A proposal for network air traffic flow management incorporating fairness and airline collaboration*

SHUBHAM GUPTA†

Faculty Advisor:
DIMITRIS BERTSIMAS‡

April 11, 2010

Abstract

There has been significant research effort in the academic literature related to Air Traffic Flow Management (ATFM). Yet, the research has not been fully implemented in practice. In our opinion, the key reasons are - i) the existing network models approach the problem from the point of view of a central decision-maker without taking into account the airlines to a significant degree; and ii) the notions of fairness as introduced under the Collaborative Decision-Making (CDM) paradigm operate under a single-airport setting and do not address network effects (presence of multiple airports, sectors and various connectivity requirements). In this paper, we address these shortfalls by presenting a proposal which alleviates both these concerns. In stage I of our proposal, we present network models that incorporate a notion of fairness - controlling number of reversals and total amount of overtaking. In stage II, we allow for further airline collaboration by proposing a network model for slot reallocation. We provide empirical results of the proposed optimization models on national-scale, real world datasets spanning across six days that show interesting tradeoffs between fairness and efficiency. We report promising computational times of less than 30 minutes for up to 25 airports and provide theoretical evidence that illuminates the strength of our formulations.

1 Introduction

The sustained growth of the aviation industry has put a tremendous strain on the available resources of the air transporation system. This is evidenced by the steady increase in flight delays and severe congestion at airports. In 2008, approximately 22% of the flights in the

*Research funded by the NSF Grant EFRI-0735905.

†Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, shubhamg@mit.edu

‡Boeing Professor of Operations Research, Sloan School of Management, Co-director of the Operations Research Center, Massachusetts Institute of Technology, E40-147, Cambridge, MA 02139-4307, dbertsim@mit.edu

United States were delayed by more than 15 minutes, while another 2% were cancelled (Bureau of Transportation Statistics [17]). Moreover, during the 12-month period ending in September 2008, 138 million minutes of system delay led to an estimated $10 billion in costs for US airlines [1].

Air Traffic Flow Management (ATFM) refers to the set of strategic processes that try to reduce congestion costs and support the goal of safe, efficient and expeditious aircraft movement. ATFM procedures try to resolve local demand-capacity mismatches by adjusting the aggregate traffic flows to match scarce capacity resources. Ground Delay Programs (GDPs) are one of the most sophisticated ATFM initiatives currently in use that attempt to address airport arrival capacity reductions. Under this mechanism, delays are applied to flights at their origin airports that are bound for a common destination airport which is suffering from reduced capacity or excessive demand. The premise for this tool is that it is better to absorb delays for a flight while it is grounded at its origin airport rather than incurring air-borne delay near the affected destination airport which is both unsafe and more costly (in terms of fuel costs). Another recently introduced tool similar to a GDP is an Airspace-Flow Program (AFP) which is used to control arrival rate into a weather affected segment of the airspace. Some of the other ATFM tools include assigning air-borne delays, dynamic re-routing and speed control. We briefly review the literature on the existing ATFM tools below.

Odoni [12] first conceptualized the problem of scheduling flights in real time in order to minimize congestion costs. Thereafter, several models have been proposed to handle different versions of the problem. The problem of assigning ground-delays in the context of a single-airport (*Single-Airport Ground-Holding Problem*) has been studied in Terrab and Odoni [15], Richetta and Odoni [13], [14]; and in the multiple airport setting (*Multi-Airport Ground-Holding Problem*) in Terrab and Paulose [16], Vranas et al. [20]. The problem of controlling release times and speed adjustments of aircraft while air-borne for a network of airports taking into account the capacitated airspace (*Air Traffic Flow Management Problem*) has been studied in Bertsimas and Stock [5], Helme [9], Lindsay et al. [11]. The problem with the added complication of dynamically re-routing aircrafts (*Air Traffic Flow Management Rerouting Problem*) was first studied by Bertsimas and Stock [6]. Recently, Bertsimas et al. [7] have presented a new mathematical model for the ATFM problem with dynamic re-routing which has superior computational performance. For a detailed survey of the various contributions and a taxonomy of all the problems, see Bertsimas and Odoni [4] and Hoffman et al. [10]. Despite the significant progress outlined here, this set of literature while addressing *network effects* (presence of multiple airports, sectors and various connectivity requirements), takes the point of view of a single decision-maker and does not address the issue of fairness among airlines.

We review next some other important concepts in the ATFM literature - namely, *Collaborative Decision-Making* (CDM) and *Ration-by-Schedule* (RBS). The decision-making responsibilities in ATFM initiatives are shared between a number of stakeholders (primarily, airlines and the FAA). This poses a major challenge as their actions are highly interdependent and demand real-time exchange of information between the FAA and the airlines. This realization of enhanced cooperation between the various stakeholders led to the adoption of Collaborative Decision-Making (CDM) philosophy (Ball et al. [2], Wambsganss [21]) by the FAA in the 1990s. Under CDM, all ATFM initiatives are conducted in a way that gives significant decision-making responsibilities to airspace users (see Hoffman et al. [10] for details on CDM). All recent efforts to improve ATFM have been guided by this philosophy. In the US, "Ration-by-Schedule" (RBS) is the fundamental principle for GDPs and all the CDM initiatives. Under this paradigm - arrival slots at airports are assigned to flights in accordance with a first-scheduled, first-served (FSFS) priority discipline (see Ball et al. [2], Wambsganss [21] for details on rationing). In the case of GDP planning, all stakeholders have agreed that this principle is fair to all parties. This allocation process is followed by a *Compression algorithm*, which fills open slots created by flights that are canceled. The compression procedure gives airlines an incentive to report accurate flight information, by rewarding them for reporting cancellations. The combined process, RBS plus Compression (formally called RBS^{++}) is the policy currently in use for slot allocations during GDPs. Despite the use of RBS in a GDP setting, there have been no network models that satisfy the RBS principle in a multi-airport setting. This is because, applying RBS to each of the airports individually might not lead to a schedule that preserves time, sector and flight connectivities. In addition, the imposition of a maximum permissible delay on each flight would mean that a feasible solution under RBS might not even exist if the capacity reduction at some airports is significant. Hence, there is no straightforward extension of RBS from a single-airport setting to an airspace context. As part of the CDM philosophy, researchers have also explored dynamic interaction with airlines. Towards this aim, Vossen, Ball [19] [18] have studied opportunities for slot trading in a single-airport setting where the aim is to formalize an optimization problem for the FAA given the offers to trade from various airlines. In summary, the CDM concepts discussed here while addressing the issues of fairness among airlines in a single-airport setting, do not address network effects.

Our goal in this paper is to propose an optimization based approach that a) incorporates network effects and builds upon the ATFM literature; and b) takes into account fairness considerations among airlines by building upon the CDM philosophy. Specifically, our proposal consists of the following two stages:

Stage I - Network ATFM model incorporating fairness:
We generalize the classical ATFM models ([5]) to incorporate fairness considerations for airlines. The objective function used in the existing network ATFM models is to minimize the total delay costs across all flights, i.e., the focus is on overall system efficiency. A disadvantage of such an approach is that the solution to such models can have a large number of reversals, i.e., the resulting order of flight arrivals can be quite different as compared to the published flight schedules. Moreover, across these reversals, there might be different number of time-periods of overtaking. Hence, the total overtaking across these reversals might be large. Because of this deviation from the original flight ordering, it becomes difficult to implement

such a solution because of the coupling in the crew assignments and the use of hub and spoke networks. We propose integer programming models that add these fairness controls. The key output in this stage is the assignment of flights to different time periods.

Stage II - Slot reallocation through airline collaboration:
We generalize the notion of Compression, a key component of the current CDM practice in a single-airport setting to network-wide slot reallocation among airlines. Specifically, we propose an optimization model which takes as input the assignment of flights to different time periods from Stage I and permits the airlines to trade these assigned slots across different airports, thereby, resulting in improved internal objective functions. The model proposed for Stage II of our proposal allow airlines to react to the schedule determined in Stage I by taking into account their flights in the entire network and making appropriate tradeoffs.

Contributions of this work.

We feel our work makes the following contributions:

1. We present a proposal for the network ATFM problem which incorporates both fairness and airline collaboration while operating under a broad CDM paradigm. Specifically, we formulate integer programming models to impose fairness controls in Stage I and a model for slot reallocation in Stage II.

2. We provide empirical results of the proposed optimization models on national-scale, real world datasets spanning across six days. We report promising computational times of less than 30 minutes for up to 25 airports and provide theoretical evidence on the strength of our formulations.

Simultaneously, Barnhart et al. [3] develop an alternative way to address fairness in the context of ATFM. They develop a fairness metric that measures deviation from FSFS and propose an integer programming formulation that directly minimizes this metric. They further develop an exponential penalty approach, and report encouraging computational results using simulated regional and national scenarios. In contrast to this work, our paper differs in the following respects: a) we provide an exact method to model overtaking within a mathematical programming framework while [3] proposes an approximate method. In addition, we propose another, although related, metric of fairness - controlling the number of reversals; b) we explicitly model network effects that [3] does not; and c) our proposal consists of a slot reallocation phase (Stage II) which enables airlines to further improve their internal objective functions. In summary, both papers contribute to the understanding of fairness in ATFM by approaching the problem from distinct perspectives.

Organization of the paper.

Section 2 summarizes an adaptation of the Bertsimas-Stock model [5] in order to accomodate our proposal. Section 3 introduces models of fairness. Section 4 introduces our model of slot reallocation. Section 5 reports computational results of the proposed optimization models on six days of national-scale, real world datasets. Section 6 summarizes our conclusions and the Appendix reports polyhedral analysis that illuminate the strength of our proposed formulations.

2 ATFM Problem : Notation, Bertsimas-Stock Model and Solutions

In this section, we reproduce the Bertsimas-Stock model [5] for the ATFM problem which provides the starting point for all the models presented herein. We use an extended version of the notation used in that paper in order to accomodate fairness and slot reallocation considerations. Finally, we illustrate difficulties relative to fairness considerations in the solutions obtained from this model.

The Decision Variables.

The decision variables are:

$$w_{j,t}^f = \begin{cases} 1, & \text{if flight } f \text{ arrives at sector } j \text{ by time } t, \\ 0, & \text{otherwise.} \end{cases}$$

This definition of the decision variables, using "*by*" instead of "*at*", is critical to the understanding of the formulation. The variables are defined only for the set of sectors an aircraft may fly through on its route to the destination airports. In addition, variables are used for the departure and the arrival airports, in order to determine the optimal times for departure and for arrival. Since we do not consider flight cancellations, at least two variables can be fixed a priori for each flight: each aircraft has to take off by the end of a feasible time window and has to land, as well, within a feasible time window, which is determined by the time of departure.

Notation.

The model's formulation requires definition of the following notation:

$\mathcal{K}$: set of airports,

$\mathcal{F}$: set of flights,

$\mathcal{T}$: set of time periods,

$\mathcal{W}$: set of airlines,

$\mathcal{F}_w \subseteq \mathcal{F}$: set of flights belonging to airline w,

$\mathcal{S}$: set of sectors,

$\mathcal{S}^f \subseteq \mathcal{S}$: set of sectors that can be flown by flight f,

$\mathcal{C}$: set of pairs of flights that are continued,

$\mathcal{R}$: set of pairs of flights that are reversible (definition below),

$\mathcal{P}_i^f$: set of sector i's preceding sectors,

$\mathcal{L}_i^f$: set of sector i's subsequent sectors,

$D_k(t)$: departure capacity of airport k at time t,

$A_k(t)$: arrival capacity of airport k at time t,

$S_j(t)$: capacity of sector j at time t,

d_f : scheduled departure time of flight f,

a_f : scheduled arrival time of flight f,

s_f : turnaround time of an airplane after flight f,

$orig_f$: airport of departure of flight f,

$dest_f$: airport of arrival of flight f,

l_{fj} : minimum number of time units that flight f must spend in sector j,

D : maximum permissible delay for a flight,

T_j^f : set of feasible time periods for flight f to arrive in sector j,

$\underline{T}_j^f$: first time period in the set T_j^f,

$\overline{T}_j^f$: last time period in the set T_j^f,

$T_{f,f'}^{reversal}$: set of time-periods common for flights f and f' where a reversal could occur (definition below),

$O_{f,f'}^{max}$: maximum amount of overtaking possible between flights f and f' (definition below),

$\mathcal{O}$: set of all possible trades (definition in Section 4),

$\mathcal{O}^f \subseteq \mathcal{O}$: set of offers containing flight f,

D_f : time period assigned to flight f from Stage I optimization.

The key additions relative to the notation used in [5] are $\mathcal{W}$ (set of airlines), $\mathcal{F}_w$ (set of flights belonging to airline w), $\mathcal{R}$ (set of pairs of flights that are reversible), $T_{f,f'}^{reversal}$ (set of time-periods common for flights f and f' where a reversal could occur), $O_{f,f'}^{max}$ (maximum amount of overtaking possible between flights f and f'), $\mathcal{O}$ (set of all possible trades), $\mathcal{O}^f \subseteq \mathcal{O}$ (set of offers containing flight f) and D_f (slot assigned to flight f from Stage I).

The set $\mathcal{R}$ (Reversible pairs of flights).

We give next the definition of $\mathcal{R}$ (set of pairs of flights that are reversible). A pair of flights (f, f') belongs to $\mathcal{R}$ if the following two conditions are satisfied:

1. $dest_f = dest_{f'}$, i.e., the destination airport of both flights f and f' is the same.
2. $a_f \leq a_{f'} \leq a_f + D$, i.e., the scheduled time of arrival of flight f' at the destination airport lies between the scheduled time of arrival of flight f and the last time period in the set of feasible time periods that the flight f can arrive at its destination airport.

For each pair of flights $(f, f') \in \mathcal{R}$, we count a *reversal*, if in the resulting solution, flight f' arrives before flight f (i.e., $\exists t$ such that $w_{dest_{f'},t}^{f'} > w_{dest_f,t}^{f}$).

The set $T_{f,f'}^{reversal}$ and parameter $O_{f,f'}^{max}$.

The set $T_{f,f'}^{reversal}$ (set of time-periods common for flights f and f' where a reversal could occur) is defined as $[\underline{T}_{dest_{f'}}^{f'}, \overline{T}_{dest_f}^{f} - 1]$. To elaborate, this is the set of time-periods t, such that it is possible to have the following assignment: $w_{dest_f,t}^{f} = 0$ and $w_{dest_{f'},t}^{f'} = 1$. This assignment would imply that a reversal occurs at time t.

The parameter $O_{f,f'}^{max}$ is defined as $|T_{f,f'}^{reversal}|$ (cardinality of the set $T_{f,f'}^{reversal}$), and hence is equal to $\overline{T}_{dest_f}^{f} - \underline{T}_{dest_{f'}}^{f'} - 1$. It is the maximum amount of overtaking possible between flights f and f' and would be attained when $w_{dest_f,\overline{T}_{dest_f}^{f}-1}^{f} = 0$ and $w_{dest_{f'},\underline{T}_{dest_{f'}}^{f'}-1}^{f'} = 1$.

Figure 1 depicts a reversible pair of flights $(f, f') \in \mathcal{R}$. In this example, the arrows correspond to the set of time-periods common for both flights. Moreover, the set of time-periods marked by these arrows (except for the last one) constitue $T_{f,f'}^{reversal}$. This is because, the model enforces $w_{dest_f,a_f+D}^{f} = 1$ at the outset and hence, it is not possible to have a reversal at $a_f + D$. Finally, $O_{f,f'}^{max} = |T_{f,f'}^{reversal}| = 6$.

The Objective Function.

We use an adapted expression for total delay costs (which is a combination of the costs of air-borne delay (AH) and ground-holding delay (GH)) introduced recently in Bertsimas et al. [7].

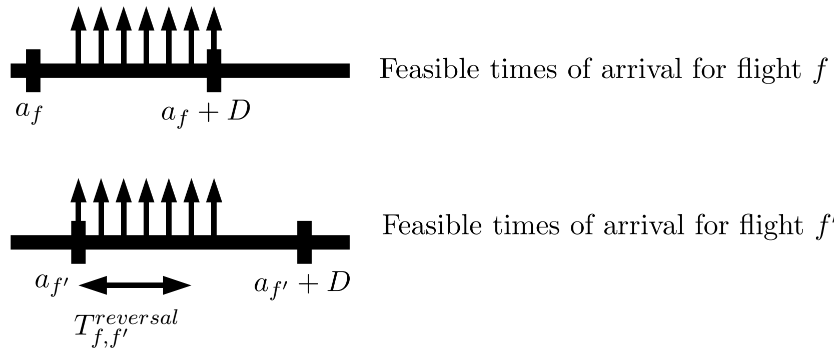

Figure 1: A reversible pair of flights $(f, f') \in \mathcal{R}$.

The objective function is composed of two terms: a first term that takes into account the cost of the total delay assigned to a flight and a second term which accounts for the cost reduction obtained when a part of the total delay is taken on the ground, before take-off. The objective function cost coefficients are a super-linear function of the tardiness of a flight of the form $(t - d_f)^{1+\epsilon_1}$ and $(t - a_f)^{1+\epsilon_2}$, with ϵ_1, ϵ_2 close to zero. Hence, for each flight f and for each time period t, we define the following two cost coefficients:

$$\begin{aligned} c_{td}^f(t) = (t - a_f)^{1+\epsilon_2} &\equiv \text{total cost of delaying flight } f \text{ for } (t - a_f) \text{ units of time,} \\ c_g^f(t) = (t - d_f)^{1+\epsilon_2} - (t - d_f)^{1+\epsilon_1} &\equiv \text{cost reduction obtained by holding flight } f \text{ on the ground} \\ &\quad \text{for } (t - d_f) \text{ units of time,} \end{aligned}$$

where a_f and d_f are the scheduled arrival and departure times of flight f, respectively. In view of the above, the objective function is as follows:

$$Min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T_{dest_f}^f} c_{td}^f(t) \cdot (w_{dest_f,t}^f - w_{dest_f,t-1}^f) - \sum_{t \in T_{orig_f}^f} c_g^f(t) \cdot (w_{orig_f,t}^f - w_{orig_f,t-1}^f) \right)$$

The TFMP model.

The complete description of the model, referred to as (TFMP), is as follows:

$$IZ_{TFMP} = Min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T_{dest_f}^f} c_{td}^f(t) \cdot (w_{dest_f,t}^f - w_{dest_f,t-1}^f) - \sum_{t \in T_{orig_f}^f} c_g^f(t) \cdot (w_{orig_f,t}^f - w_{orig_f,t-1}^f) \right)$$

subject to:

$$\sum_{f \in \mathcal{F}: orig_f = k} (w^f_{k,t} - w^f_{k,t-1}) \leq D_k(t) \qquad \forall k \in \mathcal{K},\ \ t \in \mathcal{T}. \tag{1}$$

$$\sum_{f \in \mathcal{F}: dest_f = k} (w^f_{k,t} - w^f_{k,t-1}) \leq A_k(t) \qquad \forall k \in \mathcal{K},\ \ t \in \mathcal{T}. \tag{2}$$

$$\sum_{f \in \mathcal{F}: j \in \mathcal{S}_f, j' = \mathcal{L}^f_j} (w^f_{j,t} - w^f_{j',t}) \leq S_j(t) \qquad \forall j \in \mathcal{S},\ \ t \in \mathcal{T}. \tag{3}$$

$$w^f_{j,t} - w^f_{j',t-l_{fj'}} \leq 0 \qquad \forall f \in \mathcal{F},\ \ t \in T^f_j,\ \ j \in \mathcal{S}^f\ :j \neq orig_f, j' = \mathcal{P}^f_j. \tag{4}$$

$$w^f_{orig_f,t} - w^{f'}_{dest_{f'},t-s_f} \leq 0 \qquad \forall (f, f') \in \mathcal{C},\ \ \forall t \in T^f_k. \tag{5}$$

$$w^f_{j,t-1} - w^f_{j,t} \leq 0 \qquad \forall f \in \mathcal{F},\ \ j \in \mathcal{S}^f,\ \ t \in T^f_j. \tag{6}$$

$$w^f_{j,t} \in \{0,1\} \qquad \forall f \in \mathcal{F},\ \ j \in \mathcal{S}^f,\ \ t \in T^f_j. \tag{7}$$

The first three sets of constraints take into account the capacities of the various elements of the system. Constraints (1) ensure that the number of flights which may take off from airport k at time t, will not exceed the departure capacity of airport k at time t. Likewise, Constraints (2) ensure that the number of flights which may arrive at airport k at time t, will not exceed the arrival capacity of airport k at time t. Finally, Constraints (3) ensure that the total number of flights which may feasibly be in Sector j at time t will not exceed the capacity of Sector j at time t.

The next three sets of constraints capture the various connectivities - namely sector, flight and time connectivity. Constraints (4) stipulate that a flight cannot arrive at Sector j by time t if it has not arrived at the preceding sector by time $t - l_{fj'}$. In other words, a flight cannot enter the next sector on its path until it has spent at least $l_{fj'}$ time units (the minimum possible) traveling through one of the preceding sectors on its current path. Constraints (5) represent connectivity between flights. They handle the cases in which a flight is continued, i.e., the flight's aircraft is scheduled to perform a subsequent flight within some user-specified time interval. The first flight in such cases is denoted as f' and the subsequent flight as f, while s_f is the minimum amount of time needed to prepare flight f for departure, following the landing of flight f'. Constraints (6) ensure connectivity in time. Thus, if a flight has arrived at element j by time $\tilde{t}$, then $w^f_{j,t}$ has to have a value of 1 for all later time periods $(t \geq \tilde{t})$.

Solutions from (TFMP).

Here, we illustrate the difficulties relative to fairness considerations in the solutions obtained from the formulation (TFMP). We report a solution from (TFMP) for one of the six datasets (14 July 2004) on which we have performed our experiments in this paper. First, the total number of reversals in the resulting solution is **915**. Moreover, there are **1492** units of overtaking across these reversals. This indicates that the sequence of flight arrivals in the solutions from (TFMP) differ significantly from the scheduled sequence of flight arrivals. These two observations in the solutions obtained from the formulation (TFMP) is present across all the six datasets. In the next section, we introduce discrete optimization models that control the number of reversals and the amount of overtaking.

3 Network models that incorporate concepts of fairness

3.1 Minimizing the total amount of overtaking

A notion of fairness widely agreed upon by the airlines is to have a schedule that preserves the order of flight arrivals at an airport according to the published schedules in the Online Airline Guide (OAG). As previously mentioned, this is known as Ration-by-Schedule (RBS). But, given the capacity reductions at the airports, it might not always be possible to have a feasible solution under RBS in a network setting. A close equivalent to the RBS solution would be one which has a small amount of overtaking. Hence, in such a scenario, a plan which minimizes the total overtaking while keeping the total delay cost small might be a more desirable solution. As the first approach, we present a model which achieves this objective.

For every reversible pair of flights $(f, f') \in \mathcal{R}$, let $g_{f,f'}$ denote the total amount of overtaking between flights f and f'. Then, we need to define the following set of variables to express $g_{f,f'}$.

$$s^i_{f,f'} = \begin{cases} 1, & \text{if flight } f' \text{ arrives but } f \text{ does not arrive by time } \underline{T}^{f'}_{dest_{f'}} + i, \\ 0, & \text{otherwise.} \end{cases}$$

The definition above implies the following:

$$s^i_{f,f'} = 1 \iff \left\{ w^f_{dest_f, \underline{T}^{f'}_{dest_{f'}}+i} = 0, \; w^{f'}_{dest_{f'}, \underline{T}^{f'}_{dest_{f'}}+i} = 1 \right\}$$

Table 1 summarizes the various feasible combinations of these variables under the above definition. Thus, an alternative way to express $s^i_{f,f'}$ is as follows:

$$s^i_{f,f'} = \max \left\{ w^{f'}_{dest_{f'}, \underline{T}^{f'}_{dest_{f'}}+i} - w^f_{dest_f, \underline{T}^{f'}_{dest_{f'}}+i}, 0 \right\}. \tag{8}$$

Equation (8) implies that the following constraints suffice to express $s^i_{f,f'}$ in a mathematical programming framework if the objective is to minimize $s^i_{f,f'}$:

$$s^i_{f,f'} \geq w^{f'}_{dest_{f'}, \underline{T}^{f'}_{dest_{f'}}+i} - w^f_{dest_f, \underline{T}^{f'}_{dest_{f'}}+i}, \tag{9}$$

$$s^i_{f,f'} \geq 0. \tag{10}$$

S.No.	$w^{f}_{dest_f,\underline{T}^{f'}_{dest_{f'}}+i}$	$w^{f'}_{dest_{f'},\underline{T}^{f'}_{dest_{f'}}+i}$	$s^{i}_{f,f'}$
1	0	0	0
2	0	1	1
3	1	0	0
4	1	1	0

Table 1: Truth table for modeling the overtaking variables.

The set of variables $s^{i}_{f,f'}$ are defined for $i \in [0,\ O^{max}_{f,f'}]$. Now, $g_{f,f'} \in [0,\ O^{max}_{f,f'}]$ can be defined as follows:

$$g_{f,f'} = \sum_{i=0}^{O^{max}_{f,f'}} s^{i}_{f,f'}.$$

We shall work with an alternative description of Equation (9) to make the exposition on overtaking clearer. We substitute $i = t - \underline{T}^{f'}_{dest_{f'}}$ in Equation (9) to rewrite it as follows:

$$w^{f'}_{dest_{f'},t} \leq w^{f}_{dest_f,t} + s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'}. \tag{11}$$

We prove next that the following set of constraints are required to model overtaking between $(f, f') \in \mathcal{R}$ if we use an objective function to minimize $g_{f,f'}$ in addition with $s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \geq 0,\ \ \forall t \in T^{reversal}_{f,f'}$:

$$w^{f'}_{dest_{f'},t} \leq w^{f}_{dest_f,t} + s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \qquad \forall t \in T^{reversal}_{f,f'}. \tag{12}$$

Proposition 1. *If we use an objective function of minimizing $g_{f,f'}$ (the total amount of overtaking for $(f, f') \in \mathcal{R}$) in addition with $s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \geq 0,\ \ \forall t \in T^{reversal}_{f,f'}$, then Constraint (12) correctly captures the semantics of overtaking.*

Proof. In case, there is no reversal, i.e.,

$$w^{f'}_{dest_{f'},t} \leq w^{f}_{dest_f,t},\ \ \forall t \in T^{reversal}_{f,f'},$$

then Constraint (12) becomes reduntant. Since we minimize total amount of overtaking (and $s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \geq 0,\ \ \forall t \in T^{reversal}_{f,f'}$), it forces:

$$s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} = 0,\ \ \forall t \in T^{reversal}_{f,f'},$$

leading to $g_{f,f'} = 0$. On the contrary, if there are i units of overtaking, then $\exists t \in [\underline{T}^{f'}_{dest_{f'}}, \underline{T}^{f'}_{dest_{f'}} + O^{max}_{f,f'} - i]$ such that:

$$w^{f'}_{dest_{f'},t-1} = 0, \;\; w^{f'}_{dest_{f'},t} = 1,$$
$$w^{f}_{dest_{f},t+i-1} = 0, \;\; w^{f}_{dest_{f},t+i} = 1.$$

Now, the time-connectivity constraints (Constraints (6)) imply that:

$$w^{f'}_{dest_{f'},t+m} = 1, \;\; w^{f}_{dest_{f},t+m} = 0, \; \forall\, 0 \le m \le i-1.$$

Constraint (12) then enforces

$$s^{k}_{f,f'} = 1, \;\; \forall\, t \le k \le t+i-1.$$

Again, since we minimize total amount of overtaking (and $s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \ge 0, \;\; \forall t \in T^{reversal}_{f,f'}$), therefore,

$$s^{k}_{f,f'} = 0, \;\; \forall\, 0 \le k < t, \;\; t+i-1 < k \le O^{max}_{f,f'},$$

leading to $g_{f,f'} = i$. In summary, Constraint (12) (in addition with $s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \ge 0$), correctly model overtaking if the objective function is to minimize $g_{f,f'}$. □

The proof of Proposition 1 relied critically on the assumption that we use an objective function that minimizes $g_{f,f'}$. Next, we propose a formulation to model overtaking which is independent of the objective function used. We propose a set of constraints that capture the convex hull of the four feasible integer points enumerated in Table 1, namely $(0,0,0), (0,1,1), (1,0,0)$ and $(1,1,0)$. Figure 2 depicts the convex hull of these four points. We introduce the following set of constraints to model overtaking:

$$w^{f'}_{dest_{f'},t} \le w^{f}_{dest_{f},t} + s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \qquad \forall (f,f') \in \mathcal{R}, \;\; t \in T^{reversal}_{f,f'}. \tag{13}$$

$$w^{f}_{dest_{f},t} \le w^{f'}_{dest_{f'},t} + 1 - s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \qquad \forall (f,f') \in \mathcal{R}, \;\; t \in T^{reversal}_{f,f'}. \tag{14}$$

$$w^{f}_{dest_{f},t} + s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \le 1 \qquad \forall (f,f') \in \mathcal{R}, \;\; t \in T^{reversal}_{f,f'}. \tag{15}$$

$$-w^{f'}_{dest_{f'},t} + s^{t-\underline{T}^{f'}_{dest_{f'}}}_{f,f'} \le 0 \qquad \forall (f,f') \in \mathcal{R}, \;\; t \in T^{reversal}_{f,f'}. \tag{16}$$

The TFMP model with the additional control on total amount of overtaking (referred to as TFMP-Overtake henceforth) is as follows:
$IZ_{TFMP-Overtake} =$

$$Min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T^{f}_{dest_f}} c^{f}_{td}(t) \cdot (w^{f}_{dest_f,t} - w^{f}_{dest_f,t-1}) - \sum_{t \in T^{f}_{orig_f}} c^{f}_{g}(t) \cdot (w^{f}_{orig_f,t} - w^{f}_{orig_f,t-1}) \right) +$$

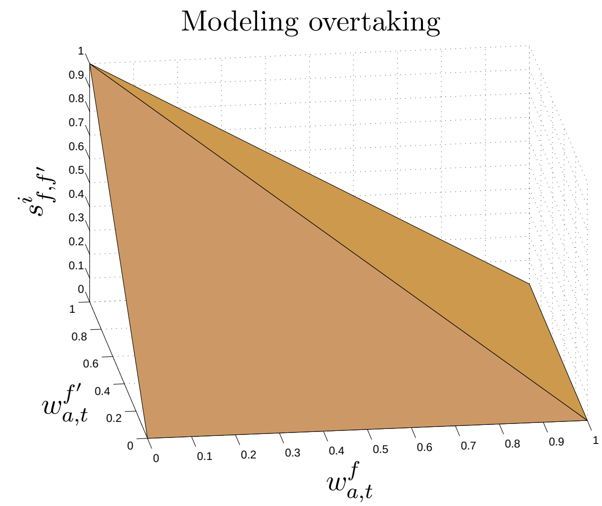

Figure 2: Convex hull of the integer points in Table 1 to model overtaking, ($dest_f = dest_{f'} = a$, and $i = t - \underline{T}_a^{f'}$).

$$\lambda_1 \cdot \left(\sum_{(f,f')\in\mathcal{R}} \sum_{i=0}^{O^{max}_{f,f'}} s^i_{f,f'} \right)$$

subject to:

$$(1) - (7).$$
$$(13) - (16).$$
$$s^i_{f,f'} \in \{0,1\} \qquad \forall (f,f') \in \mathcal{R}, \;\; i \in [0, O^{max}_{f,f'}].$$

3.2 Minimizing the total number of reversals

The model introduced in Section 3.1 took into account the magnitude of overtaking within each reversal. In this section, we introduce a model which controls the total number of reversals.

For each element $(f, f') \in \mathcal{R}$, we introduce the following new variable:

$$s_{f,f'} = \begin{cases} 1, & \text{if there is a reversal,} \\ 0, & \text{otherwise.} \end{cases}$$

Next, we relate the variables $s_{f,f'}$ (used to model a reversal) to the variables $s^i_{f,f'}$ (used to model overtaking). It is evident that a reversal occurs if and only if there is at least one time-period of overtaking. Mathematically, it translates to the following:

$$s_{f,f'} = 1 \iff \left\{ \exists i \in [0, O^{max}_{f,f'}], \;\; s^i_{f,f'} = 1 \right\} \tag{17}$$

Building upon Equation (17), we have the following:

$$s_{f,f'} = \max_{t \in T^{reversal}_{f,f'}} \left\{ s_{f,f'}^{t-\underline{T}^{f'}_{dest_{f'}}} \right\},$$

$$s_{f,f'} = \max_{t \in T^{reversal}_{f,f'}} \left\{ \max\{w^{f'}_{dest_{f'},t} - w^{f}_{dest_f,t}, 0\} \right\},$$

$$s_{f,f'} = \max \left\{ \max_{t \in T^{reversal}_{f,f'}} \{w^{f'}_{dest_{f'},t} - w^{f}_{dest_f,t}\}, 0 \right\}. \tag{18}$$

Equation (18) implies that the following constraints suffice to express $s_{f,f'}$ in a mathematical programming framework if the objective is to minimize $s_{f,f'}$:

$$s_{f,f'} \geq w^{f'}_{dest_{f'},t} - w^{f}_{dest_f,t}, \ \ \forall t \in T^{reversal}_{f,f'}. \tag{19}$$

$$s_{f,f'} \geq 0. \tag{20}$$

Equation (19) can be rearranged as follows:

$$w^{f'}_{dest_{f'},t} \leq w^{f}_{dest_f,t} + s_{f,f'}, \ \ \forall t \in T^{reversal}_{f,f'}. \tag{21}$$

Proposition 2. *If we use an objective function of minimizing $s_{f,f'}$, then Constraint (21) in addition with $s_{f,f'} \geq 0$ correctly captures the semantics of modeling a reversal.*

Proof. In case, there is no reversal, i.e.,

$$w^{f'}_{dest_{f'},t} \leq w^{f}_{dest_f,t} \ \ \forall t \in T^{reversal}_{f,f'},$$

then Constraint (21) becomes reduntant. Since we minimize $s_{f,f'}$ (and $s_{f,f'} \geq 0$), it forces $s_{f,f'} = 0$. On the contrary, if there is a reversal, then $\exists t \in T^{reversal}_{f,f'}$ such that:

$$w^{f'}_{dest_{f'},t} = 1, \ \ w^{f}_{dest_f,t} = 0.$$

Constraint (21) then implies that $s_{f,f'} \geq 1$. Again, minimizing $s_{f,f'}$ makes $s_{f,f'} = 1$ ensuring that Constraint (21) indeed models a reversal correctly. □

The proof of Proposition 2 relied critically on the assumption that we use an objective function that minimizes $s_{f,f'}$. Here, we present a formulation that models a reversal correctly independently of the objective function used. For each element $(f, f') \in \mathcal{R}$, we introduce the following constraints to (TFMP):

$$w^{f'}_{dest_{f'},t} \leq w^{f}_{dest_f,t} + s_{f,f'} \qquad \forall (f,f') \in \mathcal{R}, \ \ t \in T^{reversal}_{f,f'}. \tag{22}$$

$$w^{f}_{dest_f,t} \leq w^{f'}_{dest_{f'},t} + 1 - s_{f,f'} \qquad \forall (f,f') \in \mathcal{R}, \ \ t \in T^{reversal}_{f,f'}. \tag{23}$$

If there is a reversal between flights f and f', i.e., $s_{f,f'} = 1$, then Constraint (22) becomes redundant and Constraint (23) stipulates that if flight f has arrived by time t, then flight f' has to arrive by that time, hence ensuring that flight f cannot arrive before flight f'. Similarly, if there is no reversal, i.e., $s_{f,f'} = 0$, then Constraint (23) becomes redundant and Constraint (22) stipulates that if flight f' has arrived by time t, then flight f has to arrive by that time, hence ensuring that flight f' cannot arrive before flight f. Thus, we are able to model a reversal with the addition of only one variable ($s_{f,f'}$).

Given this additional set of constraints, the model then minimizes a weighted combination of total delay costs and total number of reversals. The parameter λ_2 is chosen appropriately to control the tradeoff between the two objectives.

(TFMP) extended with reversals: (TFMP-Reversal).

The TFMP model with the additional control on reversals is as follows:
$IZ_{TFMP-Reversal} =$

$$Min \sum_{f \in \mathcal{F}} \left(\sum_{t \in T^f_{dest_f}} c^f_{td}(t) \cdot (w^f_{dest_f,t} - w^f_{dest_f,t-1}) - \sum_{t \in T^f_{orig_f}} c^f_g(t) \cdot (w^f_{orig_f,t} - w^f_{orig_f,t-1}) \right) +$$

$$\lambda_2 \cdot \left(\sum_{(f,f') \in \mathcal{R}} s_{f,f'} \right)$$

subject to:

$$(1) - (7).$$
$$(22) - (23).$$
$$s_{f,f'} \in \{0,1\} \qquad \forall (f,f') \in \mathcal{R}.$$

For each element $(f,f') \in \mathcal{R}$, let $IP_{Reversal}$ denote the set of all feasible binary vectors satisfying Constraints (22) and (23). We show in the Appendix that the polyhedron induced by Constraints (22) and (23) is the convex hull of solutions in $IP_{Reversal}$.

$$\begin{aligned} IP_{Reversal} = \{ w^f_{dest_f,t} \in \{0,1\}, \;\; s_{f,f'} \in \{0,1\} | \\ \mathrm{w}^{f'}_{dest_{f'},t} \leq w^f_{dest_f,t} + s_{f,f'} \qquad t \in T^{reversal}_{f,f'}, \\ \mathrm{w}^f_{dest_f,t} \leq w^{f'}_{dest_{f'},t} + 1 - s_{f,f'} \qquad t \in T^{reversal}_{f,f'}. \} \end{aligned}$$

RBS Policy - a special case of (TFMP-Reversal).

When there is sufficient capacity at all airports, such that a feasible solution under RBS exists (i.e., there are no reversals), this model is capable of generating that solution (using a sufficiently high λ_2) while minimizing the total delay costs. *Hence, a solution under RBS policy is a special case of our model.* Since, a solution under RBS preserves the order of flight arrivals, therefore, for every pair of flights $(f, f') \in \mathcal{R}$, the variable $s_{f,f'} = 0$, and Constraints (22) and (23) reduce to Constraint (24) which ensures that flight f' cannot arrive before flight f:

$$w^{f'}_{dest_{f'},t} \leq w^{f}_{dest_f,t} \qquad \forall (f, f') \in \mathcal{R}, \;\; t \in T^{reversal}_{f,f'}. \tag{24}$$

Extensions.

Here, we elaborate on how the proposed models (TFMP-Reversal and TFMP-Overtake) can be extended to control reversals and overtaking in en-route airspace and how they can accomodate alternative objective functions.

- *Control reversals/overtaking in en-route airspace:*
  Suppose, in addition to the arrival airports, we want to control the number of reversals and amount of overtaking in sector j. This can be achieved by introducing variables $s^i_{f,f',j}$ (for each $(f, f') \in \mathcal{R}$) with the following semantics:

  $$s^i_{f,f',j} = \begin{cases} 1, & \text{if flight } f' \text{ arrives but } f \text{ does not arrive by time } \underline{T}^{f'}_j + i \text{ at sector } j, \\ 0, & \text{otherwise.} \end{cases}$$

  Then, the total amount of overtaking is given by: $g_{f,f',j} = \sum_{i=0}^{O^{max}_{f,f',j}} s^i_{f,f',j}$ and can be modeled similar to the way done in Constraints (13) - (16). Further, for a reversal at an intermediate sector j, we introduce the following variable:

  $$s_{f,f',j} = \begin{cases} 1, & \text{if there is a reversal between flights } f \text{ and } f' \text{ at sector } j, \\ 0, & \text{otherwise.} \end{cases}$$

  Finally, the reversal at sector j can be modeled similar to the way done in Constraints (22) - (23).

- *Incorporating alternative objective functions:*
  Although the models presented in this paper minimize the number of reversals and amount of overtaking, it is possible to extend them to accomodate alternative objective functions. For instance, suppose we want to equalize the resulting reversals and overtaking among airlines taking into account the number of flights they operate. This can be achieved as follows:
  Let d_w denote the number of reversals per flight for airline w and γ denote the mean of the d_w's across all airlines.

  $$d_w = \left(\sum_{f' \in \mathcal{F}_w} s_{f,f'} \right) / |\mathcal{F}_w|,$$

$$\gamma = \left(\sum_{w\in\mathcal{W}} d_w\right) \Big/ |\mathcal{W}|.$$

Then, we add $|d_w - \gamma|$ term to the objective function of minimizing the total delay cost with an appropriate tradeoff parameter.

Size of the Formulations.
Denoting with

$$D = \max_{f\in\mathcal{F}, j\in P_f} |T_f^j|, \qquad N = \max_{f\in\mathcal{F}} |\mathcal{S}^f|,$$

the total number of decision variables and constraints for the various models can be bounded as listed in Table 2.

Model	No. of Decision Variables	No. of Constraints
(TFMP)	$\lvert\mathcal{F}\rvert DN$	$2\lvert\mathcal{K}\rvert\lvert\mathcal{T}\rvert + \lvert\mathcal{S}\rvert\lvert\mathcal{T}\rvert + 2\lvert\mathcal{F}\rvert DN + 2\lvert\mathcal{F}\rvert N + D\lvert\mathcal{C}\rvert$
(TFMP-Reversal)	$\lvert\mathcal{F}\rvert DN + \lvert\mathcal{R}\rvert$	$2\lvert\mathcal{K}\rvert\lvert\mathcal{T}\rvert + \lvert\mathcal{S}\rvert\lvert\mathcal{T}\rvert + 2\lvert\mathcal{F}\rvert DN + 2\lvert\mathcal{F}\rvert N + D\lvert\mathcal{C}\rvert + 2\mathcal{R}D$
(TFMP-Overtake)	$\lvert\mathcal{F}\rvert DN + D\lvert\mathcal{R}\rvert$	$2\lvert\mathcal{K}\rvert\lvert\mathcal{T}\rvert + \lvert\mathcal{S}\rvert\lvert\mathcal{T}\rvert + 2\lvert\mathcal{F}\rvert DN + 2\lvert\mathcal{F}\rvert N + D\lvert\mathcal{C}\rvert + \mathcal{R}D$

Table 2: Upper bound on the size of the models.

In order to get a feeling of the size of the formulations, let us consider an example that adequately represents the U.S. network: $\mathcal{K} = 50$, $\mathcal{T} = 100$, $\mathcal{S} = 100$, $\mathcal{R} = 50000$, $\mathcal{F} = 10000$, $\mathcal{C} = 8000$, $D = 6$ and $N = 5$. For this example, the upper bound on the number of variables and constraints are listed in Table 3.

Model	No. of Decision Variables	No. of Constraints
(TFMP)	300,000	780,000
(TFMP-Reversal)	350,000	1,380,000
(TFMP-Overtake)	600,000	1,080,000

Table 3: Numerical Example: Upper bound on the size of the models.

Since we introduce only one class of variables $s_{f,f'}$ for all elements $(f, f') \in \mathcal{R}$, the number of variables in the model (TFMP-Reversal) are comparable to the original model (TFMP).

4 Network model for slot reallocation that incorporates airline collaboration

In this section, we present a model (called TFMP-Trading) for slot reallocation in a network setting that introduces only one additional variable per offer above the TFMP model. We let airlines submit offers to trade slots assigned to its flights across different airports. The executed set of trades should ensure that the resulting schedule is still feasible taking into account all kinds of network connectivities and airspace capacities.

The set $\mathcal{O}$ (Set of Airline Offers).

We give next the definition of $\mathcal{O}$ (set of airline offers). We use a structure proposed by Vossen, Ball [19] that allows the airlines to submit so-called "at-least, at-most" offers. Airlines submit offers of the following kind: $(f_d, t_{d'}; f_u, t_{u'})$ which means that the airline is willing to move flight f_d to a later time-period, but no later than $t_{d'}$; in return for moving flight f_u to an earlier time-period, but no later than $t_{u'}$. The destination airports of flights f_d and f_u are allowed to be distinct. The set $\mathcal{O}$ contains all such four-tuples $(f_d, t_{d'}; f_u, t_{u'})$ submitted by the airlines after a schedule is generated from Stage I of our proposal. Note that D_{f_d} and D_{f_u} denote the slots allotted to the two flights from Stage I, and hence, for such an offer to be useful, we must have $D_{f_d} < t_{d'}$ and $D_{f_u} > t_{u'}$. Finally, $\mathcal{O}^f \subseteq \mathcal{O}$ defines the set of offers containing flight f.

4.1 (TFMP-Trading): A model for slot reallocation in a network setting

Here, we introduce our model of slot reallocation in a network setting. The model only introduces one additional variable per offer above the variables used in the TFMP model of Bertsimas-Stock [5], namely, $w^f_{j,t}$.

The Decision Variables.

- $o_{dd'uu'} \in \{0, 1\} = 1$ if offer $(f_d, t_{d'}; f_u, t_{u'})$ is executed.

Constraints.

$$(1) - (7).$$

$$o_{dd'uu'} \leq w^{f_d}_{dest_{f_d}, t_{d'}} \qquad \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}. \tag{26}$$

$$o_{dd'uu'} \leq w^{f_u}_{dest_{f_u}, t_{u'}} \qquad \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}. \tag{27}$$

$$\sum_{j \in \mathcal{O}^f} o_j \leq 1 \qquad \forall f \in \mathcal{F}. \tag{28}$$

$$w^f_{dest_f, D_f} - w^f_{dest_f, D_f - 1} \geq 1 - (\sum_{j \in \mathcal{O}^f} o_j) \qquad \forall f \in \mathcal{F}. \tag{29}$$

$$w^f_{j,t} \in \{0, 1\} \qquad \forall f \in \mathcal{F},\ \ j \in \mathcal{S}^f,\ \ t \in T^f_j. \tag{30}$$

$$o_{dd'uu'} \in \{0, 1\} \qquad \forall (f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}. \tag{31}$$

Constraints (26) and (27) enforce that when an offer $o_{dd'uu'}$ is executed (i.e., $o_{dd'uu'} = 1$), then $w^{f_d}_{dest_{f_d}, t_{d'}} = 1$ and $w^{f_u}_{dest_{f_u}, t_{u'}} = 1$, i.e., flights f_d and f_u cannot arrive after the respective time-periods in the offer, namely, t'_d and t'_u. This ensures that the semantics of the structure of an offer are satisfied. Constraint (28) enforces that for each flight, at most one offer can get executed. Moreover, constraint (29) stipulates that if no offer for a flight f is executed (i.e., $o_j = 0,\ \forall j \in \mathcal{O}^f$), then the flight will arrive at the time-period allotted from Stage I (D_f).

Objective Function.

In the model presented above, we have not explicitly stated the objective function that should be used. It is evident that fairness in the number of executed offers across airlines would again be relevant in this stage of our proposal.

Let n_w denote the number of trades executed corresponding to airline w, and let γ denote the mean of the trades executed across all airlines.

$$n_w = \sum_{f \in \mathcal{F}_w, j \in \mathcal{O}^f} o_j,$$

$$\gamma = \left(\sum_{w \in \mathcal{W}} n_w \right) \Big/ |\mathcal{W}|.$$

In the next section, we report computational results based on the following two objective functions:

- *Objective 1*: maximize the total number of trades ($\max \sum_{(f_d, t_{d'}; f_u, t_{u'}) \in \mathcal{O}} o_{dd'uu'}$).
- *Objective 2*: minimize the difference in the number of trades executed for each airline from the mean ($\min \sum_{w \in \mathcal{W}} |n_w - \gamma|$).

5 Computational Results

In this section, we report computational results from the optimization models introduced in the previous section on national-scale, real world datasets spanning across six days. The dataset for each day encompasses 55 major airports of the US and covers operations of the top five airlines. Each dataset contains data on the actual flight arrival and departure times for that particular day which lets us compute the actual delays.

5.1 Statistics of the Datasets

Table 4 summarizes the statistics of six days of flights data. These correspond to the operations at the 55 major airports of the US. We filter in the flights corresponding to the operations of the top 5 airlines (measured by the number of flight operations) - Southwest (SWA), American (AAL), Delta (DAL), United (UAL) and Northwest (NWA) to enable us to better analyze the distribution of delays across airlines.

Day	Number of Flights	Number of Connecting Flights	Total Delay (units of 15 min.)
14th Jul'04	5092	2691	4438
4th Aug'04	5844	3298	4926
13th May'05	5780	3310	3079
16th Jul'05	4590	2301	3907
27th Jul'05	5128	2728	3326
27th Jul'06	4781	2504	3101

Table 4: Summary of the datasets.

5.2 Experimental Setup

In our experimental setup, the airspace is subdivided into sectors of equal dimensions (10 by 10) that form a grid, thereby, having a total of 100 sectors. Each of the 55 major airports of the US is then mapped to one of these 100 sectors based on its geographical coordinates. For each flight, we fix its flight trajectory (i.e., the sequence of sectors in its path) based on the shortest path from the origin to the destination airport. Using the information on its flight time, we compute the minimum amount of time that each flight has to spend in a sector. This value is then used to calculate the set of feasible times that a flight can be in a sector. By tracking the tail number of an aircraft, we form the set of connecting flights.

For a sample day, we know the scheduled departure and arrival times of a flight as well as what actually happened on that day. We use this to compute its ground and air-hold delays. Further, we compute the capacities at all the airports by noting the actual times of departure and arrival of the flights and we use these values as capacity inputs to run the optimization models. It is important to note that this capacity corresponds to the exact number of flight departures and arrivals that happened on that particular day and hence, is the most conservative estimates of the capacity. The available airport capacities on that day has to be higher than the actual number of operations. Finally, we set values for the nominal sector capacities that lead to no delays when these airport capacities are used.

A critical parameter of the optimization models is the maximum permissible delay for a flight (D). This value is used to define the set of feasible times that a flight can be in a particular sector. For example, the set of feasible times that a flight f can arrive at its destination airport is given by all values in a_f through $(a_f + D)$. The size of all the optimization models and hence, the computational times, are sensitive to the value of D. We use a value of $D = 6$, which corresponds to 6 time periods (each of length 15 minutes), hence permitting a maximum delay of 90 minutes.

To compute optimal solutions, we use the CPLEX-MIP solver 11.0, implemented using AMPL as a modeling language on a laptop with 2 GB RAM and Linux Ubuntu OS. The instances reported in this paper have a typical size of the order of 300,000 variables (this increases significantly for TFMP-Overtake) and 800,000 constraints (this increases significantly for TFMP-Reversal).

5.3 Performance of (TFMP)

Table 5 reports solutions from the (TFMP) model for two cases - first, when the capacity used is the exact number of flight arrivals and departures as happened on that particular day, and second, with the capacity increased by 20%. In the first case, the total absolute delay is not very different from the actual delays of the day under consideration. The reason is that since the total capacity in this case is exactly equal to the total number of aircraft operations of that day, hence the (TFMP) model just finds a feasible solution and there is not much scope for optimization. In the second case, where we give the optimization model some room by increasing the capacity by 20%, there is an average reduction of 23% in the total absolute delays. This illustrates the benefits that could be achieved by using a centralized optimization-based approach. Table 6 lists the number of reversals and amount of overtaking in the solutions obtained from (TFMP) when the capacity is increased by 20%. The number of reversals consistently range between 500 and 1000 and amount of overtaking range between 800 and 1500 across all days. This confirms that, although, there can be significant benefits in the total delay costs by using the model (TFMP), the number of reversals and overtaking might be high.

Day	No. of Flights	Actual Delay (units of 15 min.)	TFMP Delay (units of 15 min.)	
			(same capacity)	(cap. increased by 20%)
14 Jul'04	5092	4438	4360	3385
4 Aug'04	5844	4926	4863	3492
13 May'05	5780	3079	3034	2242
16 Jul'05	4590	3907	3851	3053
27 Jul'05	5128	3326	3282	2648
27 Jul'06	4781	3101	3050	2542

Table 5: Performance of (TFMP).

5.4 Performance of (TFMP-Overtake)

Table 7 reports the computational performance of (TFMP-Overtake) model on the six datasets. These results pertain to the parameter λ_1 set to 100. The number reported under

Day	Number of Flights	Number of Reversals	Amount of Overtaking
14 Jul'04	5092	915	1492
4 Aug'04	5844	924	1426
13 May'05	5780	753	1191
16 Jul'05	4590	769	1235
27 Jul'05	5128	801	1291
27 Jul'06	4781	526	822

Table 6: Number of reversals and amount of overtaking from (TFMP).

'Total overtaking' takes into account the relative magnitudes of overtaking within each reversal, i.e., the number of time periods by which a flight overtakes its preceding flight when a reversal occurs. The degradation in total delay costs from (TFMP-Overtake) model over the (TFMP) solution range between 13% and 41% for fairness at 25 airports, the average being 24.5%. The model on average takes less than 30 minutes to converge to optimality for up to 25 airports.

5.5 Performance of (TFMP-Reversal)

The (TFMP-Reversal) model minimizes a weighted combination of total delay costs and total number of reversals where λ_2 is the weight parameter. We study the tradeoff inherent in these conflicting objectives in two ways - a) as a function of the tradeoff parameter λ_2 and b) as a function of the number of airports where this fairness criterion is imposed.

The effect of the tradeoff paramater.

Figure 3 plots the tradeoff in the number of reversals with the total delay cost as a function of λ_2 for fairness based on controlling total reversals imposed at 25 airports. The five points on the plot for each day correspond to the result from (TFMP-Reversal) with $\lambda_2 = 0$, 1, 10, 100 and 1000. Initially, there is a significant reduction in the number of reversals at the cost of a small increase in total delay cost, but the subsequent benefits in the number of reversals come at a high cost. For all days, the model is able to achieve less than 100 reversals for a degradation of at most 10% in the total delay cost. To achieve reversals between 10 and 30, the degradation in total delay costs range between 10% and 40% across all days.

Day (# of Flights)	No. of Airports with fairness	Solution Time (sec.)	No. of Reversals	Total overtaking	Total Delay Cost (units of 15 min.)	% Increase in Delay Cost over (TFMP)
	0	261	915	1492	3525	
14 Jul	5	186	2	4	3690	4.68
2004	15	727	13	27	4243	20.36
(5092)	25	3073	26	39	4662	32.25
	30	3600	39	65	4815	36.59
	0	108	924	1426	3604	
4 Aug	5	206	1	2	3802	5.49
2004	15	596	6	9	4029	11.79
(5844)	25	806	9	12	4080	13.20
	30	3530	16	18	4510	25.13
	0	311	753	1191	2313	
13 May	5	170	3	6	2401	3.80
2005	15	397	8	18	2584	11.71
(5780)	25	295	11	22	2651	14.61
	30	3394	17	28	3096	33.85
	0	51	769	1235	3173	
16 Jul	5	150	1	1	3628	14.33
2005	15	746	5	6	4201	32.39
(4590)	25	691	13	18	4452	40.30
	30	3600	29	33	4743	49.47
	0	178	801	1291	2744	
27 Jul	5	116	0	0	2871	4.62
2005	15	492	10	18	3319	20.95
(5128)	25	1983	17	26	3505	27.73
	30	3600	25	36	3804	38.62
	0	49	526	822	2637	
27 Jul	5	143	5	7	2826	7.16
2006	15	378	9	15	3070	16.42
(4781)	25	479	15	22	3145	19.26
	30	1305	28	49	3383	28.28

Table 7: Computational performance of (TFMP-Overtake) with capacity increased by 20 percent. Note that the row with k airports corresponds to imposing fairness at k airports and no fairness at the remaining $|\mathcal{K}| - k$ airports. In particular, $k = 0$ corresponds to the (TFMP) solution.

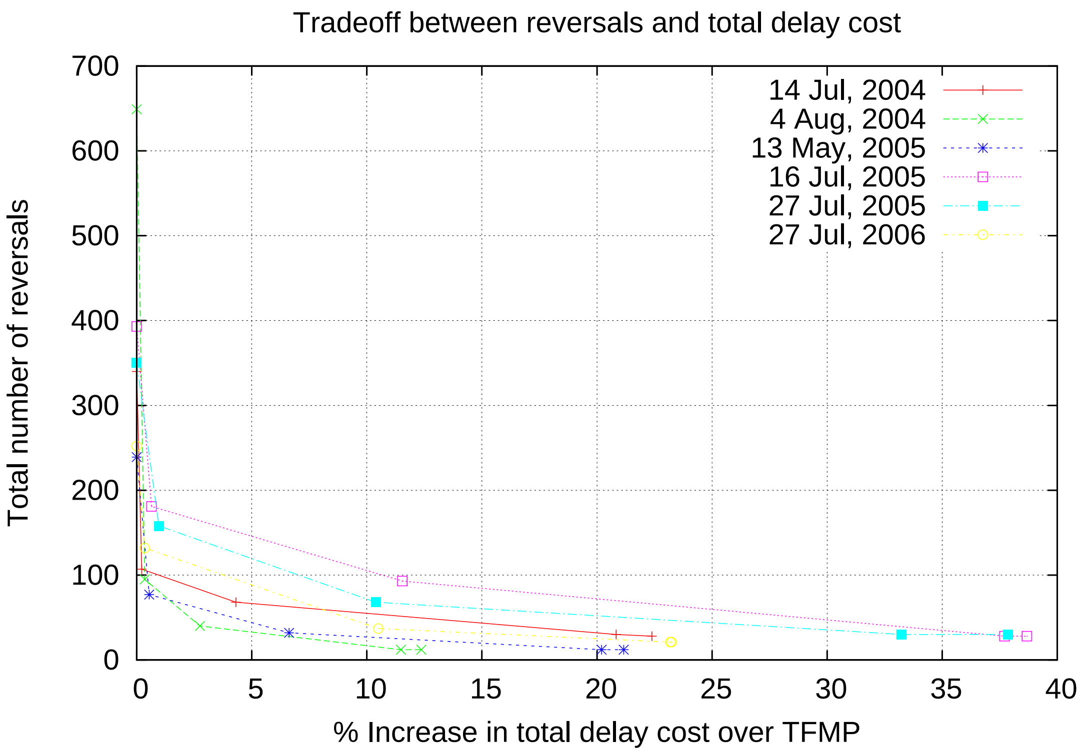

Figure 3: Effect of the tradeoff parameter λ_2. The five points for each day correspond to the result from (TFMP-Reversal) with $\lambda_2 = 0$, 1, 10, 100 and 1000.

The effect of the number of airports.

Table 8 reports the computational performance of the (TFMP-Reversal) model on the six datasets as a function of the number of airports where this fairness criterion is imposed. The capacity input used for the results in Table 8 is 20% higher than the exact number of aircraft operations that happened on the day under consideration. These results pertain to the tradeoff parameter λ_2 set to 100. As is evident from the results reported across all days, the number of reversals can be controlled up to 10-30. The degradation in total delay costs from (TFMP-Reversal) model over the (TFMP) solution range between 13% and 40% for fairness at 25 airports, the average being 24.5%. The model on average takes less than 30 minutes to converge to optimality for up to 25 airports. As expected, the total amount of reversals reported in Table 8 is always less than the corresponding number in Table 7, whereas the opposite is true for the amount of overtaking.

The computational times of both (TFMP-Reversal) and (TFMP-Overtake) are consistently less than 30 minutes for up to 25 airports, but they break down when we impose fairness at 30 airports and above. We believe that this is due to memory limitations.

5.6 Performance of slot reallocation model (TFMP-Trading)

Given the assignment of flights to various time-periods from Stage I, we generate offers to trade for each airline which maximizes its on-time performance. In other words, each airline tries to maximize the number of flights with delay less than one time-unit (15 minutes). To elaborate, suppose two flights f_1 and f_2 (belonging to the same airline) have been assigned two time-units of delay each from Stage I optimization. Moreover, let t_1 and t_2 be the time-periods assigned to the two flights respectively. Then, the owner airline generates an offer to trade which says that it is willing to delay flight f_1 further by three time-units if in return flight f_2 can arrive within one time-unit of delay, i.e., it generates the offer $(f_1, t_1{+}3; f_2, t_2{-}1)$. Thus, in case, this trade is executed, then flight f_2 will arrive on-time (given the definition of on-time performance). Table 9 reports the results from our network slot reallocation model (TFMP-Trading). As is evident, the computational times of the model TFMP-Trading is consistently less than a minute.

Comparison of TFMP-Trading between single-airport and network-wide settings.

In this section, we contrast the performance of TFMP-Trading between single-airport and network-wide settings. It is evident that in the network version, there is a tradeoff between the flexibility of trading slots at different airports versus the added constraint of satisfying all network connectivities. In contrast, in a single-airport setting, there is the advantage of not having to satisfy the network connectivity requirements at the expense of losing on trades across different airports. So, to compare the two settings, we divide the set of generated

Day (# of Flights)	No. of Airports with fairness	Solution Time (sec.)	No. of Reversals	Total overtaking	Total Delay Cost (units of 15 min.)	% Increase in Delay Cost over (TFMP)
	0	261	915	1492	3525	
14 Jul	5	228	2	4	3691	4.70
2004	15	899	13	30	4246	20.45
(5092)	25	3600	25	53	4402	24.87
	30	3600	37	84	4586	30.09
	0	108	924	1426	3604	
4 Aug	5	213	1	2	3785	5.02
2004	15	837	6	9	4028	11.76
(5844)	25	1024	9	12	4077	13.12
	30	3600	16	30	4403	22.16
	0	311	753	1191	2313	
13 May	5	219	3	6	2406	4.02
2005	15	282	8	19	2583	11.67
(5780)	25	384	11	24	2648	14.48
	30	3600	15	35	3048	31.77
	0	51	769	1235	3173	
16 Jul	5	165	1	1	3627	14.30
2005	15	1367	5	10	4138	30.41
(4590)	25	2292	12	22	4468	40.81
	30	3600	29	55	4558	43.64
	0	178	801	1291	2744	
27 Jul	5	225	0	0	2866	4.44
2005	15	654	10	24	3319	20.95
(5128)	25	1556	16	35	3504	27.69
	30	3600	25	38	3925	43.03
	0	49	526	822	2637	
27 Jul	5	213	5	7	2836	7.54
2006	15	751	9	16	3059	16.00
(4781)	25	654	15	25	3125	18.50
	30	3600	28	51	3387	28.44

Table 8: Computational performance of (TFMP-Reversal) with capacity increased by 20 percent. Note that the row with k airports corresponds to imposing fairness at k airports and no fairness at the remaining $|\mathcal{K}| - k$ airports. In particular, $k = 0$ corresponds to the (TFMP) solution.

offers between local and network offers. A local offer is one where both the flights involved have the same destination airport. On the contrary, a network offer contains flights whose destination airports are distinct. Hence, the single-airport model will not have any executed trades that correspond to the network offers. Table 10 reports the results from the two versions under the two objectives described before. The number reported under TFMP-Trading is the number of offers executed from the network model proposed in this paper, whereas, SA-Trading reports the results obtained from TFMP-Trading after removing all the network satisfiablity constraints and only taking into account the local offers. The numbers reported highlight the tradeoffs inherent in ignoring the network effects vis-a-vis trading slots at different airports. As the percentage of local offers increases, SA-Trading outperforms TFMP-Trading emphasizing that network connectivities are indeed relevant, whereas, for higher fraction of network offers, TFMP-Trading performs better reinforcing the utility of network offers.

Day	No. of Flights	TFMP-Trading			
		Objective 1		Objective 2	
		Offers Executed	Sol. Time (in sec.)	Offers Executed	Sol. Time (in sec.)
14 Jul'04	5092	283	13	190	29
4 Aug'04	5844	256	14	180	31
13 May'05	5780	150	5	140	19
16 Jul'05	4590	278	11	215	27
27 Jul'05	5128	194	9	140	22
27 Jul'06	4781	153	6	90	16

Table 9: Computational performance of TFMP-Trading.

5.7 Summary of computations

To summarize, TFMP (the optimization model without fairness) is able to reduce the total delays by 23% on average when we take reasonable estimates on the available capacity (increasing it by 20% from the actual number of flight operations). In Stage I of our proposal, we obtain solutions that are able to control the total reversals and overtaking up to less than 100 (from the models TFMP-Reversal and TFMP-Overtake respectively). In addition, we report promising computational times of less than 30 minutes for up to 25 airports from both models which make them well suited for online use. In Stage II of our proposal, there are 220 trades executed on average when the objective function used is to maximize the

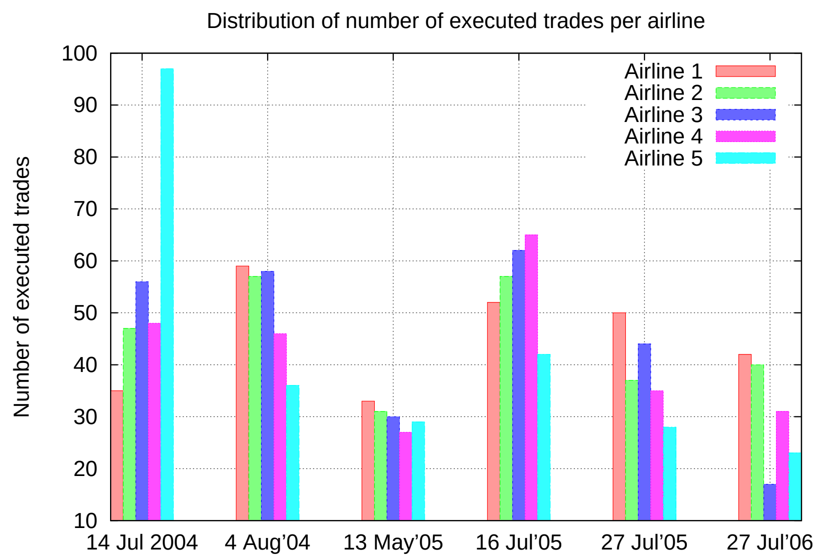

Figure 4: Distribution of number of executed trades across airlines from TFMP-Trading (*Objective 1*).

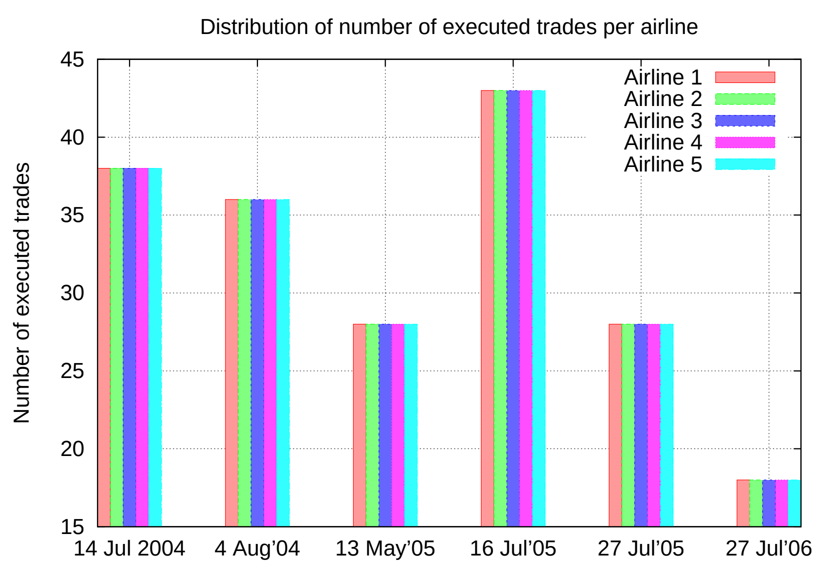

Figure 5: Distribution of number of executed trades across airlines from TFMP-Trading (*Objective 2*).

% Local Offers	% Network Offers	No. of Offers Executed (Obj. 1)		No. of Offers Executed (Obj. 2)	
		SA-Trading	TFMP-Trading	SA-Trading	TFMP-Trading
0	100	0	281	0	190
25	75	152	255	130	90
50	50	221	241	175	125
75	25	269	238	195	120
100	0	308	258	240	175

Table 10: Comparison of TFMP-Trading between single-airport and network-wide settings.

number of trades, and 160 when we impose fairness. This reinforces the utility of the slot reallocation phase of our proposal as the airlines are able to increase the number of flights arriving on-time.

6 Conclusions

In this paper, we present a proposal that extends the CDM framework from an airport setting to an airspace context while explicitly incorporating fairness and airline collaboration. Specifically, in Stage I of our proposal, we introduce models for the ATFM problem that incorporate notions of fairness. Given a schedule from this stage, in Stage II, we propose an optimization model for slot reallocation to mimic the compression and substitution-cancellation process in a GDP setting. Further, we report empirical results of the proposed models on national-scale, real world datasets. We feel the key advantages of our proposal are high-quality of solutions, consideration of network effects and promising computational times which ensure practical tractability in real-time settings.

Acknowledgement

This research has been funded by the NSF Grant EFRI-0735905. We thank Bill Moser and Mark Weber of Lincoln Labs for providing us with the data.

Appendix

Strength of TFMP-Reversal:
Let us denote the polyhedron induced by the the additional set of constraints to model a reversal for each element $(f, f') \in \mathcal{R}$ as $P_{Reversal}$.

Proposition 3. *The polyhedron $P_{Reversal}$ is integral.*

Proof. $P_{Reversal}$ can be written as follows:

$$\begin{aligned} P_{Reversal} = \{x = (w^f_{a,t},\ s_{f,f'})|\ & 0 \leq w^f_{a,t} \leq 1,\ \ 0 \leq s_{f,f'} \leq 1, \\ \mathrm{w}^{f'}_{dest_{f'},t} - w^f_{dest_f,t} - s_{f,f'} \leq 0 & \qquad t \in T^{reversal}_{f,f'}, \\ \mathrm{w}^{f}_{dest_f,t} - w^{f'}_{dest_{f'},t} + s_{f,f'} \leq 1 & \qquad t \in T^{reversal}_{f,f'}\} \end{aligned}$$

We make use of the following two facts from discrete optimization [8]:

Fact 1: Let $\mathbf{A}$ be an integral matrix. $\mathbf{A}$ is totally unimodular if and only if $\{\mathbf{x}|\ \mathbf{a} \leq \mathbf{A}\mathbf{x} \leq \mathbf{b}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}\}$ is integral, for all integral vectors $\mathbf{a},\ \mathbf{b},\ \mathbf{l},\ \mathbf{u}$.

Fact 2: A matrix $\mathbf{A}$ is totally unimodular if and only if each collection Q of rows of $\mathbf{A}$ can be partitioned into two parts so that the sum of the rows in one part minus the sum of the rows in the other part is a vector with entries only 0, +1 and -1.

Consider the following polyhedron P and let $\mathbf{A}$ be the matrix such that $P = \{\mathbf{x}|\ \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$:

$$\begin{aligned} P = \{x = (w^f_{a,t},\ s_{f,f'})| & \\ \mathrm{w}^{f'}_{dest_{f'},t} - w^f_{dest_f,t} - s_{f,f'} \leq 0 & \qquad t \in T^{reversal}_{f,f'}, \\ \mathrm{w}^{f}_{dest_f,t} - w^{f'}_{dest_{f'},t} + s_{f,f'} \leq 1 & \qquad t \in T^{reversal}_{f,f'}\} \end{aligned}$$

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \cdots & \cdots & 0 & -1 \\ -1 & 1 & 0 & 0 & 0 & \cdots & \cdots & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & \cdots & \cdots & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 & \cdots & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & -1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & -1 & 1 & 1 \end{bmatrix}$$

The matrix $\mathbf{A}$ has a special structure. If we remove the last column, the remaining matrix is a network matrix.

Let $B_1, B_2, ..., B_n$ be consecutive blocks of two rows each, i.e., block B_k contains the rows $2k-1$ and $2k$. For any collection Q of rows of the matrix $\mathbf{A}$, we show how to partition it into two parts J_1 and J_2 so that the sum of the rows in J_1 minus sum of the rows in J_2 is a vector with entries 0, +1 and -1 only. Suppose Q contains both the rows of some block B_i, then, put both these rows in J_1. The remaining rows (say m) in Q then come from different blocks, call them $R_{j_1}, R_{j_2}, ..., R_{j_m}$. These m rows are partitioned as follows:

Let Q_+ be the subset of these m rows where the last element is $+1$ and Q_- be those rows where the last element is -1. Then, put $\lceil \frac{|Q_+|}{2} \rceil$ rows of Q_+ in J_1 and the remaining rows in J_2. Similarly, put $\lceil \frac{|Q_-|}{2} \rceil$ rows of Q_- in J_1 and the remaining rows in J_2. Since the sum of two rows in the same block is all zeroes, therefore, all such blocks in J_1 do not affect the sum of all the rows in J_1. Let T denote the vector resulting from the sum of the rows in J_1 minus the sum of the rows in J_2. All the elements except the last one in T is exactly 0, +1 or -1 because of the structure of the matrix A. The contribution of the rows from Q_+ to the last element of T is either 0 or +1. Similarly, the contribution of the rows from Q_- to the last element of T is either -1 or 0. This implies that the last element of T which is the sum of these two contributions can either be +1, -1 or 0.

This shows that the matrix A is totally unimodular. Using Fact 1, we conclude that $P_{Reversal}$ is integral. □

References

[1] Air Transport Association of America Inc., 2008. *http://www.airlines.org/economics/cost+of+delays*

[2] Ball M., Chen C-Y., Hoffman R., and Vossen T. Colloborative decision making in air traffic management: Current and future research directions. In Proceedings of ATM'99, Workshop on Advanced Technologies and Their Impact on Air Traffic Management in the 21st Century. 1999.

[3] Barnhart C., Bertsimas D., Caramanis C., Fearing D. (2009). Equitable and Efficient Coordination of Traffic Flow Management Programs. Submitted to Transportation Science.

[4] Bertsimas D., Odoni A.R. (1997) A critical survey of optimization models for tactical and strategic aspects of air traffic flow management. Technical report, NASA.

[5] Bertsimas D., Stock S. (1998) The Air Traffic Management Problem with Enroute Capacities. Operations Research **46**, 406-422.

[6] Bertsimas D., Stock S. (2000). The Traffic Flow Management Rerouting Problem in Air Traffic Control: A Dynamic Network Flow Approach, Transportation Science **34**, 239-255.

[7] Bertsimas D., Lulli G., Odoni A.R. (2008). The Air Traffic Flow Management Problem: An Integer Optimization Approach. IPCO 2008.

[8] Bertsimas D., Weismantel R. (2005). Optimization Over Integers. Dynamic Ideas.

[9] Helme M. (1992). Reducing air traffic delay in a space-time network, IEEE International Conference on Systems, Man and Cybernetics **1**, 236-242.

[10] Hoffman R., Mukherjee A. and Vossen T. (2007). Air Traffic Flow Management. Working Paper.

[11] Lindsay K., Boyd E., Burlingame R. (1993). Traffic flow management modeling with the time assignment model. Air Traffic Control Quarterly **1**, (3), 255-276.

[12] Odoni A.R. (1987). The Flow Management Problem in Air Traffic Control, in *Flow Control of Congested Networks*, A.R. Odoni, L. Bianco and G. Szego eds., 269-288, Springer-Verlag, Berlin.

[13] Richetta O. and Odoni A.R. (1993). Solving Optimally the Static Ground-Holding Policy Problem in Air Traffic Control. Transportation Science **27**, 228-238.

[14] Richetta O. and Odoni A.R. (1994). Dynamic Solution to the Ground-Holding Policy Problem in Air Traffic Control. Trans. Res. **28a**, 167-185.

[15] Terrab M. and Odoni A.R. (1993). Strategic Flow Control on an Air Traffic Network. Operations Research **41**, 138-152.

[16] Terrab M. and Paulose S. (1993). Dynamic Strategic and Tactical Air Traffic Flow Control. RPI Technical Report.

[17] U.S. Dept. of Transportation, Bureau of Transportation Statistics (2008). Airline On-Time Statistics. *http://www.transtats.bts.gov/HomeDrillChart.asp*

[18] Vossen T., Ball M. Slot Trading Opportunities in Collaborative Ground Delay Programs. Transportation Science. 2006.

[19] Vossen T., Ball M. Optimization and Mediated Bartering Models for Ground Delay Programs. Naval Research Logistics. 2006.

[20] Vranas P.B., Bertsimas D., Odoni A.R. (1994). The Multi-Airport Ground Holding Problem in Air Traffic Control. Operations Research, **42**, 249–261.

[21] Wambsganss M. Collaborative decision making through dynamic information transfer. Air Traffic Control Quarterly, **4**:107-123, 1996.