Block Time Estimation and Robust Airline Scheduling

1. Introduction

In a recent article (Associated Press 2007) the Associated Press reported that the U.S. airline industry's on-time performance (OTP), through the first eleven months of the year 2007, was the second worst on record. According to the U.S. Department of Transportation (DoT), a flight is delayed if it arrives at its destination gate 15 minutes or more after its scheduled arrival time. Flight delays and cancelations have been attributed to several causes some of which include weather conditions, airport congestion, national air-space congestion, aircraft maintenance related issues, and more recently airline security related services. Consequently, such delays lower service reliability and adversely affect a commuter's travel experience. While some of the causes of delays, such as weather conditions, are beyond the control of the airlines, previous research shows that some causes of delays are attributable to the network and scheduling design decisions of an airline. For example, while an airline develops its hub-and-spoke network it typically does not account for the congestion externality imposed on other carriers operating out of the same hub stations. In a recent paper Mayer and Sinai (2003a) empirically demonstrate that the gains from hubbing activities offset the costs incurred by flight delays and congestions. In such cases congestion pricing, at certain capacity constrained airports, may be a solution to elevate the problem. In a companion paper Mayer and Sinai (2003b) also hypothesized that wage cost minimization and aircraft utilization maximization result in airlines flying with very tight schedules. Such objectives are typical in most airline planning systems, which are designed to achieve cost efficient resource utilization.

An airline schedule comprises of a list of flights and specifies the origin, destination, scheduled departure, and arrival time of each flight in the airline's network. A critical component of the schedule development activity is the estimation of flight block-times. A flight block-time is defined as the total elapsed time between the time an aircraft pushes back from its departure gate and arrives at its destination gate. The block-time comprises of several components including taxi-out time, enroute time, and taxi-in time. Each of these components is subject to different causes of delay and the total block-time delay is the sum of all individual component delays. Since airline schedules must be published well in advance of the actual day of operation, block-times for all the flights in the schedule, are typically estimated using block-times of similar flights operated in the past. The DoT OTP metric is computed against these published flight block-times. Most airline operations are compared based on their OTP rankings and hence airlines perceive their OTP as an important operational measure of their schedule reliability. However, research indicates that airlines fail to adequately estimate block-times and typically do not incorporate uncertainty in their planned schedules. Since most planned resource costs, such as aircraft and crew utilization costs, depend on the cumulative hours in a schedule, airlines face a key trade-off decision between adjusting (increasing) flight block-times to improve schedule reliability and incurring additional planned costs. Using data made available by the Bureau of Transportation (BTS), Deshpande and Arikan (2007) argue that airlines systematically "under-schedule" flights, i.e., the amount of block-time allocated for a flight is less than the average block-time expected for the flight. Conversations with planners at a large U.S. carrier suggested that airlines do not judiciously allocate block-hours to scheduled flights to balance costs versus operational benefits. Typically, planners use ad-hoc techniques to either lower or raise block-hours across the entire flight network in the hope of increasing OTP. Results in Deshpande and Arikan (2007) also corroborate these findings and indicate that airlines do not maintain consistent service levels by adjusting their schedules based on the time of the day, origin airport congestion, and destination airport congestion.

Improving block-time estimates and planning for uncertainty in the schedule building process becomes necessary not just to improve OTP rankings but also to improve passenger service levels. The goal of this paper is to develop a robust optimization approach to schedule planning by specifically incorporating passenger centric goals in the planning models. A key trade-off in such a process is between higher service levels achieved through increasing (and better allocation of) flight block-times and higher planned costs (i.e., lower planned profits). In this paper we develop a model that re-times (perturbs) a proposed flight schedule by considering block-time probability distributions. First, we explicitly define notions of passenger and network service levels. Then we develop a model that maximizes the expected profit while guaranteeing minimum service levels. This models allows us to impose a set of minimum service level guarantees. While the optimization model is complex, we also develop computational procedures, based on cut generation techniques, to efficiently solve these models. To this end, this paper also has a methodological contribution to the development of computationally efficient procedures. We validate our model using data from a large US carrier and demonstrate potentially large revenue gains for the airline. However, in this extended abstract we do not include our computational results. We restrict our discussion to describing the main optimization model and providing a brief discussion on issues related to the computational tractability.

2. Model Description

Perturbing a flight schedule implies adjusting the scheduled departure times of flights in the network within an allowable time window. While perturbing the incumbent schedule, however, we must guarantee that the resulting schedule continues to remain feasible with respect to the *aircraft turns* built of the incumbent schedule. Every flight in the incumbent schedule is assigned to exactly one aircraft. An *aircraft turn* is essentially a pair of consecutive flights flown by the same aircraft. We assume that the set of turns associated with the incumbent schedule is known a priori.

A passenger travel plan, or *itinerary*, may comprise of multiple flight legs. Broadly defined, a *fare class* is the price an airline charges to book a passenger in a particular booking class. Airline seats are divided into several *booking classes*. Next, we define the important modeling notation and parameters.

N : The set of all flights (legs) in the airlines flight network,
B : the total available planned budget (depends on the total block-time across all flights)
O : the set of all passenger itineraries,
T : the complete set of aircraft turns,
F : the set of all fare classes,
α_i : the origin station of flight i,
β_i : the destination station of flight i,
m_{ij} : minimum passenger connection time between two flights i and j,
t_{ij} : minimum turn-time between flights i and j on an aircraft rotation
D_{of} : expected demand for itinerary o and fare class f,
$[l_i, u_i]$: the allowable departure time-window for flight i,
c_i : the per time unit cost incurred for flight i, which includes unit costs corresponding, to crew pay and aircraft utilization,
$\mathcal{B}_{if}$: booking limit for fare class f on flight i,
r_{of} : the average fare of itinerary o and fare class f,
d_i^s : the previously scheduled departure time of flight i in the incumbent schedule,
e_i : the penalty for deviating from the preferred departure time of flight i, and
δ : DoT OTP measure for flight delay (typically 15 minutes after the scheduled arrival time).

Next, we define the decision variables of the model.

d_i : The published departure time of flight i,
a_i : the published arrival time of flight i,
X_{of} : demand of itinerary o and fare-class f satisfied, and
z_{ij} : binary variable indicating if passenger connections between flights i and j is feasible.

The only random variables in the model are the block-times and are denoted by Y_{it} where t represents the departure time of the flight i. The relation between a flights departure time, arrival time, and the corresponding block-time is as follows: $A_i = d_i + Y_{id_i}$, where A_i is the random actual arrival time of flight i. The probability density function of a flight's block-time is represented by $p_i(\cdot, t)$ and is assumed to depend on the flight's departure time t. The cumulative density function is assumed to have a finite support $[\delta_l^i, \delta_u^i]$. To reduce the complexity of our computational experiments we assume the following: The expected demand for an itinerary, D_{of}, does not vary significantly for reasonable deviations in departure time.

A flight j is said to follow-on flight i if passengers of flight i can connect to flight j. The set of all passenger connections for a flight i depends on the arrival time of flight i and departure times of possible connection flights. We define the connection set for flight i, with respect to the incumbent schedule, is defined as $C_i(d,a) = \{j \in N : d_j - a_i \geq m_{ij} \ \& \ \beta_i = \alpha_j\}$. Next, we define the Service Level, SL_i, of any flight $i \in N$. SL_i is the probability that passengers from flight i can connect to any follow-on flight included in the set $C_i(d,a)$, i.e., $SL_i = \Pr[A_i + m_{ij} \leq d_j] \quad \forall j \in C_i(d,a)$. Observe that it follows from the definition that $SL_i = \Pr[Y_{id_i} \leq \min_{j \in C_(d,a)} d_j - d_i - m_{ij}]$. The Network Service Level (NSL) is defined as the minimum service level across all the flights in the airlines network, i.e., $NSL = \min_i SL_i$. Finally, the Flight Service Level (FSL), also referred to as the OTP, is defined as the probability that a particular flight is not delayed based on the DoT measure δ, i.e., $FSL_i = \Pr[Y_{i,d_i} \leq a_i - d_i + \delta]$. Lastly, for notational convenience we denote the fact that flight j follows flight i in an itinerary $o \in O$ by $i \to j$. Next, we describe the optimization model in Section 2.1.

2.1. Maximizing Operational Profits

We consider the case when an airline must maintain a minimum FSL, γ_f, over all flights in the network and simultaneously guarantee a minimum NSL of γ_n. The profit maximizing model (PMM) reads:

$$\text{(PMM)} \max \sum_{o,f} r_{of} \ X_{of} - \sum_{i \in N} e_i \, |d_i - d_i^s| - \sum_{i \in N} c_i(a_i - d_i) \tag{1}$$

$$\text{s.t. } \Pr\left[Y_{id_i} \leq d_j - d_i - m_{ij}\right] \geq \gamma_n \quad i \in N \quad j \in C_i(d,a) \tag{2}$$

$$\Pr[Y_{i,d_i} \leq a_i - d_i + \delta] \geq \gamma_f \qquad i \in N \tag{3}$$

$$\sum_{i \in N} c_i(a_i - d_i) \leq B \tag{4}$$

$$X_{of} \leq D_{of} \qquad o \in O, \ f \in F \tag{5}$$

$$\sum_{o \in O, i \in o} X_{of} \leq \mathcal{B}_{if} \qquad i \in N, \ f \in F \tag{6}$$

$$\sum_{o,f,j\in o,i\to j} X_{of} \leq \bar{K}_{ij} z_{ij} \qquad i\in N,\ j\in N, \tag{7}$$

$$d_j - a_i \geq m_{ij} z_{ij} - K(1-z_{ij}) \qquad i\in N,\ j\in N, \tag{8}$$

$$d_j - a_i - t_{ij} \geq 0 \qquad (i,j)\in T, \tag{9}$$

$$l_i \leq d_i \leq u_i \qquad i\in N \tag{10}$$

$$z_{ij}\in\{0,1\}, d, a \text{ unrestricted} \tag{11}$$

The first term in the objective function (1) corresponds to the net revenue due to satisfied itinerary demand, the second term is the net penalty due to deviation from preferred departure time, and the third term represents the total operational cost. Constraint (2) ensures that the minimum NSL is at least as large as the desired value of γ_n. It is not difficult to observe that $NSL \geq \gamma_n$ if and only if $SL_i \geq \gamma_n$ for every $i\in N$. The latter condition holds only if the constraint (2) is satisfied. Constraint (3) guarantees that the minimal FSL is at least γ_f. Constraint (4) restricts the total network operating cost incurred and constraint (5) restricts the fare-class itinerary demand to the maximum allowable. Since every flight i within an itinerary o can carry $\mathcal{B}_{if}$ of a particular fare-class f, constraint (6) ensures that the booking limit constraint on each flight is satisfied. Constraint (7) ensures that we only account for those itineraries whose flight sequence is legal with respect to the minimum connection time for passengers. The constant $\bar{K}_{ij} = \sum_f \mathcal{B}_{if} + \sum_f \mathcal{B}_{jf}$. Constraint (9) guarantees that the schedule is not perturbed to disrupt the pre-determined aircraft turns and hence the aircraft routing solution always remains feasible. The constant K is the length of the time horizon, i.e., $K = \max_{i\in N} u_i - \min_{i\in N} l_i$. Constraints (10) and (11) bound the variables appropriately.

3. Computational Tractability

In the model PMM constraints (2) and (3) are non-linear. This makes the model difficult to solve computationally. However, a technical assumption allows us to simplify the model and reduce its computational complexity. We assume that the block-time distributions are log-concave and stationary with respect to departure time. Consequently, constraints (2) and (3) transform as follows.

$$\log\left(\Pr[Y_i \leq d_j - d_i - m_{ij}]\right) \geq \log\gamma_n \quad i\in N, \quad j\in C_i(d,a) \tag{12}$$

$$\log\left(\Pr[Y_i \leq a_i - d_i + \delta]\right) \geq \log\gamma_f \qquad i\in N. \tag{13}$$

Given that we assume the block-time distribution is independent of the departure time we drop the departure time subscript, i.e., $Y_{id_i} = Y_i$. Further, recollect that the distribution has a finite support $[\delta_l^i, \delta_u^i]$. Let us denote this interval for flight i as $\mathcal{K}_i$. Now, for every flight i, we define a function $g_i(x)$ as $g_i(x) \equiv \log \Pr[Y_i \leq x], \quad x \in \mathcal{K}_i$. To build an outer linear approximation to $g_i(x)$, we consider the set of all linear functions, U_{ik}, defined over interval $\mathcal{K}_i$. We show the form of these linear functions later. For now, using these linear functions we rewrite $g_i(x)$ as follows:

$$g_i(x) = \min_{k \in \mathcal{K}_i} U_{ik}(x). \tag{14}$$

Using equation (14) we now rewrite the NSL constraint as follows: $g_i(d_j - d_i - m_{ij}) \geq \log \gamma_n \quad j \in C_i(d,a)$. This can be further simplified as: $z_{ij} g_i(d_j - d_i - m_{ij}) \geq \log \gamma_n \quad j \in \bar{C}_i$, where $\bar{C}_i$ denotes the longest set of possible connections for flight i when the departure and arrival times are adjusted. For example, $\bar{C}_i$ could be assumed to be the set of all flights originating at station α_i, i.e., $\bar{C}_i = \{j \in N : \beta_j = \alpha_i \text{ and can connect to } i \text{ regardless of re-timing}\} = \left\{j \in N : \beta_j = \alpha_i,\ d_j^s + w_i - (d_i^s - \delta_l^i + \delta_u^i) \geq m_{ij}\right\}$. Thus, constraint (12) is equivalent to

$$z_{ij} \min_{k \in \mathcal{K}_i} U_{ik}(d_j - d_i - m_{ij}) \geq \log \gamma_n \quad j \in \bar{C}_i \ i \in N. \tag{15}$$

It is noteworthy that in equation (15) if $d_j - d_i - m_{ij} \leq 0$, then $z_{ij} = 0$ and hence we need not worry about negative arguments, i.e., we restrict our attention to positive values only. We now characterize the functions $U_{ik}(x)$. Given the probability density function $p_i(\cdot)$, for block-time Y_i, we can write these functions as follows: $U_{ik}(x) = \frac{p_i(k)}{\int_0^k p_i(t)\ dt}(x-k) + \log \int_0^k p_i(t)\ dt.$

We still need to linearize constraint (15). To this end, notice that $U_{ik}(\delta_l^i) < 0$ and $U_{ik}(\delta_l^i) \leq U_{ik}(x)$ for all $x \geq \delta_l^i$. We now define additional continuous decision variables, $s_{ijk}^{(n)}$, for all $i \in N$, $j \in \bar{C}_i$, and $k \in \mathcal{K}_i$. Constraint (15) can then be replaced by the following set of linear constraints:

$$\begin{aligned} &s_{ijk}^{(n)} \geq \log \gamma_n \quad i \in N\ , j \in \bar{C}_i,\ k \in \mathcal{K}_i, \qquad z_{ij} U_{ik}(\delta_l^i) \leq s_{ijk}^{(n)} \leq 0 \quad i \in N,\ j \in \bar{C}_i,\ k \in \mathcal{K}_i, \text{ and} \\ &s_{ijk}^{(n)} \leq U_{ik}(d_j - d_i - m_{ij}) \quad i \in N,\ j \in \bar{C}_i,\ k \in \mathcal{K}_i. \end{aligned} \tag{16}$$

Similarly, the FSL constraint given by equation (13), can also be linearized as follows: $\min_{k \in \mathcal{K}_i} U_{ik}(a_i - d_i + \delta) \geq \log \gamma_f \quad i \in N$, where the function $U_{ik}(x)$ has the same form as described before.

References

Associated Press, The. 2007. Airlines' on-time marks in '07 near record low. URL `http://www.msnbc.msn.com/id/22490808/`.

Deshpande, V., M. Arikan. 2007. Are airlines newsvendors? or, an empirical estimation of the impact of airline flight schedules on flight delays. *Working paper.* Purdue University, IN.

Mayer, C., T. Sinai. 2003a. Network effects, congestion externalities, and air traffic delays: or why not all delays are evil. *The American Economic Review* **93** 1194–1214.

Mayer, C., T. Sinai. 2003b. Why do airlines systematically schedule their flight to arrive late? *Working paper.* University of Pennsylvania, PA.