction between the two depends on the relative duration of the maintenance to the available maintenance time window. In particular, long maintenance requires more than half of the maintenance time window to complete and short maintenance requires no more than half of the maintenance time window to complete. However, it is almost always the case that B
checks are long, it being very unlikely that the maintenance time window is 20 hours or more, and A checks are short.

2.1. Long Maintenance

The data provided to the model for scheduling the B checks is a list specifying the location, fleet number, aircraft, duration and maintenance time window for each 'maintenance requirement'. We denote this list by the set \( PL \). For example, part of the daily requirement for B checks could be two Boeing 757's in Chicago for 11 hours between 6 pm and 8 am. We call the assignment of one 757 to go to maintenance at Chicago at 6:35 pm a maintenance 'visit' to distinguish it from the maintenance requirement. The maintenance time window for this example is 6 pm to 8 am, but maintenance must begin between 6 pm to 9 pm (11 hours before 8 am). Thus, the scheduler must find two aircraft that are either on the ground at Chicago at 6 pm or arrive before 9 pm to send to the maintenance hangar, returning to the terminal area 11 hours later. This problem is similar to selecting a departure time for an 11 hour out-and-back flight at Chicago.

Having presented the problem in terms of selecting departure and arrival times, we now show how the maintenance visit can be modeled by constructing a set of arcs that act like flight arcs of the required duration. There are two main differences between these maintenance arcs and the regular flight arcs. First, the maintenance arcs depart and arrive at the same station. We say that they 'leapfrog' the maintenance visit. Second, for each maintenance requirement there is a set of arcs corresponding to allowable maintenance visits and the model chooses the correct number (2 for this example) of visits from the arcs in this set.

Let \( M_p \) be the number of aircraft needed to satisfy maintenance requirement \( p \). For fixed \( p \), \( m_{pj} \) is the set of arcs that could be used to satisfy maintenance requirement \( p \). Thus we have

\[
\sum_{j} m_{pj} = M_p \quad \forall p \in PL.
\]  

The variable \( m_{pj} \) is a nonnegative integer. It also appears in two balance rows since the arc goes from one node to another node. It also may appear in the aircraft fleet size constraint depending on its begin and end time. Since maintenance requirement \( p \) is associated with a specific fleet, station and time, these indices do not appear in the above constraint. A representation of the long maintenance construction is shown in Figure 2. Note that the \( x_m \) arcs are for short maintenance and are discussed later.

The only way this construction could be invalid is if one aircraft could be counted twice. To double count an aircraft it must depart to the hangar, return and depart to the hangar again. This is impossible if the maintenance duration exceeds half the length of the maintenance time window. Hence, this construction is generally valid for long maintenance, but not for short maintenance.

Adding the minimum number of \( m_{pj} \) arcs is crucial since each additional arc potentially destroys a consolidated node, adding another ground arc and node. Thus, each \( m_{pj} \) arc can add two variables and a row to the formulation. The key to determining which \( m_{pj} \) arcs are required in the formulation is the realization that each arc should terminate at a flight's departure node. If some arc's head is not coincident with a departure, then the corresponding maintenance visit can be delayed to make the aircraft return just in time for a departure. This has no effect on the set of connections the aircraft can make and thus does not affect the feasible set of assignments. Thus nothing can be gained by adding more maintenance arcs, but if any of these were removed, then the model would miss an opportunity for allowing a maintenance visit followed by that departure.

Continuing with the Chicago example, the earliest return from maintenance is 5 am (6 pm + 11 hours). If the first departure at Chicago after 5 am occurs at 5:40 am, then the first arc for this requirement begins at 6:40 pm and continues until 5:40 am. Thus, we do not need separate arcs for each arrival between 6 pm and 6:40 pm, a large savings during a busy time of day. If the next arrival after 6:40 pm

![Fig. 2. Maintenance, with leapfrog and split arcs.](image-url)
occurs at 6:55 pm, then we know the earliest an aircraft assigned to this flight can return to service is 5:55 am. Now, we proceed as before by sliding this arc from (6:55 to 5:55) to (6:55 + t to 5:55 + t) where 5:55 + t is the time of the next departure.

2.2. Short Maintenance

Fleet types with many maintenance stations can have good rotations without explicitly including constraint (6). Other situations require (6) and there are also times when (6) is not sufficient. In the last case an alternative approach can be used to guarantee short maintenance opportunities. Let \( PS \) be the set of required short maintenance visits.

Leapfrog arcs, similar to those used for the B checks, can be added to indicate specific maintenance visits. As stated in the previous section this construction is not always valid for short maintenance. The reason is that it may be possible for a plane to depart to the hanger, return to the gate and depart again to the hanger within the same maintenance time window, thus being double counted. Therefore, unless the maintenance time window is less than twice the length of the maintenance time, a more complex construction is required.

The easiest modeling method for avoiding double counting is to associate the maintenance visit with a particular flight leg. Let \( x_{pi} \) be a variable associated with a flight arriving at a maintenance base such that there is a time window in which maintenance \( p \) can be done. Now create the new \([0,1]\) variable \( x_{mpfi} \) which has value 1 if the aircraft from fleet \( f \) that flies flight \( i \) goes directly to maintenance \( p \) upon arrival. The departure time of \( x_{mpfi} \) is the same as \( x_{pi} \) and its ready time is flight \( i \)'s arrival time plus the time for maintenance \( p \). Thus, the variable \( x_{mpfi} \) has all the same coefficients as \( x_{pi} \) except the ready time at the arrival station and \( x_{mpfi} \) appears with coefficient 1 in the constraint for maintenance \( p \). Then, since each flight is only flown by one aircraft and \( x_{pi} \) and \( x_{mpfi} \) appear in the same cover row, the model cannot double count. We call this construction 'splitting' \( x_{pi} \). For each arriving flight that could have maintenance completed in the maintenance time window, (1) is replaced by

\[
\sum_{f} x_{pi} + \sum_{p} x_{mpfi} = 1. \quad (1')
\]

The objective value of \( x_{mpfi} \) is \( c_{pi} \) minus a small quantity to encourage the model to schedule more maintenance visits. These variables have the added benefit of making the number of maintenance visits easily countable since no solution would have an aircraft available during a short maintenance time window without maintenance being scheduled.

The drawback of this formulation is the excessive number of integer variables it may add to the model. It adds one integer variable for each type of maintenance for each arrival that can have maintenance completed in the maintenance time window. One way to avoid this large number of variables is to add leapfrog arcs at the beginning of the maintenance time window, then add \( x_{mp} \) arcs for those flights that arrive after the end of the first leapfrog arc and could have maintenance. Figure 2 shows that no leapfrog arc can be added after the arrival of \( x_{p} \) since that would allow flow on the first leapfrog arc to be double counted. Therefore, the arc \( x_{mpfi} \) is added.

The maintenance constraint

\[
\sum_{j} (m_{pj} + x_{mpj}) = M_{p} \quad \forall p \in PS \quad (10)
\]

is similar to the one for long maintenance.

3. CREW CONSTRAINTS

This section addresses the modeling issues involved in avoiding lonely overnights. Our basic approach is to put a cost on each lonely overnight and then to let the optimization model balance the costs between between lonely overnights and fleeting.

3.1. Legal Rest Arcs

This modification of the model encourages aircraft to remain with the crew by adding leapfrog arcs with a duration of the legal rest time, which is usually 11 hours. The differences between the long maintenance and crew modifications are the crew arcs have a less well posed crew time window, and the crew arcs do not have a specified amount of flow. The legal rest variables, \( m_{p1} \), are given a negative cost in the objective to encourage their use. These variables are required to be nonnegative and integral, but are not necessarily \( 0, 1 \).

In order to determine the correct crew time window for legal rest arcs, we must make an assumption about which arrivals end work periods and which departures begin them. This assumption is easily made at low traffic spoke stations where the schedule usually has a few nighttime arrivals and the first flights in the morning are departures to the hubs. In this case, the nighttime arrivals end duties and the morning departures must begin them.

When the station is busier, we must identify an island that will contain all the legal rest arcs at that station. All legal rest arcs must have both endpoints within one island, otherwise these arcs would violate the no aircraft assumption we used to make the island. We choose the island containing 3 am local time.
From the first arc in the overnight island a legal rest arc of the appropriate length is constructed. The head of the arc is extended to reach the next departure. This procedure is repeated for the next arrival. If the head of the rest arc of the second arc falls in the same location as the head of the previous arc, the previous arc is deleted. Deleting the first rest arc still allows the first crew to receive the rest benefit by using the second arc. This process is repeated for each arrival, deleting the previous arc when appropriate. Thus an island with several arrivals may have only one or two strategically placed rest arcs.

In Figure 3 a rest arc is draw from arrival a (20:00 hrs) and extending 11 hours and terminating at 07:00. The arc is then extended to the next departure, which is labeled e. The process is repeated for arrival b, which generates an arc from b to e. Since these two arcs both arrive at point e, the arc beginning at point a is deleted leaving the one arc shown in the figure. If an 11 hour rest arc is extended from arrival c, it falls outside of the overnight window. Therefore, only one rest arc is used for this set of flights.

3.2. Midday Breakouts

Another option to reduce the number of lonely overnights is assigning a midday flight to the same fleet. If the only activity for DC-9's at Dullsville is an 11 pm arrival and a 7 am departure, then the planners may want to schedule a connecting pair of flights between 11 am and 3 pm for a DC-9. With two DC-9 arrivals, one crew arrives in the middle of the day and leaves at 7 am the next morning and the other arrives at 11 pm and departs in the middle of the next day. This is called a midday breakout. While this solution still requires two crews to overnight at Dullsville, it reduces the time-away-from-base which is the primary source of the excess cost.

For the model to take advantage of this option it must be able to count the number of potential lonely overnights and midday departures assigned to the same fleet. Then, the actual number of lonely overnights for a fleet at a station is the positive part of the number of crews without a legal rest minus the number of midday departures assigned to that fleet. The legal rest arcs allow us to count the number of crews without a legal rest as the flow into the node of the first departure after the last legal rest arc. The number of midday breakouts is counted as the departures in a midday time window that are assigned to fleet f, that is \( \sum_{i \in MD(f, o)} x_{fi} \), where \( MD(f, o) \) is the set of midday departures for fleet f at station o.

The rules for determining which departures belong in \( MD(f, o) \) are quite simple. The specific aircraft that is assigned a departure in the midday set arrived earlier that morning. It is desirable for its incoming crew to have accumulated enough time to form a reasonable duty period (about 5 hours) because it will now be taken off duty until the next morning. Thus, we do not want the midday time window to begin too early. It is also desirable for the crew that flies the mid-day departure to be able to accumulate enough time for a duty period. Thus we should end the midday time window early enough.

The actual constraint added to the model after adding all the legal rest arcs is

\[
\sum_{i \in MD(f, o)} x_{fi} - LO_{fo} = 0 \quad \forall o \notin CB(f), f \in F
\]

where \( y_{for} - \gamma \) is the ground arc entering the node of the first departure after the last legal rest arc, hence its flow is the number of crews that cannot remain paired to the aircraft with which they arrived and \( LO_{fo} \) is a nonnegative variable which carries an objective coefficient equal to the cost of one lonely overnight for fleet f at station o. This penalty will force \( LO_{fo} \) to be the number of crews actually experiencing a lonely overnight for that fleet at that station. The set of crew bases for fleet f is \( CB(f) \).

In Figure 3, the crew arriving on flight c did not have a legal rest. The plane for flight c may depart on flight d. The crew can go out on flight d the following day (a lonely overnight) or get a midday
breakout on flight $h$. If the crew coming in on flight $c$ goes out on flight $h$, then the crew coming in on $g$ must go out on flight $d$.

3.3. Crew Networks

The previous two sections discussed how minimizing the number of lonely overnights can be handled within the fleet assignment model when each crew can only fly one fleet of aircraft. This is not always the case, as some fleets need to be distinguished in the model for revenue purposes but have the same crews. In these cases a crew can arrive at a station via fleet $f$ and fly out on fleet $k$.

If a crew can fly the aircraft of fleets $f$ and $k$, then we say $f$ and $k$ are crew compatible fleets. To allow a crew to switch between compatible fleets, we create a separate network for the crews. This network is a time line like the line for the aircraft except that it is only created for the overnight island. It is this crew island at each non-crew base station that contains the legal rest arcs and counts the number of crews that do not have legal rests. Thus if crew compatible fleets are present in the model then the legal rest arcs appear in the crew time line, not the aircraft time line, and are denoted $l_{r_{kot}}$, where $k$ is a crew type, $o$ a station and $t$ the time of the tail of the arc.

Now each $x_p$ at value 1 brings an aircraft to the aircraft time line and a crew to the crew time line at the destination. The only new arcs needed are ground arcs for the crew network that count the number of crews at the station during the crew island. Denoting the fleets flown by crew type $k$ by $F(k)$ and the ground arcs for crew type $k$ by $y_{c_{kot}}$, we write the balance constraint for a node in the crew time line at station $o$ as

$$\sum_{i \in F(k)} x_{f_{di}} + y_{c_{kot} \cdot t} + l_{r_{kot}} - \sum_{i \in F(k)} x_{f_{ot}} - y_{c_{kot} \cdot t} - l_{r_{kot}} = 0 \quad (12)$$

The midday breakout constraint is also moved to the crew network and is written as

$$y_{c_{kot} \cdot t} - \sum_{i \in F(k)} x_{f_i} - L_{O_{ko}} \geq 0 \quad (13)$$

These networks and constraints are created only for stations that are not crew bases for the crew types.

3.4. Co-Terminals

There are still more options available for avoiding lonely overnights. In metropolitan areas serviced by more than one major airport, e.g. New York, with JFK, La Guardia and Newark, a crew that arrives at one airport can be driven to another airport and depart from there. These airports are called co-terminals. For example, a crew arrives at JFK at 7 pm and is scheduled to depart at 9 am (a legal rest), another compatible crew arrives at 9 pm at La Guardia and departs at 7 am (not a legal rest). The crew that arrives at JFK can be driven to a hotel near La Guardia and fly the 7 am departure, while the La Guardia arrival flies out the JFK departure. As long as the travel time between the airport hotels is less than 60 minutes, the crew planner has avoided a lonely overnight at La Guardia.

This ferrying of crews between co-terminals can be captured by arcs representing the transport of the crews linking the crew networks. The cost of flow on these arcs is the ground transportation rate between the airports. For each crew that is ferried from $A$ to $B$ there must be a crew that is ferried from $B$ to $A$ to preserve the balance constraints of the crew network.

To cover every possible ferry opportunity, each arrival at $A$ generates a tentative ferry arc to $B$ that terminates at $B$, $t_{AB}$ minutes after its arrival at $A$, where $t_{AB}$ is the travel time from $A$ to $B$. There are two reasons a crew should be ferried from $A$ to $B$. The ferry may allow the crew to make a legal rest, or it balances the number of crews that are ferried back and forth. A crew that is ferried to get a legal rest should have the ferry arc it travels terminate with the beginning of a legal rest arc at $B$, and a crew that is balancing the flow should have its ferry arc terminate with a departure at $B$. Both of these statements are motivated from the same arguments as sliding the long maintenance arcs presented earlier. Thus, the tentative ferry arc from an arrival at $A$ is delayed until its terminus at $B$ coincides with a departure or the head of a legal rest arc.

Adding this many arcs at all co-terminals and crews may increase the size of the model unnecessarily because they are used so infrequently. However, crew planners can identify specific co-terminals and crews that are more likely to use the ferry opportunities and only these anticipated opportunities can be modeled, or after an initial solution is obtained the ferry opportunities can be added to the model.

4. COMPUTATIONAL RESULTS

4.1. Data Sets

Computational trials were performed on three data sets provided by a major U.S. airline representing different times of the year. For most airlines winter and summer schedules are very different in terms of both the number of flights and the destinations. For each of the three data sets we solved the fleet assignment problem for three different constraint sets. The first set is an assignment without maintenance or crew constraints. The second set has
only long maintenance constraints and the third set has both long maintenance and crew constraints. We have not implemented the special modeling devices for dealing with short maintenance because solutions without explicit consideration of M and A checks have met the requirements satisfactorily.

The sizes of the different models for each data set are shown in Table I. The basic model (without maintenance or crew constraints) is represented in the table by 1B–3B. The constraints used in this model are (1)–(5) and (8).

When the B check maintenance is modeled, constraint (9) is added to the formulation. This model is represented by rows 1M–3M and adds several integer variables that represent the opportunity of a flight to go to maintenance. These variables are in special ordered sets, where only one variable can have a nonzero value. The number of variables added depends upon the flights that are entering and leaving the time line. One long maintenance requirement can add as many as 50 integer variables to the model. Each long maintenance requirement also adds one cover constraint that forces the maintenance requirement to be satisfied. The maintenance variables also change the structure of the time line. This changes the balance constraints in the model. The number of balance constraints usually increases.

When crew considerations are included in the model, constraints (11)–(13) are added to the formulation. This model is represented by rows 1MC–3MC. For a particular crew and station, an entire crew time line is built. The flights in the time line do not generate additional variables. However, the ground arcs and the rest arcs do generate additional variables with the number dependent on the structure of the time line. There have been as many as 15 ground arcs and 5 rest arcs added to a crew time line. One mid-day breakout variable is also added.

Each of these variables is modeled as continuous and naturally are integer due to the structure of the model. Balance constraints are added (one more than the number of ground arcs) for each time line and one mid-day break out constraint is added. For the problems solved, co-terminals were not part of the model.

In addition to the variables and constraints added to each of the models for the representation of the special needs the preprocessing described in Section 1.3 results in minor differences in the flights that are allowed to be flown by each fleet.

### 4.2. Algorithm/Implementation

There are two distinct phases to get a solution to the fleet assignment problem as defined in this paper. The first phase is generate the model as described in the previous sections. This phase includes the preprocessing described in Section 1.3.

The second phase is to solve the model using the steps given below.

#### Solution Steps

1. Aggregate using optimizer’s algebraic preprocessor.
2. Solve LP by dual steepest-edge simplex.
3. Fix flight variables with value ≥ 0.99 to 1.
4. Aggregate using optimizer’s algebraic preprocessor.
5. Solve LP by dual steepest edge simplex.
7. Solve MIP by branch and bound.

We use the mathematical programming callable library OSL Release 2 (Drucker et al., 1991). OSL has an LP preprocessor (step 1) that can reduce the size of the problem. Solving the dual with steepest-edge pricing was found to be the best way to optimize the LP. Fixing variables in step 3 takes away the guarantee of optimality. However, we have empirically determined that the change in objective function caused by this step is insignificant. Once variables have been fixed, the preprocessing in step 4 reduces the problem to a size significantly smaller than the original. For this reason the LP solution time in step 5 is small. OSL’s MIP preprocessor (without supernodes) reduces the size further. The branch and bound portion (step 7) is the most complex part of the algorithm.

Special ordered sets (SOS) and priorities are used to improve the efficiency of the branching. Constraint (1) is a SOS of the variables $x_f$. SOS branching works well for this partitioning constraint if the flight variables are ordered by the seating capacity of the aircraft.
The general branching strategy employed is to make major decisions first. Variables representing the number of planes used receive the highest priority. Variables representing where planes overnight receive second highest priority. These variables are normally designated continuous, and are naturally integers after the flight variables are resolved. Designating these variables as integers helps to find good solutions early. The next variables in priority rank are the maintenance variables. Removing a plane from service for 10 hours is a significant decision. The flight variables have the lowest priority. Different sets of flight variables receive different priorities. The flight sets with widely varying costs receive higher priority. Fleet types are ordered by seating capacity, which has some correlation to cost. A variance number is determined by taking the absolute value of the cost difference for adjacent fleet types. This variance number is used to place the flights into a small number (5 was used for our computations) of different priority levels. Sets with higher variance received a higher priority.

4.3. Results

<table>
<thead>
<tr>
<th>Name</th>
<th>Node First</th>
<th>Time First</th>
<th>Node Last</th>
<th>% Gap</th>
<th>Number Solutions</th>
<th>Number Nodes</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1B</td>
<td>8</td>
<td>17</td>
<td>69</td>
<td>optimal</td>
<td>4</td>
<td>69</td>
<td>23</td>
</tr>
<tr>
<td>2B</td>
<td>3</td>
<td>13</td>
<td>3</td>
<td>optimal</td>
<td>1</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>3B</td>
<td>5</td>
<td>13</td>
<td>5</td>
<td>optimal</td>
<td>1</td>
<td>52</td>
<td>16</td>
</tr>
<tr>
<td>1M</td>
<td>145</td>
<td>55</td>
<td>468</td>
<td>0.237</td>
<td>2</td>
<td>1000</td>
<td>236</td>
</tr>
<tr>
<td>2M</td>
<td>145</td>
<td>47</td>
<td>344</td>
<td>1.397</td>
<td>2</td>
<td>1000</td>
<td>207</td>
</tr>
<tr>
<td>3M</td>
<td>37</td>
<td>22</td>
<td>108</td>
<td>0.366</td>
<td>2</td>
<td>1000</td>
<td>160</td>
</tr>
<tr>
<td>1MC</td>
<td>30</td>
<td>49</td>
<td>30</td>
<td>0.487</td>
<td>1</td>
<td>1000</td>
<td>325</td>
</tr>
<tr>
<td>2MC</td>
<td>19</td>
<td>44</td>
<td>896</td>
<td>0.288</td>
<td>3</td>
<td>1000</td>
<td>295</td>
</tr>
<tr>
<td>3MC</td>
<td>71</td>
<td>48</td>
<td>466</td>
<td>0.057</td>
<td>5</td>
<td>1000</td>
<td>254</td>
</tr>
</tbody>
</table>

The most important conclusion is that the additional constraints provide much better solutions. Without the maintenance constraints, it was not possible to find any place to do long maintenance for any of the fleets. This is easy to see since a plane needs to be on the ground for 12 continuous hours. Short maintenance does not cause a problem and have not been explicitly considered.

In an attempt to show the benefits gained by adding the maintenance constraint and the crew constraints, we developed a relaxed version of the problem where all the maintenance and crew constraints were in the model with minor modifications. The maintenance constraints (9) were modified to read

\[ \sum_j m_{pj} \leq M_p \quad \forall p. \quad (14) \]

This is a relaxation that allows fewer than the required number of planes to receive maintenance. A cost was placed on the maintenance arcs of $-1$. Since ground arcs have zero cost, the small negative cost on the maintenance arcs will attract all possible maintenance opportunities. Since the value is very small it will not change the fleeting to get a maintenance opportunity. In other words, we are avoiding alternate optima where the plane can be sitting on the ground or in maintenance with no effect on the objective function value. The crew network is not modified. However, the negative cost in the objective function that is placed on the legal rest arcs, $r_{rest}$, is reduced to $-1$. This will not alter the fleeting to take a legal rest arc. The penalty for lonely overnights (13) is removed.

Table III shows a comparison of the relaxed formulation and the proper formulation. The intent of this table is to show what is gained in the quality of the solution when the maintenance and crew constraints are added. The quality measures used are in the first column. There are 33 long maintenance visits specified in the input. The 'rest' rows indicate the number of crews that have a legal rest at the end of the day. The higher this number the better the
solution. The 'lonely' rows count the number of crews that sit on the ground more than 24 hours. It is desirable for this number to be small. The 'midday' rows indicate the number of crews that do not have a legal rest and begin their next day's flying several hours later than is legally possible. These crews do not have a lonely overnight. A midday breakout is much better than a lonely overnight, but not as good as a legal rest. The cost row has been scaled to show the relative change in the objective value. The cost shown is only the fleeting costs, all bonuses or penalties have been removed.

The columns represent slightly varying formulations. The 'relaxed' column is the basic model with the maintenance and crew constraints present only to count the number of occurrences. The relaxed formulation was not capable of satisfying any maintenance with the current fleeting. The 'mtc' column satisfies the maintenance constraints by changing the fleeting. Surprisingly, this is done with no increase in operating cost. In evaluating the ability to crew a given fleeting we would like as many legal rests as possible and as few lonely overnights as possible. The midday breaks become dependent on these previous two numbers. The crew considerations are modeled as soft constraints, violations pay a penalty cost. Also, legal rests are encouraged by a bonus in the objective. Therefore, the value of the penalties and bonuses will determine how much emphasis is placed on a good crew solution versus the basic fleeting solution. In Table III the relaxed model can be compared against three different sets of bonus/penalty values for the crew considerations.

Each of the crew models has progressively more rests and fewer lonely overnights. This is achieved with progressively higher costs in the fleeting.

We have demonstrated that extensions of the basic fleet assignment model to include maintenance and crew issues can be modeled and solved with a reasonable increase in computer time. These enhancements address the most important shortcomings of the basic model and improve the quality of the solution.

ACKNOWLEDGMENTS

This work was partially supported by AFORS and NSF grants DDM-9115768 and SES-9122674. We thank the employees of Delta Air Lines Inc., for their assistance. We also thank two referees and the associate editor for comments and suggestions that led to improvements over the original paper.

REFERENCES


(Received August 1994; revision received: May 1995; accepted August 1995)