Evaluation of an Auction Mechanism for Allocating Airport Arrival Slots
by
Eric J. Cholankeril
Massachusetts Institute of Technology
Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of Master of Engineering in Electrical Engineering and Computer Science at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2003
© Eric J. Cholankeril, MMIII. All rights reserved.
The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.
Evaluation of an Auction Mechanism for Allocating Airport Arrival Slots

by

Eric J. Cholankeril

Submitted to the Department of Electrical Engineering and Computer Science on May 21, 2003, in partial fulfillment of the requirements for the degree of Master of Engineering in Electrical Engineering and Computer Science

Abstract

A sequential, sealed-bid Vickrey auction without package bidding is proposed as a mechanism for allocating arrival slots during Ground Delay Programs (GDPs). The auction mechanism was simulated on historical flight data and compared with two other methods—the Collaborative Decision Making method currently in use, and a “global optimization” method that determines an upper bound on the delay reduction achievable. Additionally, an integer program was formulated to simulate how airlines might optimally reroute aircraft in response to a GDP, and to determine the total amount of passenger delay resulting from a particular allocation of slots. Overall, it was found that at least 5-7% of passenger delay was reduced through the auction method for the scenarios tested, and as much as 75% of the maximum possible delay reduction was achieved in some scenarios.

Thesis Supervisor: John-Paul Clarke
Title: Associate Professor of Aeronautics and Astronautics

Thesis Supervisor: William D. Hall
Title: Senior Member of Technical Staff, Draper Laboratory
Acknowledgments

First and foremost, I’d like to acknowledge Bill Hall, my advisor at Draper Laboratory, for helping me find a really interesting thesis topic, for offering me countless suggestions over the past year, and for being a really nice, approachable guy. Much of my research was inspired by his previous work on airport operations.

I’d also like to thank John-Paul Clarke, my faculty advisor, for all of his advice and suggestions, and for being so available despite having many other students. J.P. is truly a man of ideas, and I’m grateful for his guidance.

I’d also like to acknowledge Bruno Miller and Francis Carr from the MIT International Center for Air Transportation for some very useful conversion tables, and everyone else from ICAT who offered suggestions. Thanks also to Draper Laboratory and the Education Office for funding my research and offering me a really great opportunity.

Finally, I’d like to thank Joy, for keeping me around.
Acknowledgments (continued)

This thesis was prepared at The Charles Stark Draper Laboratory, Inc., under Internal Research and Development Contract IRD03-0-5043.

Publication of this thesis does not constitute approval by Draper Laboratory of the findings or conclusions contained herein. It is published for the exchange and stimulation of ideas.

Eric J. Cholankeril
May 21, 2003
[This page intentionally left blank.]
# Contents

1 Introduction  
1.1 Arrival Slots: A Scarce Resource  
1.2 The Need for Incentives in Slot Exchange  
1.3 Overview of the Thesis  

2 Background and Related Work  
2.1 Design of Auction Mechanisms  
2.1.1 Open vs. Sealed Bidding  
2.1.2 Sequential vs. Simultaneous Bidding  
2.1.3 Individual vs. Package Bidding  
2.1.4 The Vickrey Auction  
2.1.5 Application to the Arrival Slot Auction  
2.2 Previous Work on Auctioning Arrival Slots  
2.2.1 Combinatorial Auction (Rassenti)  
2.2.2 Multi-Object Auction (Milner)  
2.2.3 Groves Mechanism (Hall)  
2.3 The Airline Recovery Problem  
2.3.1 Set-Packing Model (Clarke)  
2.3.2 Set-Packing with Selection Heuristic (Rosenberger)  

3 Arrival Slot Allocation Methods  
3.1 Initial Slot Allocation via Ration-By-Schedule  
3.2 Collaborative Decision Making
3.2.1 Substitution and Cancellation .................................. 37
3.2.2 Compression ....................................................... 37
3.3 Global Optimization .................................................. 38
3.4 The Auction Mechanism ............................................... 39
  3.4.1 Determination of Bid Values .................................. 40
  3.4.2 Determination of Reservation Price ......................... 41
  3.4.3 Alternative Airline Strategies in the Auction ............... 42

4 Modeling the Airline Recovery Problem .......................... 45
  4.1 Overview of the Model ............................................ 46
  4.2 Generating Alternative Routes .................................... 46
    4.2.1 Restricting Route Possibilities Using Subroutes ........ 48
    4.2.2 Route Generation Algorithm ............................... 49
  4.3 Objective Function and Constraints ................................ 51
    4.3.1 Precomputing Route Delay Values for Slotted Routes .... 54
    4.3.2 Estimating Flight Cancellation Delays .................. 55
    4.3.3 Estimating Delays for Interrupted Connections .......... 58
  4.4 Modeling Considerations ......................................... 59

5 Data Reconstruction and Generation .............................. 61
  5.1 Reconstruction of Schedule Data ............................... 61
    5.1.1 ASQP Flight Reconstruction ............................... 62
    5.1.2 ASQP Route Reconstruction ................................ 62
    5.1.3 Anomalies in ASQP Data .................................... 63
    5.1.4 OAG Reconstruction ......................................... 64
  5.2 Stochastic Generation of Passenger Data ....................... 65
    5.2.1 Determination of Passenger Loads ....................... 65
    5.2.2 Creation of Passenger Connections ....................... 66

6 Results and Analysis ................................................ 69
  6.1 Implementation Notes ............................................ 69
6.2 Comparison of Slot Allocation Methods ......................................... 70
6.3 Sensitivity of Results to Delay Cutoff ........................................... 73
6.4 Analysis of Alternative Airline Behaviors ...................................... 74
   6.4.1 Effect of Increasing Airline Cautiousness in the Auction .............. 76
   6.4.2 Effect of Increasing Airline Cautiousness in CDM ...................... 79
   6.4.3 Effect of Predictive Bidding in the Auction ............................... 80
6.5 Running Time of the Optimization ............................................... 82

7 Conclusions ................................................................................. 85
   7.1 Future Research ................................................................. 86
   7.2 Summary of Thesis Contributions ........................................... 88
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-1</td>
<td>Example Original Routes for Two Aircraft</td>
<td>47</td>
</tr>
<tr>
<td>4-2</td>
<td>Airline Recovery Optimization Problem</td>
<td>52</td>
</tr>
<tr>
<td>5-1</td>
<td>Determining Candidate Flights for Connections from a Flight $f$</td>
<td>67</td>
</tr>
<tr>
<td>5-2</td>
<td>Generation of Passenger Itineraries</td>
<td>68</td>
</tr>
<tr>
<td>6-1</td>
<td>Effect of Increasing Maximum Individual Passenger Delay on Total RBS Airline Delay</td>
<td>75</td>
</tr>
<tr>
<td>6-2</td>
<td>Effect of Increasing One Airline's Caution Level on its Passenger Delay in the Auction</td>
<td>78</td>
</tr>
<tr>
<td>6-3</td>
<td>Effect of Increasing One Airline's Caution Level on its Net Passenger Delay in the Auction</td>
<td>78</td>
</tr>
<tr>
<td>6-4</td>
<td>Effect of Increasing Number of Cautious Airlines on Overall Passenger Delay in the Auction</td>
<td>79</td>
</tr>
<tr>
<td>6-5</td>
<td>Effect of Increasing Number of Cautious Airlines on Overall Passenger Delay in CDM</td>
<td>80</td>
</tr>
<tr>
<td>6-6</td>
<td>Effect of Increasing Number of Predictive Airlines on Overall Passenger Delay in the Auction</td>
<td>81</td>
</tr>
</tbody>
</table>
List of Tables

4.1 Decision Variables .................................................. 52
4.2 Optimization Parameters .............................................. 53

6.1 Percentage of Reducible Delay Captured by Auction, for Various GDPs 70
6.2 CDM Passenger Delay Reduced Using Auction, for Various GDPs .......... 72
6.3 Percentage of Reducible Delay Captured by Auction, for Various Delay Cutoffs .................................................. 75
6.4 CDM Passenger Delay Reduced Using Auction, for Various Delay Cutoffs 75
6.5 Airline Codes and Optimization Running Times .......................... 84
List of Algorithms

3.1 Ration By Schedule ................................................. 36
3.2 Compression .......................................................... 38
3.3 Auction Mechanism ................................................... 41
4.1 Construction of Alternative Routes .............................. 50
4.2 Generating Slotted Routes and Computing Delay Values .... 56
Chapter 1

Introduction

Flight delays continue to be a source of frustration for wearied travelers, even though in recent years, aircraft on-time statistics have actually improved. According to the Bureau of Transportation Statistics, over 80 percent of scheduled flights arrived on time in 2002—a higher on-time percentage rate than for each of the previous seven years [5]. (A flight is defined as arriving “on time” if it arrives at the gate within 15 minutes of its originally scheduled arrival time.)

In spite of these improved on-time rates, delay is still problematic. Barnhart and Bratu [4] found that while there were fewer delayed flights in 2000 than in 1995, there was more total passenger delay. Passenger delay is defined as the sum of the delays to individual passengers in arriving at their destinations, and is measured in passenger minutes.

In part, passenger delay has increased despite improved on-time rates because the delays that did occur were longer. For example, the number of flights delayed more than 45 minutes nearly doubled between 1995 and 2000. Total passenger delay has also increased because passenger load factors are higher—in other words, planes are carrying more passengers. Finally, more and more of those passengers are connecting through hub airports rather than flying directly to their destinations. Connecting passengers often incur more delay than terminating passengers, because they may miss their connections.

A significant fraction of passenger delay is unavoidable. When a major snowstorm
causes the runways at an airport to ice over, it may be completely shut down for hours.

However, some amount of passenger delay is due to inefficient use of airport resources, and is thus recoverable. The goal of this thesis is to evaluate a proposed mechanism for allocating these resources, and to determine whether it would reduce overall passenger delay.

1.1 Arrival Slots: A Scarce Resource

When the arrival capacity of an airport is reduced due to poor weather or high traffic volume, the Federal Aviation Administration (FAA) issues a Ground Delay Program (GDP), holding planes on the ground at their origins to avoid incurring costly in-air delays.

As part of the GDP, airlines are assigned arrival slots, or time slots at which the airline has permission to land a plane at the reduced-capacity airport. For many years, under a system known as “Grover Jack,” these slots were simply assigned according to the original schedule of arrival times. The airlines were then free to reassign slots to any of their flights, provided that flights did not arrive more than twenty minutes before their originally scheduled arrival times. Airlines were also free to cancel one or more of their flights in response to a GDP.

The major problem with this system was that airlines would cancel their flights and fail to report their unused arrival slots to the FAA. If reported, these slots could potentially be redistributed to other airlines and used to reduce systemwide delay.

The current slot allocation system, known as Collaborative Decision Making (CDM) and first implemented in 1998, attempts to give airlines more of an incentive to report flight cancellations. Under CDM, airlines are encouraged to release their unused arrival slots to the FAA, and any released slots are then redistributed. To ensure that the airlines have an incentive to report their cancelled flights, an airline that releases a slot receives priority for any slots that may open up as a result.
1.2 The Need for Incentives in Slot Exchange

Despite the incentive provided by CDM, airlines do not always inform the FAA when they decide to cancel a flight. For example, on a typical GDP day at one airport implementing CDM, 45 of 124 cancellations, or 36%, went unreported [3].

There are a few possible causes for this high percentage of unreported cancellations. First, CDM may not provide enough direct incentive for airlines to report cancellations. Second, there could be some extra value to the airline for withholding an unused slot, and thereby preventing it from being redistributed to other airlines. For example, concerns about market share can drive airlines to withhold information about cancelled flights, so that rival airlines do not benefit from a better slot time. Finally, risk-averse airlines may hold onto unused arrival slots in order to preserve their options in the case that the GDP is revised.

This thesis proposes an auction mechanism that would address these issues. Under the auction mechanism proposed, slots are first assigned to the airlines according to the original schedule. The slots are then auctioned in sequential order by the FAA on behalf of the initial slot owners. The proceeds from an individual slot auction go to the initial owner. Additionally, the owner of the slot can set a minimum bid price for the slot; setting this price to a very large number ensures that the slot will not be sold.

Because proceeds from the auction go to the initial slot owners, each airline has a direct monetary incentive to give up slots it does not need. Additionally, any extra value to the airline for withholding an unused slot would be captured within the price at which the slot is sold. This extra value would include both the value of withholding slots from rivals, and the premium necessary to compensate the airline for the risk of a GDP revision.

The proposed auction mechanism was simulated using actual flight data. Additionally, two other slot allocation methods were simulated for comparison—the CDM slot allocation method currently in use, and a “global optimization” method, which simulates how slots would be allocated if a single airline owned all flights and all slots.
The global optimization method was simulated to determine an upper bound on the delay reduction that is achievable.

Determining the total amount of passenger delay for each method requires that airlines decide how to delay or cancel particular flights, given some set of available arrival slots. As part of the thesis, a model for solving this problem, the airline recovery problem, was developed and formulated as an integer program.

1.3 Overview of the Thesis

Chapter 2 gives an overview of issues in auction mechanism design and reviews previously proposed auction mechanisms for airport arrival slots. Previous work on the airline recovery problem is also reviewed.

Chapter 3 describes the three arrival slot allocation methods that were simulated—the current method known as Collaborative Decision Making, the global optimization method simulated to determine a lower bound on the amount of achievable delay, and finally, the proposed auction mechanism.

Chapter 4 describes the model of the airline recovery problem used to determine, in each of the three allocation methods, how airlines decide how to delay or cancel flights in response to a GDP.

Chapter 5 provides detail on how actual flight data was reconstructed for the simulation, and how passenger itineraries were stochastically generated.

Chapter 6 describes the results of the simulation. In particular, the auction mechanism is compared to the other two slot allocation methods. Also examined is the effect of varying an airline’s bidding strategy in the auction.

Chapter 7 concludes the thesis and describes ideas for future work.
Chapter 2

Background and Related Work

This thesis proposes a particular mechanism for auctioning airport arrival slots. However, several questions arise in the design of any auction mechanism. These issues are examined in 2.1. Previously suggested mechanisms for auctioning arrival slots in particular are reviewed in 2.2.

Additionally, this thesis addresses the question of how an airline might determine the value of a particular arrival slot. Determining this value requires that the airline decide how to optimally delay or cancel flights, given some set of arrival slots. Previous work on this problem, known as the airline recovery problem, is reviewed in 2.3.

2.1 Design of Auction Mechanisms

Auctions range in complexity from the simplest ascending auctions on eBay, to multiple-round, simultaneous auctions such as those used by the Federal Communications Commission (FCC) to auction off regions of the electromagnetic spectrum. Therefore, it is important to keep the underlying goals of the auction in mind when designing an appropriate mechanism.

Many of the issues relevant to the auctioning of airport arrival slots are similar to those considered in the design of the FCC spectrum auctions. The main goal of the spectrum auction was to ensure efficient usage of the medium, while preventing monopolies, promoting licenses for minority and women-owned businesses, and lastly,
maximizing revenue for the government [13].

The main goal of any auction of airport arrival slots would also be focused on efficiency, and in particular, on reducing systemwide delay. An additional goal of a slot auction would be to preserve fairness of the slot distribution, so that small airlines with low liquidity are not locked out of the bidding process.

In general, there are a few different variables that are considered when designing auction mechanisms. Bidding can be open, or sealed to preserve anonymity. Auctions for several items may occur in sequence, or all at the same time. Separate auctions may be held for each item, or items may be “packaged” together. Finally, the rules of the auction must specify how the winner is determined, as well as the amount the winner must pay.

2.1.1 Open vs. Sealed Bidding

The most well-known type of auction is the English auction, also known as the ascending price auction. Typically used to sell artwork and other high-priced items, the English auction is open in that bid values and bidder identities are public information.

The advantage of an open auction is that it reduces the effect of the “winner’s curse.” This term describes the situation where the winner overestimates the value of an item, and thereby ends up paying more than it is actually worth. Open auctions alleviate the winner’s curse because bidders learn what the item may be worth to other bidders; this allows them to obtain a better estimate of the item’s true value [13].

However, open bidding is more susceptible to bidder collusion than sealed bidding, and may actually yield lower revenues if the bidders are risk averse [7]. Cramton and Schwartz [8] found that some bidders in the FCC spectrum auctions, where bidding was open, colluded by encoding information directly within their bids. Essentially, bidders used the low digits of their bid values to signal other bidders away from particular licenses they were pursuing, and thereby forced an early halt to bidding on those licenses. In doing so, the colluding bidders won 40% of the spectrum and were able to pay significantly less than other bidders.
Note that it is not possible to maintain bidder anonymity and also reveal bid values, because bidder identities can be coded within bids!

If airlines are mainly concerned with reducing their delays and recovering quickly from a GDP, then it is unlikely that they would gain much information about the “true value” of the slot from knowing other airlines’ bids. In fact, an airline can calculate the total reduction in delay due to owning a particular slot, as described in Chapter 4. Because each airline has a different flight schedule, a slot should actually be worth different amounts to different airlines.

Since the GDP might be revised, it is true that the airlines may not know the final value of the slot. However, it is unlikely that other airlines would have better information about possible GDP revisions, which might be revealed through open bidding. Therefore, open bidding on arrival slots should not significantly reduce the effect of the winner’s curse.

Furthermore, sealed bidding helps prevent airlines from colluding or retaliating against competitors. If bidders submit a single sealed bid, they cannot collude unless they manage to come to an agreement outside the auction framework—but they could potentially do so under any auction mechanism.

Thus, because collusion could potentially have a large effect on reducing the efficiency of the auction, and because the role of the winner’s curse is negligible, sealed bidding is preferable to open bidding for a real-time auction of airport arrival slots.

### 2.1.2 Sequential vs. Simultaneous Bidding

When multiple items are up for auction, they can be auctioned sequentially, or all auctions can occur simultaneously. In the latter case, bidders continue to place bids on all slots until the auction is declared over.

If bidding is open, sequential auctions are simpler to implement in practice and generally run more quickly. However, sequential bidding is less efficient than simultaneous bidding if items are more valuable when aggregated. FCC spectrum licenses, for example, are more valuable if put together to cover larger blocks of the spectrum. Since bidding on many licenses occurs at the same time, bidders can coordinate their
bids on licenses that would aggregate well [13].

The main problem with open simultaneous auctions is that they can potentially run for a very long time, since bidders are unwilling to bid on particular slots before seeing the high bids for other slots. Even if a time limit is put on the auctions, bidders may wait until just before the auction is closed to place their bids, in which case the outcome may not be efficient [13].

If bidding is sealed and items are auctioned individually, then it is likely more efficient to hold auctions sequentially rather than simultaneously. There is no way for bidders to coordinate bids on different items to create a desired package if the high bids of other bidders are unknown. Also, sequential bidding allows bidders to use the information from earlier, finalized auctions when deciding how to bid on later items. Simultaneous bidding may lead to bidders ending up with an undesired set of items if bidding is sealed.

If items are auctioned in packages as well as individually, then clearly simultaneous bidding is required, since the auctions for individual items in a package must conclude at the same time as the auction for the package.

Therefore, the decision on whether to implement sequential or simultaneous bidding hinges on whether bidding is sealed and on whether package bidding is allowed; sequential bidding is only preferable if bidding is sealed and package bidding is not allowed, or if a simultaneous auction is too complex to implement.

2.1.3 Individual vs. Package Bidding

Bidding on items individually rather than in packages can lead to inefficiencies. In particular, if two items are worth more when aggregated, the result of auctioning them individually may not be optimal.

For example, flights often arrive and take off from airport hubs in banks. A set of bank slots may be worth more than the sum of the individual slot valuations, since passengers transfer to a departing flight from among several flights in an arriving bank.

Suppose that two arrival slots are being auctioned, and all airlines value either
one of the slots at $100, but they value the pair of slots at $250. Then, auctioning
the slots individually may result in two different airlines winning the slots for $100
each. However, it would be more efficient to distribute both slots to a single airline.

Allowing bids on packages of items resolves this particular inefficiency. However,
package bidding can still be inefficient if bidding is open. For example, suppose that
two arrival slots are being auctioned, and airline A values the pair of slots at $250 and
either slot alone at $0. Assume also that airline B values the first slot alone at $150,
and airline C values the second slot alone at $150. Then, it is optimal for airlines B
and C to each receive one slot.

But if airline A bids $250 when B and C’s high bids thus far are only $120, then
B would prefer to hold his bid at $120 and allow C to raise his bid such that the total
of B’s and C’s individual bids exceeds $250. Likewise, C would prefer that B raise
his bid. This is an example of the free-rider problem [7].

Allowing package bidding can also be extremely complex to implement. If \( n \) items
are up for auction, then \( 2^n \) possible packages can be assembled and bid on.

Thus, package bidding has both advantages and disadvantages. If there are rela-
tively few combinations of packages and bidding is sealed, allowing package bidding
improves the efficiency of the auction. If bidding is open, it is unclear whether there
is an efficiency gain, and if there are many possible combinations of items, an auction
with package bidding may simply be unworkable.

2.1.4 The Vickrey Auction

In any auction, it is necessary to specify how to determine the winner, and also how
much the winner must pay for the item being auctioned.

In a Vickrey auction, also known as a second-price auction, the highest bidder
wins, but pays only the amount of the second highest bid. The item’s seller can also
set a minimum sell price, or reservation price, to ensure that the item is only sold if
the bids surpass a certain amount. The nice property of the Vickrey auction is that
it induces bidders to bid their true valuations of the item being auctioned, since that
is a weakly optimal strategy [10].

27
For example, suppose that two agents (A and B) are bidding for a house in a sealed-bid, Vickrey auction. A’s valuation for the house is $500,000, and A does not know B’s bid value $V_B$.

If A wins the auction, he will pay $V_B$ no matter what he bids. However, if A bids below $500,000, he is less likely to win the auction than if he bids $500,000. Therefore, A cannot gain by lowering his bid below $500,000.

If A bids above $500,000, then he may end up winning and paying more than $500,000. For example, suppose that A bids $510,000 and B bids $505,000. Then A wins and pays $505,000 for the house, which results in a net loss of $5,000. In fact, A cannot gain by raising his bid above $500,000; if he does so, he will lose utility unless $V_B \leq 500,000$.

Since A cannot gain by deviating from a bid of $500,000, that value is an optimal bid for A.

There are several other general types of auctions, including the English auction described above, and the first-price sealed bid auction, where the winner pays the amount of the second-highest bid. However, several of these auctions actually yield the same result. The Revenue Equivalence Theorem states that certain types of auctions where the highest bidder wins, and where the bidder with the lowest feasible valuation receives no surplus, yield the same revenue [10]. For example, the English auction and the second-price auction yield the same revenue, since bidding in the English auction stops when the bidder with the second highest valuation bids his valuation.

### 2.1.5 Application to the Arrival Slot Auction

The auction considered in this thesis is a Vickrey auction. All slots were auctioned sequentially and individually, and bidding was sealed.

Package bidding was not allowed due to the complexity of implementation. Also, to ensure fairness of slot allocation, it was decided that slots would first be distributed to airlines and sold by the individual airlines through the FAA, rather than sold directly by the FAA. Since items cannot be packaged if there are multiple sellers, this
decision precluded the possibility of allowing package bidding. Bidding was sealed in order to prevent collusion among bidders. Slots were bid on sequentially rather than simultaneously, because sequential bidding is likely more efficient given that bidding is also sealed and package bidding is not allowed, as explained in 2.1.2.

The Vickrey auction was chosen largely because it simplifies the simulation of bidding. It is fairly straightforward to calculate the value of a slot to an airline. Since the optimal strategy in the Vickrey auction is to bid the valuation, determining the bid values is also straightforward.

2.2 Previous Work on Auctioning Arrival Slots

The idea of auctioning airport arrival slots assigned during GDPs has been examined before. A few different auction mechanisms have been proposed, although few have been simulated on actual flight data. Three of the proposed mechanisms are described below; all are simultaneous, sealed-bid auctions where the FAA acts both as the seller and the auctioneer.

2.2.1 Combinatorial Auction (Rassenti)

Rassenti, Smith, and Bulfin [15] developed a sealed-bid, “combinatorial” auction for allocating arrival slots to airlines. In the Rassenti auction, airlines are free to submit bids on all possible packages of slots. Additionally, airlines are allowed to place contingency bids, or constraints on accepting certain combinations of packages. For example, an airline is allowed to specify that it will only buy package A if it also wins package B, and that it will buy at most \( n \) packages of some set.

The auction was simulated using human subjects to represent the airlines. Packages were assigned fixed values. After subjects submitted sealed bids for their desired packages, a linear optimization was solved via computer to maximize total revenue, and per-slot marginal prices were approximated from the bids. (The true marginal price of a slot would be the lowest amount the winner would be able to pay and still
win the slot. Because package bids were allowed, these prices could not be determined exactly.) The winner of a package was paid the fixed value of the package, minus the sum of the marginal prices of all the slots in that package.

It was found that in an auction of six possible packages of six slots, over 90% of the total possible payoff was realized by the auction’s participants. Thus, the Rassenti combinatorial auction may be an efficient method of allocating slots. However, it does not guarantee that the resulting slot distribution is fair, and would have to be simulated on actual flight data to determine its true efficiency.

2.2.2 Multi-Object Auction (Milner)

Milner [14] proposed a simultaneous, sealed-bid, second-price auction without package bidding. In Milner’s auction, airlines first report the value of assigning a particular flight to a particular slot, for each possible flight-slot combination. The FAA then solves a linear optimization to assign flights to slots based on the reported values, thereby maximizing total value. After the assignment, each airline awarded a slot then pays the FAA a fee equal to the second highest bid for the slot.

A few major problems with the multi-object auction have been noted [11].

First, it is not incentive-compatible. Since flights from the same airline compete with each other for the same slots, airlines do not have the incentive to report their true slot valuations. This is known as the coalition problem. By stating a slot value lower than the actual valuation, a coalition of flights (in this case, the airline) may reduce the price it pays and make a larger overall profit, despite winning fewer slots in doing so.

For example, suppose that two airlines are bidding for three available arrival slots in a second-price auction. Airline A needs only one slot for its one flight, and any one of the slots would result in reducing systemwide delay for the airline by 90 passenger minutes. Airline B has three flights, and each of the slots would result in reducing delay by 100 passenger minutes. Assume that each minute of reduced passenger delay is worth $1 to either airline.

The most efficient outcome in terms of reducing overall delay is for B to be assigned
all three slots, which occurs if A and B both bid their valuations for each slot. A bids $90 for each slot and B bids $100 for each slot. Then, B wins all three slots at $90 each, and B’s profit is $30.

However, if B bids $0 for each of the three slots, it will win two of the slots and pay nothing; in that case, B’s total profit is $200. Airline 2 makes a bigger profit in the second case, but the overall outcome is less efficient in terms of reducing overall delay.

Another major problem with the multi-object auction is the inefficiency of auctioning slots individually rather than in packages, as described in 2.1.3. Airlines may be willing to bid more for packaged slots, since flights are often grouped into banks.

Thus, Milner’s auction is less efficient than other possible auction mechanisms. Furthermore, it does not guarantee that the resulting slot distribution is fair.

2.2.3 Groves Mechanism (Hall)

Hall [11] proposed modifying Milner’s multi-object auction to correct incentives, using a Groves mechanism. Under the Groves mechanism, each airline is charged a fee equal to the lost value it causes all other airlines through its presence in the solution. The effect of this fee is to align each airline’s incentive with the incentive of the FAA, which is to maximize the sum of all airlines’ valuation functions.

For example, suppose that two airlines, A and B, are bidding for a single slot in the Milner auction, and that A’s valuation for the slot is 2 and B’s valuation is 1. The assignment variable $x_A$ equals 1 if the slot is assigned to airline A, and 0 otherwise; the assignment variable $x_B$ equals 1 if the slot is assigned to airline B, and 0 otherwise. Let $Z$ represent the value of the second-highest bid, unknown to the bidding airline. Then A seeks to maximize $2x_A - Z$, and B seeks to maximize $x_B - Z$. The FAA seeks the global optimum obtained by maximizing $2x_A + x_B$.

Under the Groves mechanism, the fee assessed A is the value that would be achieved by all other airlines if A received no slots, minus the value achieved by all other airlines in the actual assignment. In the example, if the value achievable by all other airlines without A is $V$, then this fee is equal to $V - x_B$. Then A maximizes
2x_A + x_B - V, which is equivalent to maximizing the FAA’s global objective.

In effect, the Groves mechanism generalizes the per-slot marginal prices used in the Rassenti and Milner mechanisms, to an overall, per-airline marginal price.

However, a few other problems that are not resolved by the Groves mechanism can be pointed out. One problem is that any system of imposing fees on airlines is politically infeasible and would be difficult to implement. This is particularly true if the fees are seemingly unrelated to the bids in an auction, as is the case with the Groves fees. Additionally, in practice it is difficult for airlines to come up with individual valuations for every possible flight-slot combination. These problems have prevented a practical auction system from being implemented to date.

Thus, the Groves mechanism is a very efficient auction mechanism, although it is politically infeasible to implement and does not guarantee that the slot distribution is fair.

### 2.3 The Airline Recovery Problem

In any auction of arrival slots assigned for GDPs, an airline is faced with determining the value of some slot (or of some set of slots). In order to determine this value, an airline must determine how it will cancel or delay each of its scheduled flights, given some set of arrival slots. This determination of how to best recover from the GDP is characterized as the airline recovery problem.

The airline recovery problem can be decomposed into several subproblems [6]:

- **Fleet assignment** – Each flight must be assigned an appropriate aircraft type, or *fleet* type.

- **Aircraft rerouting** - Each aircraft must be assigned some sequence of flights, such that each scheduled flight is included in exactly one route, and all other flights are cancelled. Also, the aircraft’s maintenance requirements must be satisfied.
• **Gate assignment**  Each scheduled flight arriving at or departing from some airport must be assigned to a gate at that airport.

• **Slot allocation**  – Each arriving flight must be assigned to some arrival time slot, if a GDP has been issued for the arrival airport.

• **Crew scheduling**  – Each flight must be assigned a flight crew, such that the crew begins and ends at their home airport, and such that any work regulations are satisfied.

• **Passenger rerouting**  – The airline must reassign passengers to flights. This is necessary for passengers whose flights have been cancelled, and for passengers whose flights are delayed so much that they miss a connecting flight. Also, an airline may reroute passengers in order to reduce their delay and prevent them from jumping to a competing carrier.

2.3.1 **Set-Packing Model (Clarke)**

Clarke [6] proposed a set-packing model based around solving the fleet assignment and aircraft rerouting problems. All of the other airline recovery subproblems are incorporated as additional constraints.

First, the operating cost of assigning a particular aircraft to a particular flight is determined for all such assignments, and the cost of cancelling a particular flight is determined for all flights. Then, all possible flight sequences that meet the operational constraints are generated.

An integer program is then solved to minimize the total cost of the operated flights and the cancelled flights, subject to constraints on aircraft utilization and maintenance, crew availability, slot allocation, and gate allocation. Also included are constraints on flight covering (every flight must either be flown or cancelled), aircraft covering (every aircraft is assigned to exactly one route), demand covering (all passengers on a flight must be accommodated by the aircraft type), and aircraft balance (at particular times, each airport must contain the correct number of aircraft).
After each aircraft is assigned a flight sequence using the integer program, the scheduled arrival and departure times of all flights are then revised to match the resulting assignment.

Clarke’s set-packing approach is comprehensive, and addresses almost every consideration an airline might take into account when deciding how to reroute aircraft in response to a GDP. However, the breadth of the model makes it too complex for quickly solving large instances of the airline recovery problem.

2.3.2 Set-Packing with Selection Heuristic (Rosenberger)

The main problem with the Clarke model is that it is computationally complex, and requires generating all possible routes. Rosenberger, Johnson, and Nemhauser [16] reduced the complexity of the problem by eliminating some of the subproblem constraints, and by adding a heuristic that selects a subset of aircraft. Route possibilities were generated only for this subset of aircraft.

The heuristic first selects all disrupted aircraft, or aircraft affected by the GDP. It then chooses additional aircraft whose original routes are especially likely to combine well with the disrupted aircraft to form valid alternative routes.

Using the heuristic, the Rosenberger model was able to solve instances of the airline recovery problem with 96 aircraft and 469 legs in only 16 seconds per instance.

In practice, any real-time auction of arrival slots would have to occur only hours before the start of a GDP. This requires that airlines decide on bid values very quickly.

The approach taken in this thesis simplifies the airline recovery problem even further, by eliminating almost all constraints other than the aircraft covering and flight covering constraints, and by pregenerating possible assignments of slots to routes. This approach allows large instances of the airline recovery problem to be solved within seconds. Thus, it is most appropriate to the auction, where practical constraints may limit bidding time per slot to under a minute.
Chapter 3

Arrival Slot Allocation Methods

Three slot allocation methods were simulated in this thesis—the auction mechanism, and two other methods that establish benchmarks for measuring the performance of the auction.

The allocation method currently in practice, known as Collaborative Decision Making, was simulated to determine the total amount of delay that results under the current system. The global optimization method effectively simulates what would occur if the FAA were to allocate arrival slots in the most efficient way possible, given information about passenger itineraries. This method determines a lower bound on the total amount of achievable delay.

Both of these methods, as well as the auction mechanism, rely on an initial slot allocation procedure known as Ration-By-Schedule.

3.1 Initial Slot Allocation via Ration-By-Schedule

The Ration-By-Schedule (RBS) procedure (Alg. 3.1) assigns available slots to the airlines in the order of the original schedule of flights. The goal of the RBS algorithm is to ensure that the initial allocation of slots is fair.

The Airport Arrival Rate (AAR), or the number of allowed arrivals per hour, is assumed to be known for any given period in the day, and determines the available slot times. For example, if the airport arrival rate at some airport is reduced to 20
Algorithm 3.1 Ration By Schedule

1: begin
2: First, knowing the airport arrival rate (AAR) for each time interval throughout the day, determine the available slot times throughout the day, starting with the initial period of reduced capacity.
3: Next, determine the set of flights $F$ arriving at the GDP airport after the start of reduced capacity, and order the flights by their original scheduled time of arrival.
4: for all $f$ in $F$ do
5: Create a slot $s$ for $f$ at the first available unassigned slot time it can meet, and assign $s$ to the owner of $f$.
6: if the original arrival time for $f$ equals the slot time for $s$ then
7: stop
8: end if
9: end for
10: end

arrivals per hour between 10:00am and 11:00am, arrival slot times are available every three minutes within that hour. However, whether a slot is actually created and assigned for each of these times depends on the original schedule. If 30 flights were scheduled to arrive within that hour, but all of them were scheduled to arrive after 10:30, then no slots would be created in the first half hour. At most 10 of the flights could be assigned to slots between 10:30 and 11:00, since there are only 10 available slot times in that period. The rest of the flights would be assigned to slot times in subsequent periods.

The end of a GDP is defined as the time, after the final period of reduced capacity, at which operations return to normal. It is determined by the first flight that incurs no delay in the RBS allocation, i.e. whose slot time is equal to its originally scheduled arrival time.

3.2 Collaborative Decision Making

The Collaborative Decision Making approach to slot allocation has three stages. In the first stage, the FAA provides an initial assignment of slots to the airlines using the RBS procedure. In the second stage, substitution and cancellation, airlines are allowed to reassign their slots among their scheduled flights and then release any
unused slots to the FAA. In the third stage, known as compression, the FAA tries to swap these released slots among the airlines in order to maximize slot usage.

Typically, the initial RBS allocation occurs once, and the remaining phases are repeated at regular intervals. Compressing slots regularly ensures that an airline may still release an unused arrival slot if conditions change and it decides at a later point to reprioritize or cancel particular flights. In this thesis, GDP conditions were assumed to be fixed, so it was unnecessary to repeat the substitution and cancellation phase or the compression phase.

### 3.2.1 Substitution and Cancellation

During the substitution and cancellation phase of CDM, airlines are free to reassign slots to flights. Assigning slots to airlines rather than to individual flights gives the airline some measure of flexibility; an airline may decide, based on its own operational data, which flights it would like to prioritize over others. Additionally, airlines may decide to cancel individual flights.

One restriction on substitution, imposed by the FAA, is that no flight may be assigned to a slot time more than 20 minutes before its originally scheduled arrival time.

The substitution and cancellation phase can be modeled as an assignment problem, where the airline maximizes its objective value (or equivalently, minimizes its cost) while assigning flights to slots.

### 3.2.2 Compression

The compression phase of CDM (Alg. 3.2) allows for released slots to be redistributed to airlines that can better take advantage of them [18]. At regular intervals, the FAA redistributes all arrival slots that were released by the airlines in the previous substitution and cancellation phase.

The key idea behind compression is that the airline that released the slot receives priority for the slot that is freed up as a result of compression. If it cannot use the
Algorithm 3.2 Compression

1: begin
2: Determine the set \( C_s \) of slots released by all airlines in the substitution and cancellation phase, and order this set by slot time.
3: Let \( c \), the current open slot, equal the first slot in \( C_s \), and let \( a \) be the airline to which \( c \) is assigned.
4: while \( C_s \) is not empty do
5: Let \( g \) be the first flight from airline \( a \) that can be assigned to \( c \).
6: If there is no such flight, let \( g \) be the first flight from any other airline that can be assigned to \( c \).
7: if \( g \) is not null then
8: Let \( d \) be the slot assigned to flight \( g \).
9: Swap slots \( c \) and \( d \), by assigning slot \( c \) to flight \( g \) and airline \( b \), and slot \( d \) to airline \( a \).
10: Let \( c \), the new current open slot, equal \( d \).
11: else
12: Remove \( c \) from \( C_s \) and let \( c \) equal the next slot in \( C_s \).
13: end if
14: end while
15: end

slot that is freed up, it receives priority for any subsequent slots that are opened up as a result of compressing the newly freed slot. This priority-based swapping method is what provides an incentive in CDM for airlines to release resources they do not need.

3.3 Global Optimization

The global optimization algorithm simulates the slot assignment and passenger delay that would result if one airline owned every slot and every flight. In such a case, the airline would try to allocate all slots to GDP arrival flights in the most efficient way, given the information available about passenger itineraries.

A single airline would also be able to exploit other efficiencies, unrelated to slot allocation, in order to reduce passenger delay. For example, by combining the fleets of the individual airlines, the single airline would have more options for rerouting aircraft. While each individual airline's aircraft are restricted to flying that airline's flights, the combined airline's aircraft would be able to fly routes with flights pieced
together from all airlines.

Additionally, the single airline would be able to more easily reroute passengers disrupted by the GDP. For example, the single airline might be able to cancel an American Airlines flight arriving at the GDP airport, and reroute its passengers onto a Continental flight scheduled to depart ten minutes later for the same destination.

The global optimization was modeled simply as a large instance of the airline recovery problem. All flights, slots, and aircraft are assigned to one large airline. After the FAA creates slots according to the RBS procedure, this large airline solves an integer program to determine an efficient allocation of slots to flights.

Because the global optimization was simulated only to benchmark the delay achievable by allocating slots optimally, the simulation purposefully did not take advantage of the efficiencies unrelated to slot allocation. In particular, the possibility of generating routes using flights from multiple airlines was excluded. Also excluded was the possibility of rerouting passengers from one airline to flights for a different airline.

3.4 The Auction Mechanism

The auction mechanism (Alg. 3.3) allows for slots to be exchanged between airlines through the FAA as an intermediary. The major goals in the design of the auction mechanism were to provide a greater incentive for airlines to release unneeded slots, and to maintain fairness of slot allocation. The slot auction provides a direct monetary incentive for airlines to give up slots they do not need. Fairness is preserved in that just as in CDM, an airline has the option of retaining any slot it is initially assigned.

The mechanism is a sealed bid, second price (Vickrey) auction; that is, bids are unknown to other bidders, and the winner of the slot pays only the amount of the second highest bid. Slots are initially assigned to airlines using the RBS algorithm, and then each slot is auctioned off one at a time, in order of the original schedule. The FAA serves as an auctioneer, and facilitates the exchange of slots between the airlines.
The airline that is initially assigned a slot is designated as the seller of the particular slot, and all other airlines have the opportunity to bid for the slot. In the event that the slot is successfully won by a bidder, the winning airline pays the seller for the right to the slot. It is assumed that there are no liquidity constraints—in other words, there are no limits on the amounts airlines can “spend” when purchasing slots.

Before a particular slot is auctioned off, an airline sets its reservation price, or the minimum price for which it is willing to sell the slot; the reservation price is unknown to the bidders. By setting the reservation price to a very high value, the airline can ensure that it keeps the slot. This mechanism achieves the goal of fairness, by allowing airlines to keep the slots they were originally assigned if they so desire.

Note that if the second highest bid is higher than the reservation price for the slot, the price paid by the winner is set equal to the second highest bid. However, if the second highest bid is not higher than the reservation price, the winner pays the reservation price. [17]

Each airline places only one bid per slot, in contrast to Milner’s multi-object auction [14], where each airline places a bid for each possible flight-slot pairing. Therefore, the “coalition” problem described in Chapter 2 is avoided. This implies that the airlines will each bid the true value to them of owning the slot, since that is an optimal strategy for any bidder in a second price auction [10].

3.4.1 Determination of Bid Values

Since airlines are assumed to exhibit rational behavior, the bidders were modeled as bidding their valuation of the slot. However, it is not immediately obvious how to calculate this value. In this simulation, the valuation was assumed to be the same as the marginal value of the slot with respect to the bidder’s slot holdings at the time of the auction.

Since slots are considered one at a time, in the order of the original schedule, each airline can determine the marginal value of the slot up for auction in order to place a bid. That is, an airline evaluates the slot value as the overall reduction in cost to the airline that would result if the airline owned the slot.
Algorithm 3.3 Auction Mechanism

1: begin
2: Create the set of slots $S$ using the Ration-By-Schedule (RBS) procedure, and order the slots by slot time.
3: for all $s$ in $S$ do
4: Determine the reservation price $r$ for the airline $a$ to which $s$ was initially assigned. ($r$ is unknown to the bidding airlines.)
5: Determine the bid values for $s$ for all airlines other than $a$. (Although it was not necessary in the simulation, in practice bids would be accepted only up until some deadline.)
6: Let $v_1$ be the maximum bid value, and $b$ be the highest bidder:
7: if $v_1 > r$ then
8: Assign $s$ to $b$.
9: Let $v_2$ be the second highest bid value.
10: if $v_2 > r$ then
11: Increment $a$’s account by $v_2$ and decrement $b$’s account by $v_2$.
12: else
13: Increment $a$’s account by $r$ and decrement $b$’s account by $r$.
14: end if
15: end if
16: end for
17: end

This cost reduction was calculated by solving the flight-slot assignment problem twice. In the first instance, the airline assigns its flights among the set of slots $S$ it currently owns. In the second instance, the airline additionally includes the slot $s$ up for auction, assigning its flights to $S \cup \{s\}$. The value of the objective function is calculated for each case, and the two values subtracted to obtain the marginal value of the slot.

3.4.2 Determination of Reservation Price

The reservation price, or minimum sell price of the RBS slot owner, was also determined as the marginal value of the slot to the owner. The airline assigned an arrival slot $s$ through RBS determines the slot value by first solving the assignment problem with the current set of slots $S$, and then solving the assignment problem with the set of slots $S \setminus \{s\}$. Subtracting the two objective function values yields the marginal value of the slot.
3.4.3 Alternative Airline Strategies in the Auction

A few alternative slot valuation strategies were explored to determine whether it is possible for airlines to improve on their final slot assignments. The bid and reservation price determination functions were modified for one or more airlines to determine whether an alternative strategy confers an advantage in the auction.

“Predictive Bidding”

The assumption that airlines bid the marginal value of the slot partially resolves the problem described in 2.1.3, where slots are worth more in aggregation. Airlines bid the additional value of the slot relative to their current slot holdings at the time of the auction, rather than the value of the slot independent of all other slots. A complete combinatorial auction, where airlines could bid on different packages of slots rather than on one slot at a time, would eliminate the aggregation problem but might introduce other inefficiencies.

However, the airline would bid differently for a particular slot if it knew what the final allocation of slots was likely to be. A true solution that addresses aggregation would be for airlines to bid the marginal value of the slot relative to the final set of slots allocated to the airline. In fact, it may be that an airline has a better estimate of the final set of slots it will end up with. If so, it could bid the additional value of the slot relative to the airline’s expected final slot holdings, rather than relative to the current slot holdings.

This “predictive bidding” strategy was simulated by running the auction twice. In the first auction, all airlines bid using the “naive” strategy of determining the marginal value of a slot relative to the airline’s current slot holdings. In the second auction, one or more airlines were allowed to compute their bids by determining the value of the slot relative to the final slot holdings that resulted from the first auction. The underlying assumption was that the result of the first auction serves as a better predictor of the final allocation in the second auction than the airline’s current slot holdings!
The appeal of predictive bidding as a strategy was tested by comparing the total amount of passenger delay resulting for a particular airline in the case that all airlines bid naively, to the delay in the case that the one airline bid predictively and all others bid naively. Also, the effect of increasing the number of predictive airlines on the overall amount of passenger delay was tested.

"Cautious Selling"

Given that airlines currently sometimes refuse to give up slots in CDM even if the slots are unused, it is reasonable to think that an airline might refuse to give up a slot in the auction as well. Usually, the purpose of holding onto a slot is to hedge against the possibility that the airline may decide at a later point in time to use the slot, for instance if airport conditions change. Or, it may be that an airline fears that giving up a particular slot would induce a shift in market share to other airlines in the long run.

The tendency of an airline to hold onto its slot, for whatever reason, is referred to as cautiousness. It can be simulated in the auction by setting the airline’s reservation price to positive infinity with some probability $p$. This behavior can likewise be simulated in the CDM slot allocation method by assuming that with the same probability $p$, the airline refuses to release the slot in the substitution and cancellation phase.

The “cautious selling” strategy was tested in the auction by varying the caution probability $p$ for one particular airline $a$, setting all other airlines to bid naively, and observing the change in total passenger delay for $a$. The effect of increasing the number of cautious airlines (each with some fixed caution probability) on the overall amount of passenger delay in the auction was also tested. Finally, the effect of increasing the number of cautious airlines on the overall amount of passenger delay resulting under the CDM slot allocation method was tested.
Chapter 4

Modeling the Airline Recovery Problem

In all three slot allocation methods described in Chapter 3, it is assumed that each airline has some sort of mechanism to determine which flights to delay or cancel in response to the GDP, and to determine which flights to assign to particular arrival slots. In the auction, this mechanism is used by airlines to determine the marginal value of a slot. In Collaborative Decision Making, it is necessary for the substitution and cancellation phase. Finally, such a mechanism is the entire basis of the global optimization allocation method.

The problem of assigning flights to slots and determining how flights are cancelled and delayed is referred to as the airline recovery problem, and was formulated as an integer program. While the airline recovery problem may in general be made arbitrarily complex, incorporating subproblems such as crew scheduling and gate assignment, the model has been simplified here to one of assigning each aircraft one of several alternative routes and simultaneously assigning an arrival slot to each GDP arrival flight.

The simplification was made in order to reduce the computational complexity of the problem. It can be assumed that subproblems such as crew scheduling and gate assignment affect all three slot allocation methods equally; therefore, the simplified model can be used to estimate the percentage differences in passenger delay between
the three methods. Determining the actual delay values would require a more complete model.

4.1 Overview of the Model

The airline recovery model takes as input a set of aircraft belonging to one particular airline. It computes alternative routes for all aircraft affected by the GDP, and then solves an integer program to determine the most efficient way to reroute the aircraft and, consequently, delay or cancel specific flights.

First, the set of disrupted aircraft is determined. A disrupted aircraft is any aircraft whose original route contains at least one flight affected by the GDP, i.e. any flight assigned a slot through the RBS procedure.

For each disrupted aircraft, alternative routes are constructed using a simple queueing algorithm, which is described in 4.2. The disrupted aircraft can be grouped by fleet type, and route possibilities are generated by constructing all valid sequences of flights requiring the same fleet type. Flights are assumed to require the same fleet type as originally scheduled.

For each possible route alternative, GDP arrivals are then paired with available slots to produce slotted routes. “Slotting” GDP arrivals within a route allows the delays that would result from flying that route to be calculated, based on the assigned slot times, the length of each flight, and the minimum allowed turnaround time between flight legs in a route. Additionally, the total passenger delay that would result from cancelling a particular flight can be estimated.

Finally, an integer program is solved to simultaneously assign routes to aircraft and slots to flights so as to minimize total passenger delay.

4.2 Generating Alternative Routes

The key requirement for generating routes is that flow balance is maintained; essentially, if an aircraft arrives at an airport on one leg of a route, it must depart from the
same airport on its next leg. Another requirement is that an aircraft cannot depart on one flight leg before having landed on its previous flight leg and spent some amount of turnaround time on the ground.

Rosenberger [16] builds route alternatives for a particular aircraft $\phi$ by exhaustively considering all possible flight sequences. In his notation, if the time horizon for the GDP is $(t_0, T)$, and original route $r(\phi)$ contains flights $r = (f_1(\phi), \ldots, f_{n(\phi)-1}(\phi))$ which depart after $t_0$ and arrive before $T$, then the commencing flight $f_0(\phi)$ for aircraft $\phi$ is the leg previous to $f_1(\phi)$ in the original schedule. The terminating flight $f_{n(\phi)}(\phi)$ for route $r$ is the next leg after $f_{n(\phi)-1}(\phi)$. Then, when rerouting a set of aircraft $\Phi$ of the same fleet type, every route alternative for a particular aircraft $\phi \in \Phi$ must begin with $f_0(\phi)$ and end with some terminating flight $f_{n(\phi')}(\phi')$ for any aircraft $\phi'$ in the fleet.

Additionally, for any set of route assignments, each terminating flight must be assigned to exactly one aircraft. This guarantees that all aircraft operations proceed normally after the GDP ends. If the terminating flight $f_{n(\phi)}(\phi)$ is assigned to aircraft $\phi' \in \Phi$, this implies that any subsequent flights in $r(\phi)$ would also be assigned to $\phi'$. Since $\phi'$ is an aircraft of the correct fleet type for those flights, operations can proceed normally. Effectively, $\phi'$ is swapped in for $\phi$ for all flights after the GDP ends.

Note that in the case that the first known flight in the route departs after $t_0$ or the last known flight in the route arrives before $T$, “dummy flights” can be added before and after the known flights in the route to ensure that a commencing flight and a terminating flight exist for all aircraft.

Fig. 4-1 illustrates an example of original routes for two aircraft. The arrows
represent flights that depart and arrive at particular points on the timeline, and the

dotted regions correspond to periods during which the aircraft is on the ground at
the airport indicated.

In the figure, flights 1 and 6 are commencing flights for aircraft A and B, respectively; flights 5 and 11 are the terminating flights. The possible routes that could be assigned to aircraft A are (1,2,3,4,5), (1,2,9,10,11), and (1,2,11). The possible routes that could be assigned to aircraft B are (6,7,8,9,10,11), (6,7,4,5), (6,5), and (6,7,8,11). Flight sequences such as (6,7,8,3,4,5) and (1,2,3,4,7) are excluded because they contain flight legs scheduled to depart before earlier flight legs have landed.

4.2.1 Restricting Route Possibilities Using Subroutes

Generating every possible flight sequence for each aircraft can lead to an enormous
number of possible routes being generated. The model simulated in this thesis takes
advantage of one additional assumption in order to reduce the number of alternative
routes generated. In particular, it is undesirable to cancel flights that do not depart
or arrive from the GDP airport. In the example, if the GDP airport is LAX, one
valid solution is to assign route (1,2,11) to aircraft A and route (6,5) to aircraft B;
however, this assignment would have the effect of cancelling flights 9 and 10, which
neither depart from nor arrive at LAX!

To avoid generating these extraneous routes, each original route is first converted
from a sequence of flights into a sequence of subroutes. A subroute is either a sequence
of flights of which none depart from or arrive at the GDP airport during the GDP
time horizon, or a single flight that departs from or arrives at the GDP airport. The
idea is to ensure that only one aircraft flies any sequence of non-GDP airport flights
within some original route, because there is no need for those flights to be flown by
different aircraft.

Using subroutes rather than flights to generate alternative routes results in vastly
fewer possibilities, and thus saves both memory and computation time in any imple-
mentation. An additional side benefit is that routes can be stored more efficiently
in memory, because subroutes represent sequences of flights that would otherwise be
replicated in many different alternative routes.

The sequence of subroutes for a particular original route \( r(\phi) \) is generated by concatenating any consecutive flights that neither arrive nor depart from the GDP airport. After this is done, all subroutes that arrive before \( t_0 \) are combined into one large subroute, and all subroutes that depart after \( T \) are joined into another large subroute. The reason for the second step is that flights arriving before \( t_0 \) or departing after \( T \) are never cancelled or delayed, and therefore it is unnecessary to generate sets of route alternatives which would assign those flights to multiple aircraft.

In the example, again assuming a GDP at LAX, aircraft A is composed of subroutes (1,2), (3), (4), and (5). Aircraft B is composed of subroutes (6), (7), (8), and (9,10,11). Subroutes (1,2) and (6) are the commencing subroutes, and subroutes (5) and (9,10,11) are the terminating subroutes, for aircraft A and B. These are simply the first and last subroutes \( u_0(\phi) \) and \( u_n(\phi) \) in the original route for the aircraft.

If route alternatives are generated by combining subroutes rather than by combining flights, the possible routes for aircraft A are (1,2,3,4,5) and (1,2,9,10,11). The possible routes for aircraft B are (6,7,8,9,10,11), (6,7,4,5) and (6,5).

Routes (1,2,11) and (6,7,8,11), which were generated as described in 4.2, are no longer generated—any route which includes flight 11 is now required to also include flights 9 and 10, which neither depart from nor arrive at LAX. This excludes any possible solutions in which flights 9 and 10 are cancelled, and any solutions in which flights 9, 10, and 11 are assigned to different aircraft.

Additionally, routes such as (6,7,8,3,4,5) and (1,2,3,4,7) are excluded as before because they contain flight legs scheduled to depart before earlier flight legs have landed.

### 4.2.2 Route Generation Algorithm

The algorithm used to generate alternative routes is shown as Alg. 4.1.

For an aircraft \( \phi \), the algorithm maintains a queue INCOMPLETE of partially completed flight sequences, as well as a list \( R(\phi) \) of completed flight sequences. The queue is initialized to contain the route consisting of only \( \phi \)'s commencing subroute.
Then, for each partially completed route \( r \) in the queue, the algorithm tries to create new partially completed subroutes by adding every possible subroute to \( r \). If \( r \) is already complete, i.e. its last subroute is a terminating subroute, then it is identified as such and added to the set of complete routes.

Subroutes \( u \) that are added must obey the requirement of flow balance, i.e. the first flight in the subroute must depart from the arrival airport of the last flight in \( r \). Also, the original departure time of the added subroute must be no earlier than the original arrival time of the last flight in \( r \), plus the minimum turnaround time \( \gamma \) of the aircraft. Finally, the aircraft whose original route contains \( u \) must be of the same fleet type as \( \phi \).

In order to speed up route construction, a data structure mapping airports to a list of subroutes departing from that airport was constructed, allowing the set of eligible subroutes to be quickly narrowed down.

\[\textbf{Algorithm 4.1} \text{ Construction of Alternative Routes}\]

1: begin
2: for all aircraft \( \phi \in \Phi \) do
3: \( R(\phi) := \emptyset \).
4: INCOMPLETE := \emptyset.
5: Add a new route \((u_0(\phi))\) to INCOMPLETE.
6: for all routes \( r \in \) INCOMPLETE where \( r \) consists of subroutes \((u_1, \ldots, u_k)\) do
7: Remove \( r \) from INCOMPLETE.
8: if \( u_k \) is a terminating subroute then
9: Add \( r \) to \( R(\phi) \).
10: else
11: for all subroutes \( u \) where \( u \) belongs to original route \( r(\phi_u) \) for aircraft \( \phi_u \) do
12: if \( u.\text{orig} = u_k.\text{dest} \)
and \( \phi_u \) has same fleet type as \( \phi \)
and \( \delta_{u_k} > \alpha_u + \gamma \) then
13: Add a new route \((u_1, \ldots, u_k, u)\) to INCOMPLETE.
14: end if
15: end for
16: end if
17: end for
18: end for
19: end
4.3 Objective Function and Constraints

Each airline, after generating route possibilities, solves an integer program in order to determine a route for each aircraft, and in order to assign flights to slots. The decision variables for the integer program are shown in Table 4.1. The objective function and constraints are shown in Fig. 4-2, and all other variables and parameters are shown in Table 4.2.

In the notation, \( \Phi \) represents the airline's disrupted aircraft, \( F \) represents the scheduled flights owned by the airline, and \( S \) represents the airline's slots. \( R \) represents the airline's non-slotted routes, and \( V \) represents its slotted routes.

The route assignment variable \( X_v \) equals 1 if slotted route \( v \) is assigned. The flight cancellation variable \( K_f \) equals 1 if flight \( f \) is cancelled. The slot cancellation variable \( K_s \) equals 1 if slot \( s \) is unused; although \( K_s \) does not appear in the objective function, it is useful for later determining which slots were unassigned in the solution.

The objective function (Eq. 4.1) seeks to minimize the total number of passenger minutes of delay resulting from any particular set of aircraft routings. It is assumed that all costs are due to passenger delay—any other costs are ignored. It is also assumed that there is a fixed cost per minute of passenger delay for all airlines. If these assumptions are true and airlines seek to minimize cost, then minimizing cost is equivalent to minimizing passenger delay, and the objective function is correct. In reality, airlines are likely to consider other components of cost.

The total amount of passenger delay is equal to the sum of the passenger delays for all the flights that are flown but delayed \( (C_v X_v) \), plus the sum of the passenger delays for all the flights that are cancelled \( (d_f K_f) \). \( C_v \) equals the total passenger delay incurred by flights \( (f_1, \ldots, f_n) \) in slotted route \( v \) (Eq. 4.2). \( d_f \) is the total passenger delay incurred by cancelling flight \( f \).

The first constraint (Eq. 4.4) ensures that every aircraft is assigned exactly one routing. The second constraint (Eq. 4.5) requires that every flight is either cancelled or flown on some route. The third constraint (Eq. 4.6) ensures that every arrival slot either goes unassigned or is assigned to a flight in exactly one route.
minimize
\[ \sum_{v \in V} C_v X_v + \sum_{f \in F} d_f K_f \]

where
\[ C_v = \sum_{f_i \in v} ((p_{f_i} - i p_i) D_{f_i} + m_{f_i}^v) \]

and
\[ D_{f_i} = \begin{cases} 
0 & \text{if } i \text{ is 1 and } f_i \text{ is not a GDP arrival,} \\
\max(\tau_s - \alpha_f, 0) & \text{if } f_i \text{ is a GDP arrival and is assigned to slot } s \text{ in } v, \\
\max(D_{f_{i-1}} - \delta_f + \alpha_{f_{i-1}} + \gamma, 0) & \text{otherwise.}
\end{cases} \]

for \( f_i \) in some slotted route \( v \) composed of flights \( (f_1, \ldots, f_n) \).

subject to
\[ \sum_{v \in V(\phi)} X_v = 1, \ \forall \phi \in \Phi \]  
\[ \sum_{v \ni f} X_v + K_f = 1, \ \forall f \in F \]  
\[ \sum_{v \in V(s)} X_v + K_s = 1, \ \forall s \in S \]  
\[ X_v, K_f, K_s \in \{0, 1\} \ \forall v \in V, f \in F, s \in S \]

Figure 4-2: Airline Recovery Optimization Problem

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
</table>
| \( X_v = \) | \begin{cases} 
1 & \text{if slotted route } r \text{ is assigned,} \\
0 & \text{if unassigned}
\end{cases} | route assignment variable |
| \( K_f = \) | \begin{cases} 
0 & \text{if flight } f \text{ is flown,} \\
1 & \text{if cancelled}
\end{cases} | flight cancellation variable |
| \( K_s = \) | \begin{cases} 
0 & \text{if slot } s \text{ is assigned,} \\
1 & \text{if unassigned}
\end{cases} | slot cancellation variable |

Table 4.1: Decision Variables
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Phi$</td>
<td>set of all disrupted aircraft</td>
</tr>
<tr>
<td>$F$</td>
<td>set of all flights</td>
</tr>
<tr>
<td>$S$</td>
<td>set of all slots</td>
</tr>
<tr>
<td>$R$</td>
<td>set of all possible non-slotted routes for all disrupted aircraft</td>
</tr>
<tr>
<td>$R(\phi)$</td>
<td>set of all possible non-slotted routes for disrupted aircraft $\phi$</td>
</tr>
<tr>
<td>$V$</td>
<td>set of all possible slotted routes for all disrupted aircraft</td>
</tr>
<tr>
<td>$V(\phi)$</td>
<td>set of all possible slotted routes for disrupted aircraft $\phi$</td>
</tr>
<tr>
<td>$V(s)$</td>
<td>set of all slotted routes containing a flight assigned to slot $s$</td>
</tr>
<tr>
<td>$D_f$</td>
<td>delay for flight $f$ (in minutes)</td>
</tr>
<tr>
<td>$p_f$</td>
<td>number of passengers with seats on flight $f$</td>
</tr>
<tr>
<td>$ip_f^v$</td>
<td>number of passengers on flight $f$ who miss their connections in slotted route $v$</td>
</tr>
<tr>
<td>$m_f^v$</td>
<td>total passenger minutes of delay due to passengers on flight $f$ who would miss their connections in slotted route $v$</td>
</tr>
<tr>
<td>$d_f$</td>
<td>total passenger minutes of delay equivalent to cancelling flight $f$</td>
</tr>
<tr>
<td>$\tau_s$</td>
<td>arrival slot time for slot $s$</td>
</tr>
<tr>
<td>$\alpha_f$</td>
<td>originally scheduled arrival time for flight $f$</td>
</tr>
<tr>
<td>$\delta_f$</td>
<td>originally scheduled departure time for flight $f$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>minimum turnaround time between the arrival of one flight leg and the departure of the next flight leg, for any aircraft</td>
</tr>
<tr>
<td>$\nu$</td>
<td>absolute minimum time between any two flights $f_1$ and $f_2$, such that it is possible for a passenger to connect from $f_1$ to $f_2$</td>
</tr>
<tr>
<td>$\rho$</td>
<td>maximum delay to a passenger due to cancelling that passenger’s flight</td>
</tr>
</tbody>
</table>

Table 4.2: Optimization Parameters
4.3.1 Precomputing Route Delay Values for Slotted Routes

In order to precompute the route delay for a particular route \( r \) generated in 4.2.2, it is necessary to know the arrival slot assignments for each of the GDP arrival flights within \( r \). This is accomplished by generating a set of slotted routes from \( r \) by pairing each GDP arrival in \( r \) with some arrival slot.

The total number of passenger minutes of delay \( (C_v) \) for a particular slotted route equals the sum of the delays for all flights in the route (Eq. 4.2). The passenger delay for a particular flight \( f_i \) in \( v \) is equal to the total passenger delay incurred for passengers who miss their connections as a result of assigning route \( v \) \( (m^v_{f_i}) \), plus the total passenger delay incurred for all passengers who are merely delayed and not interrupted \( ((p^v_{f_i} - ip^v_{f_i})D_{f_i}) \). \( D_{f_i} \) represents the number of minutes that flight \( f_i \) is delayed.

The delay \( m^v_{f_i} \) for passengers who miss their connections is calculated as described in 4.3.3. The individual flight delays \( D_{f_i} \) are determined using Eq. 4.3, which guarantees that the following requirements are satisfied:

1. If a GDP arrival flight \( f \) is assigned to a slot \( s \), the flight arrives at the designated slot time. Therefore, the amount of delay for \( f \) is equal to the difference between the slot time \( \tau_s \) and \( f \)'s originally scheduled arrival time \( \alpha_f \).

2. Delays are propagated along a route, so that delaying one flight leg \( f_{i-1} \) causes the succeeding flight leg \( f_i \) to be delayed as well. In other words, the new departure time for the second flight, \( \delta_f + D_{f_i} \), must be no earlier than the new arrival time for the first flight, \( \alpha_{f_{i-1}} + D_{f_{i-1}} \), plus the minimum turnaround time \( \gamma \).

When calculating flight delays, it is assumed that the in-flight time for any particular flight \( f \) from its origin to its destination is constant. It is also assumed that flights may not take off before their originally scheduled departure times \( \delta_f \); this excludes the possibility of negative flight delays. Also, it is assumed that the minimum turnaround time \( \gamma \) between successive flight legs is also the maximum turnaround
time—that is, the possibility that an aircraft may have to remain on the ground for a longer period, for example to be serviced, is ignored.

Note that the second requirement above excludes certain slotted route possibilities. Assigning a slot to a flight early in a route might cause enough propagated delay to prevent a flight occurring later in the same route from arriving at a particular slot time.

The other requirements for slotted routes are that a slot may not be assigned to a flight originally scheduled to arrive after the time of the slot, and that slots may not be assigned to multiple GDP arrivals within the same route.

To generate slotted routes quickly, a data structure mapping GDP arrival flights to eligible arrival slots is first constructed. A slot $s$ is eligible to be assigned to a particular flight $f$ if the flight’s original arrival time $\alpha_f$ is no earlier than the slot time $T$.

The complete algorithm is shown as Alg 4.2. For a particular route $r$, a list INCOMPLETE of partially completed slotted routes is maintained. Each GDP arrival flight $f$ in $r$ is considered in order, and the set of eligible slots $S(f)$ is obtained. For each partially completed slotted route $v$ in INCOMPLETE, $v$ is replaced with the set of slotted routes formed by assigning $f$ to each possible arrival slot $s$ in $S(f)$. $f$ can only be assigned to $s$ if $s$ is not already assigned to another flight in $v$, and if the assignment of $f$ to $s$ would not require that the flight delay $D_f$ be less than the propagated delay implied by the delay of the previous flight in $v$.

The algorithm ends when all GDP arrival flights in $r$ have been considered for assignment, and therefore, all possible slotted routes have been generated.

### 4.3.2 Estimating Flight Cancellation Delays

For the objective function, it is necessary to determine the total number of minutes of passenger delay $d_f$ that would result if flight $f$ were cancelled.

In reality, when a flight is cancelled, affected passengers either find a seat on a later flight with the same destination, connect to their destination through another airport, find a flight on a different airline, or decide not to fly at all. In the first two
Algorithm 4.2 Generating Slotted Routes and Computing Delay Values

1: begin
2: For each flight \( f \), construct the set of eligible arrival slots \( S(f) \).
3: for all aircraft \( \phi \) in \( \Phi \) do
4: \( V(\phi) := \emptyset \).
5: for all routes \( r \) in \( R(\phi) \) where \( r = (f_1, \ldots, f_n) \) do
6: \( \text{INCOMPLETE} := \emptyset \).
7: Create a new slotted route \( v_0 \) from \( r \), with no flights assigned to slots.
8: Add \( v_0 \) to \( \text{INCOMPLETE} \).
9: for all \( f_i \) in \( r \) do
10: \( \text{INCOMPLETE}' = \emptyset \).
11: for all slotted routes \( v \) in \( \text{INCOMPLETE} \) do
12: if \( f_i \) is not a GDP arrival then
13: if \( i = 1 \) then
14: \( D_{f_i}^v := 0 \).
15: else
16: \( D_{f_i}^v := \max(D_{f_{i-1}} - \delta_{f_i}, \alpha_{f_{i-1}} + \gamma, 0) \).
17: end if
18: Add \( v \) to \( \text{INCOMPLETE}' \).
19: else
20: for all slots \( s \) in \( S(f_i) \) s.t. no flight in \( v \) is assigned to \( s \) do
21: if \( i > 1 \) then
22: \( p := \max(D_{f_{i-1}} - \delta_{f_i}, \alpha_{f_{i-1}} + \gamma, 0) \).
23: \( q := \max(\tau_s - \alpha_f, 0) \).
24: if \( p \leq q \) then
25: Let \( v' \) be a copy of \( v \).
26: Let \( f_i \) be assigned to \( s \) in \( v' \).
27: \( D_{f_i}^{v'} := \max(\tau_s - \alpha_f, 0) \).
28: Add \( v' \) to \( \text{INCOMPLETE}' \).
29: end if
30: else
31: Let \( v' \) be a copy of \( v \).
32: Let \( f_i \) be assigned to \( s \) in \( v' \).
33: \( D_{f_i}^{v'} := \max(\tau_s - \alpha_f, 0) \).
34: Add \( v' \) to \( \text{INCOMPLETE}' \).
35: end if
36: end for
37: end if
38: end for
39: \( \text{INCOMPLETE} := \text{INCOMPLETE}' \).
40: end for
41: Add all slotted routes \( v \) in \( \text{INCOMPLETE} \) to \( V(\phi) \).
42: end for
43: end for
44: end
cases, passengers are delayed by some finite number of minutes, and the delays can be calculated by rerouting individual passengers on later flights. However, since it is unknown at the time of the optimization which other flights are cancelled and which are delayed, it is impossible to calculate exactly the total delay that would result by rerouting passengers to other particular flights. Instead, the flight cancellation delays are estimated by rerouting passengers according to the original schedule.

The latter two cases can be characterized as instances where the passenger delay to the airline is immeasurable; however, the true cost to the airline of cancelling those passengers' seats is not infinitely large, but rather bounded by some amount. For example, the airline might have to refund the airfare for the affected passengers, and it might also lose profits if those passengers decide not to patronize the airline in the future.

Accordingly, it was assumed that the delay to a single passenger from cancelling that passenger's flight is bounded above by some fixed amount $\rho$, in order to prevent instances where cancelling flights forced passengers to bear too large an amount of delay. It is reasonable to assume that a passenger will wait up to some length of time before either switching to a different airline or deciding not to fly.

The cutoff was assumed to be six hours, as in [11]. Because the value of the cutoff is inexact and somewhat subjective, the sensitivity of the results to this assumption was later tested by varying the cutoff value.

In the simulation, $d_f$ was calculated for each flight $f$ by determining what the delay would be if passengers were rerouted to later flights with available seats. Passengers with connections from $f$ were rerouted first, and then terminating passengers were rerouted. $d_f$ was estimated by filling empty seats on available flights until all bumped passengers were accounted for.

To reroute connecting passengers to later flights, it was first attempted to place them on direct flights $f'$ from $f$'s origin to their final destinations. Only flights $f'$ that departed after $f$ ($\delta_{f'} \geq \delta_f$) and arrived within the maximum delay period ($\alpha_{f'} - \alpha_f \leq \rho$) were considered.

If some connecting passengers for the cancelled flight could not be rerouted on
direct flights, they were rerouted on later connections through their original hub. Such passengers were bumped onto later flights $f_1$ from $f$'s origin to $f$'s destination (i.e. to the hub airport), and also onto connecting flights $f_2$ from the hub airport to their final destinations. Only flight pairs $f_1$ and $f_2$ such that $f_1$ departed after $f$ ($\delta_{f_1} \geq \delta_f$), $f_2$ arrived within the maximum delay period ($\alpha_{f_2} - \alpha_f \leq \rho$), and enough time existed to connect between $f_1$ and $f_2$ ($\delta_{f_2} \geq \alpha_{f_1} + \nu$, where $\nu$ is the absolute minimum time that would allow passengers to connect between flights), were considered.

Any non-connecting passengers on $f$ were rerouted on later flights $f'$ from $f$'s origin to $f$'s destination. Flights $f'$ that were considered were those that departed after $f$ ($\delta_{f'} \geq \delta_f$) and arrived within the maximum delay period ($\alpha_{f'} - \alpha_f \leq \rho$).

Any connecting or terminating passengers who were unsuccessfully rerouted as above were assumed to be delayed by the maximum delay cutoff $\rho$.

For example, if there are 100 passengers flying direct on flight $f$, and 40 available seats on each of the next three flights to the same destination, and if each of the four flights is an hour apart, then the total amount of passenger delay $d_f$ would be $40 \times 60 + 40 \times 120 + 20 \times 180 = 10800$ passenger minutes of delay, since the first 40 passengers would be delayed an hour, the next 40 passengers would be delayed 2 hours, and the last 20 passengers would be delayed 3 hours. If instead 10 of the passengers were making a connection, and could be rerouted on a direct flight an hour after the first flight, the total delay would be $10 \times 60 + 40 \times 60 + 40 \times 120 + 10 \times 180 = 9600$ passenger minutes of delay.

Note that this procedure is just an estimate of the actual delay, since it relies not only on the assumption that the three later flights take off on time, but also on the assumption that there are no other cancelled flights whose bumped passengers occupy the available seats on the three flights.

### 4.3.3 Estimating Delays for Interrupted Connections

The procedure used to estimate $m^\gamma_f$, the total number of passenger minutes of delay due to passengers missing their connections on flight $f$ in slotted route $\nu$, is very similar to the procedure used to determine $d_f$, the delay due to cancelling flight $f$. 

58
The delay $D_f$ of a flight $f$ determines which connections from $f$ are interrupted. A connection is interrupted if there is not enough time to connect from $f$ to the connecting flight $g$, i.e. if $\delta_g - (\alpha_f + D_f) \leq \nu$.

Any interrupted connecting passengers are assumed to be rerouted onto new connecting flights $g'$ with the same itinerary as $g$, that depart after $g$ ($\delta_{g'} \geq \delta_g$), and that arrive within the maximum delay period ($\alpha_{g'} - \alpha_g \leq \rho$). Any connecting passengers who are unsuccessfully rerouted this way are assumed to be delayed by the maximum delay cutoff $\rho$.

Just as in the calculation of $d_f$, using this procedure of rerouting passengers to calculate $m'_f$ is strictly an estimation. The number of available seats on any flight may actually be affected by passengers who are bumped from other cancelled or delayed flights not taken into account.

### 4.4 Modeling Considerations

A few considerations are worth noting.

First, slotted routes and delay values were pregenerated in order to speed the linear optimization. However, the technique is subject to memory requirements; there is no limit to the number of possible slotted routes, and this number can skyrocket for large hub airports with extended GDPs. In the scenarios tested in this thesis, memory was not a limiting factor.

Second, due to the way alternative routes are constructed, the optimization problem posed in 4.3 is not necessarily always feasible. Because flights are combined into subroutes, it is not always possible to cancel any arbitrary flight $f$.

In particular, if an airline decides to sell a slot in the auction, there may be too few slots to satisfy all of the constraints. To prevent this, the airline can check whether the corresponding optimization problem would be infeasible when setting its reservation price. If so, it can set the reservation price to a high value ($+\infty$).

Finally, it is possible that the turnaround time between two flight legs in an original route $r$ (as calculated from the original schedule data) violates the minimum
turnaround time requirement, either because the data is poor, or because of exceptionally low turnaround times. However, any such flights become delayed when slotted routes are generated from $r$, so that the minimum turnaround requirement is obeyed in the slotted routes.
Chapter 5

Data Reconstruction and Generation

In order to obtain accurate passenger delay values for each of the three slot allocation methods, historical flight schedule data and passenger itinerary information were used to reconstruct aircraft routes and passenger itineraries. Actual routes for aircraft were recreated from databases containing detailed schedule information. Because per-flight passenger itinerary data was unavailable, passenger itineraries were stochastically generated using statistics derived from aggregate itinerary data.

5.1 Reconstruction of Schedule Data

Original flight schedules were reconstructed from the Airline Service Quality Performance (ASQP) database of flight data for major US airlines. Actual flown routes were constructed from the ASQP data by tracking the tail numbers of the aircraft that actually flew each flight.

Because the ASQP database does not include data for small airlines and international airlines, the Official Airline Guide (OAG) database of flight schedules was used to reconstruct flight schedules for those airlines not in ASQP. Because the OAG database does not contain the tail numbers of the aircraft assigned to each flight, routes for these airlines were estimated according to arrival and departure times of
particular flights.

5.1.1 ASQP Flight Reconstruction

For each flight, the ASQP database contains information on the airline operating the flight, the scheduled departure date of the flight, the origin and destination airport codes, the scheduled departure and arrival times, and the actual departure and arrival times. Additionally, the tail number of the aircraft that actually flew the flight is listed. All of this information was used to reconstruct the set of flights $F$.

Scheduled and actual arrival times are given in the local time zone of the flight’s origin airport. Likewise, departure times are given in the local time zone of the flight’s destination airport. Because it was necessary to piece several flights together in order to form a route, all times were first converted to a standard time zone using a mapping of airport codes to time zone offsets from Greenwich Mean Time (GMT). Also, to ensure that all flights within a continuous 24-hour period were correctly reconstructed, three days worth of data was parsed, even though routes were constructed only from flights departing and arriving within the second day.

The ASQP database does not list the actual departure date for each flight, so the actual departure date was assumed to be the same as the scheduled departure date. Also, since the actual arrival dates were not listed, the arrival date was initially assumed to be the same as the scheduled departure date, and then corrected if necessary so that the arrival time was within 24 hours after the departure time.

The fleet type of the aircraft flying a particular flight was determined from the tail number using a publicly available aircraft registry [12]. The capacity, or the number of seats available on the flight, was determined from the aircraft type according to aircraft descriptions listed online at www.landings.com.

5.1.2 ASQP Route Reconstruction

The original routes for aircraft in the ASQP schedule data were constructed according to the tail numbers listed for each flight.
First, the set of flights $F$ was sorted by scheduled departure date and time. Then, each flight was added to the original route for the aircraft indicated by the flight’s tail number. A flight $f$ was successfully added to a route $r$ if $f$ departed from the destination of the last flight in $r$. If an attempt to add a flight to a route was unsuccessful, the data for the entire route was discarded.

Sorting the flights ensures that they are added to routes in the correct order. It would actually be more accurate to sort flights by the *actual* departure, rather than by the *scheduled* departure. However, because actual departure dates are not listed in the ASQP data, it is necessary to sort by scheduled departure times in order to avoid adding flights in the wrong order.

On the other hand, all route delay calculations described in Chapter 4 were performed using the *actual* departure dates and times, because the routes themselves are the actual routes as constructed from aircraft tail numbers, rather than the originally scheduled routes. If the actual departure time of any particular flight was for any reason incorrect due to an anomaly in the ASQP dataset, the flight would have been delayed appropriately during the construction of slotted routes, as described in 4.3.1.

In the 24-hour period examined, 81 out of 2943 routes, or 2.7%, were discarded due to unsuccessful route construction.

### 5.1.3 Anomalies in ASQP Data

A few anomalies in the data prevented all flights from being reconstructed. One problem was that for a very small number of flights, the tail number was unknown or unreported. Such flights were discarded, because it was impossible to reconstruct the original route containing that flight.

The other major problem with the ASQP data was that the listing of the scheduled departure date for a particular flight is sometimes incorrect. In particular, it was determined that when several flights flown in sequence by the same aircraft are assigned the same flight number, the departure date listed in the ASQP data actually corresponds to the scheduled departure date of the *first* flight in the sequence, rather than the departure date of each individual flight.
This issue affects red eye flights in particular. For example, an aircraft might take off from BOS at 11pm on Jan. 1, fly to EWR and land at 12:30am, then continue to MCO and land at 4am. If both the BOS → EWR and EWR → MCO segments were assigned the same flight number, then both segments would be listed as departing on Jan. 1, even though the second segment actually departed on Jan. 2.

Most of these red eye flights were discarded as a consequence of route construction because the flights for a particular tailnumber are ordered incorrectly, flow balance is not satisfied, and route construction fails. In a few cases, route construction may not fail; for example, if there is another red eye flight the following night assigned to the same aircraft, the route can be successfully constructed.

5.1.4 OAG Reconstruction

Because the ASQP database includes data for only the 10 major domestic US airlines with revenues above $10 billion, it was necessary to create additional routings for small and international airlines by reconstructing flights from the OAG database, which contains schedule data for all airlines. Reconstructing flights for airlines not in ASQP is straightforward; for each flight, the OAG database includes information on departure and arrival times (in GMT), the aircraft type, and the number of seats on the aircraft.

However, the OAG database does not contain tail numbers for each flight. Therefore, the exact aircraft routings were guessed rather than accurately reconstructed. The operating assumption was that most routes for the local and international airlines consist of back and forth flights between two airports. For example, in March 1998, Business Express, a commuter airline, operated several shuttles a day between Logan International Airport (BOS) and several surrounding airports, including Bangor, ME (BGR). British Airways, an international airline, operated only a few flights a day from Logan, and all were either to or from London’s Heathrow Airport (LHR).

Using this assumption, a list of aircraft $L$ was maintained for each airline and pair of airports. $L$ contains all aircraft with routes consisting of back-and-forth flights between the two airports. Routings were constructed by examining, in scheduled
order, each OAG flight $f$ arriving at or departing from the GDP airport, and then adding $f$ to the original route of some aircraft in the appropriate list $L$. The procedure is essentially a greedy algorithm—it tries to serve all of the flights in the schedule using the fewest aircraft necessary.

A flight $f$ was successfully added to a route $r$ if $f$ departed from the destination of the previous flight in $r$, and if the turnaround time between flight legs exceeded 20 minutes. If it was not possible to add the flight to any route in $L$, a new aircraft was allocated to $L$ and assigned to fly $f$.

5.2 Stochastic Generation of Passenger Data

In order to determine total passenger delay, it is necessary to know how many passengers are on each flight, as well as the itineraries for passengers connecting to other flights. The number of passengers on a flight was stochastically generated from aggregate statistics on passenger load factors. The itineraries of each of the passengers on the flight were also randomly generated, based on the known probability of a passenger flying a particular itinerary.

5.2.1 Determination of Passenger Loads

The passenger load factor of an airline, or percentage of seating that is utilized, is commonly defined as the airline’s total revenue passenger miles divided by its available seat miles. A revenue passenger mile is equivalent to transporting one paying passenger one mile, while an available seat mile is equivalent to transporting one available seat one mile. [1]

While data on passenger load factors for large and medium size airlines is obtainable from data reported on FAA Form 41, the data for small and commuter airlines as reported on FAA Form 298C is currently unreliable [9], and implies unrealistic load factors as low as 25% for some small airlines. Therefore, per-airline passenger load factors were not used to determine the number of filled seats on each flight.

Instead, an industry-wide figure was used to generate the number of filled seats
for all airlines’ flights. The industry-wide passenger load factor was 70%, according to traffic and operations data for 1998 [1]. The number of passengers on a particular flight was generated using a normal distribution with mean 70% of the aircraft’s capacity, and standard deviation 25% of the aircraft’s capacity. This distribution was truncated so that the number of passengers on a flight was no more than 100% of the aircraft’s capacity, and no less than 0%.

5.2.2 Creation of Passenger Connections

As described in Chapter 4, the amount of delay to a single passenger is equal to his delay in reaching his final destination, which is not necessarily the destination of the delayed flight. Most airlines operate hub-and-spoke networks in order to aggregate passengers from many origins flying to a single destination, by first shuttling them through a single airport. To simulate this, connections were created for some percentage of the passengers on the flight, after the number of passengers on a flight was determined as above.

Connections were stochastically generated using itinerary probabilities calculated using the Airline Origin and Destination Survey Market (DB1B) Market database, a 10 percent sample of airline tickets from major airlines available from the Bureau of Transportation Statistics. For each ticket sampled, this database contains the number of passengers on the ticket, as well as the passengers’ itinerary in the form of the sequence of airports visited. This data was used to compute the total number of passengers that flew each itinerary, in order to later create a probability distribution over itineraries. Additionally, the data was used to create a probability distribution over passenger group sizes, by computing the total number of tickets sold for each group size.

To create the connections, each flight $f$ was considered in scheduled order. The candidate connecting flights from a flight $f$ flying from airport A to airport B were determined to be those that departed from B within some time window after $f$’s arrival. In this thesis, that time window includes all flights originally scheduled to depart between 30 minutes and three hours after $f$’s arrival.
In Fig. 5-1, Flights 3, 4, 5, 6, and 7 are considered candidate connecting flights from $f$, since they depart within the indicated time interval after $f$'s arrival at time $t$. Flights 1 and 2 depart too early for passengers to connect from $f$, and flights 8 and 9 depart too late.

For simplicity, passengers were assumed to connect at most once. Thus, all passengers on flight $f$ originating at A either terminated at B (itinerary $A \to B$), or continued on one of the candidate connecting flights from B to some airport C (itinerary $A \to B \to C$).

The total number of passengers $I_i$ that flew each possible itinerary $i$ was determined according to the DB1B Market data. Each candidate itinerary $i$ was then assigned a probability $p_i = \frac{I_i}{\sum_{j=1}^{i} I_j}$, where $n$ is the number of candidate itineraries.

After the probabilities were determined, the actual itineraries were determined for each of the passengers on flight $f$ originating at A. (Because connections are created in scheduled order and passengers can connect to $f$ from some other flight, some number of filled seats on $f$ has already been accounted for.) First, the size of the group was randomly generated according to the probability distribution over passenger group sizes. Then, the itinerary of the group was determined using the probability distribution over itineraries. This process was repeated until all filled
seats on $f$ were accounted for.

Fig. 5-2 illustrates the result of filling an example 30-passenger plane. First, the number of filled seats was determined to be 23; the gray boxes represent the seven empty seats. Itineraries were then generated for all seated passengers.

The five striped boxes indicate that the first five filled seats were allocated to passengers connecting from earlier flights. Connections were then generated for a three-passenger group connecting to flight 6, a single passenger connecting to flight 3, a two-passenger group connecting to flight 5, a single passenger connecting to flight 7, a single passenger connecting to flight 6, a pair of terminating passengers, a single passenger connecting to flight 4, a single passenger connecting to flight 7, a two-passenger group connecting to flight 5, two passengers connecting to flight 3, and two terminating passengers.

Note that connections were created only for flights obtained from the ASQP data; it was assumed that all passengers on local and international airlines flew nonstop to their destinations.
Chapter 6

Results and Analysis

The algorithms and models described in previous chapters were simulated on historical flight data in order to quantitatively analyze the differences between the various slot allocation methods. Details of the implementation are summarized in 6.1, and the results from comparing the three slot allocation methods are described in 6.2. The sensitivity of the objective function to an assumed parameter, the maximum delay for passengers rerouted from cancelled flights, is analyzed in 6.3. The results of simulating alternative airline behaviors for both the auction and CDM slot allocation methods are presented in 6.4. Finally, 6.5 analyzes the running time of the airline recovery optimization model described in Chapter 4.

6.1 Implementation Notes

All three arrival slot allocation methods were implemented in Java and simulated on ASQP and OAG flight data from March 1-3, 1998, and DB1B Market data from the first quarter of 1998. The Java interface to DashOptimization’s XpressMP software was used to compute and solve the integer programming (IP) formulation of the airline recovery problem. XpressMP solves the integer program by first solving the linear programming (LP) relaxation of the IP problem, and then performing a branch-and-bound search on the LP solution to determine integer values for the decision variables.

The simulation was performed on a 2.20 GHz Pentium 4 processor with 1GB of
RAM running Red Hat Linux 7.3. The maximum Java heap size was specified to be 600MB.

The minimum slack time between flights in a route, $\gamma$, was assumed to be 25 minutes. The minimum time allowed for passengers to connect between flights, $\nu$, was assumed to be 20 minutes. The maximum passenger delay due to a cancelled flight, $\rho$, was assumed to be six hours.

### 6.2 Comparison of Slot Allocation Methods

The three slot allocation methods—CDM, the global optimization, and the auction—were simulated for a few different GDP scenarios. In all scenarios, the GDP airport was Boston’s Logan International Airport (BOS), and the default Airport Arrival Rate (AAR) was assumed to be 60 arrivals per hour. Each GDP is specified by a period of reduced arrival capacity, and the AAR for the reduced period.

The results of the simulation are shown in Tables 6.1 and 6.2. Each scenario was simulated at least fifteen times, and the results of all runs for a particular scenario were averaged.

Table 6.1 shows, for each GDP scenario, the average percentage of reducible passenger delay captured by the auction. The amount of reducible delay is calculated as the difference between the total passenger delay incurred under CDM and the total passenger delay incurred under the global optimization method. The percent of reducible delay captured by the auction is calculated as the difference between the delay resulting under CDM and the delay resulting under the auction method, divided by

<table>
<thead>
<tr>
<th>Reduced AAR (arrivals/hr)</th>
<th>Reduction Period (hrs)</th>
<th>% of Reducible Delay Captured by Auction</th>
<th>Std. Dev.</th>
<th>Number of Runs</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>74.51%</td>
<td>9.54%</td>
<td>23</td>
<td>37.46</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>69.48%</td>
<td>11.78%</td>
<td>15</td>
<td>22.84</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>22.42%</td>
<td>21.37%</td>
<td>16</td>
<td>4.20</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>56.64%</td>
<td>33.29%</td>
<td>19</td>
<td>7.42</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>36.25%</td>
<td>18.77%</td>
<td>15</td>
<td>7.48</td>
</tr>
</tbody>
</table>

Table 6.1: Percentage of Reducible Delay Captured by Auction, for Various GDPs
the amount of reducible delay. The table also shows the standard deviation for the percentage captured, the number of times the simulation was run for the particular scenario, and a $t$ statistic calculated for each set of runs.

The results are quite varied. For a GDP that completely shuts down BOS for two hours, 74.51\% of the reducible delay was captured by the auction. On the other hand, for a GDP that reduces the arrival rate to 20 arrivals per hour for those two hours, only 22.42\% of the reducible delay was captured. The data indicate that for a fixed reduction period, a greater fraction of the reducible delay is captured by the auction if the reduction in arrival rate is more severe.

The data also indicate that for a fixed reduced arrival rate, a greater fraction of the reducible delay is captured if the reduction period is longer. For a GDP that shuts down BOS for only one hour, 56.64\% of the reducible delay was captured by the auction; this is lower than 74.51\%, the percentage captured for the two-hour shutdown.

Also, for less severe GDP scenarios, the standard deviation is fairly high. For example, in the case of the two-hour reduction to 20 arrivals per hour, the standard deviation was 21.37\%; in some simulation runs for that scenario, very little if any delay was reduced through the auction mechanism. However, the delay incurred under the auction mechanism did not exceed the delay incurred under CDM, for any simulation run.

Given the high standard deviations, it is important to test the possibility that the percentage of reducible delay captured by the auction is not statistically significant. This null hypothesis was tested using a $t$-test. The $t$ statistic was calculated as $\frac{\bar{X}}{s_X/\sqrt{N}}$, where $\bar{X}$ is the mean, $s_X$ is the standard deviation, and $N$ is the number of runs. Since each GDP scenario contains at least 15 sample points, there are 14 degrees of freedom in the $t$ distribution. The critical value of the $t$ distribution with 14 degrees of freedom, at the 1\% significance level, is 2.977.

Because $t > 2.977$ for each GDP scenario, the null hypothesis can be rejected at the 1\% significance level. That is, the probability that the reduction in delay is greater than zero is at least 99\%, for each GDP scenario tested.
Table 6.2 shows, for each GDP scenario simulated, the average amount of passenger delay reduced through the auction allocation method. This is calculated as the difference between the total passenger delay resulting under CDM, and the delay resulting under the auction method. The table also lists the standard deviation for the amount of delay reduced, the percentage reduction in delay, the standard deviation for the percentage reduction, the number of simulation runs, and the $t$ statistic for the percentage reduction.

The results indicate that for the GDP scenarios simulated, there is a wide range of reductions in delay, and also a high variance within each scenario for the amount of delay reduced. For a two-hour reduction to 10 arrivals per hour, the delay reduction achieved through the auction was on average 28.80% of the CDM passenger delay. On the other hand, for a one-hour shutdown of BOS, only 5.93% of the CDM passenger delay was reduced through the auction.

Also, the $t$ statistics are all greater than 2.977. Therefore, the null hypothesis that there is zero reduction in delay can be rejected at the 1% significance level.

The actual amount of reduced passenger delay is highly dependent on the particular GDP scenario, and the numbers in Table 6.2 are provided only for context; the true measure of the effectiveness of the auction is the percentage of reducible delay it is able to capture. For extended GDPs, the percentage delay reduced through the auction may be very small, simply because there is a fixed amount of delay that is impossible to reduce. For example, if an airport is completely shut down for four hours due to a snowstorm, those four hours are irrecoverable.

<table>
<thead>
<tr>
<th>Reduced AAR (arr/hr)</th>
<th>Reduction Period (hrs)</th>
<th>Delay Reduced (passenger minutes)</th>
<th>Std. Dev.</th>
<th>Percent Reduced</th>
<th>Std. Dev.</th>
<th>Num. Runs</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>42,757</td>
<td>30,548</td>
<td>6.84%</td>
<td>4.30%</td>
<td>23</td>
<td>7.63</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>61,420</td>
<td>23,493</td>
<td>28.80%</td>
<td>10.86%</td>
<td>15</td>
<td>10.27</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>5,023</td>
<td>5,363</td>
<td>8.83%</td>
<td>10.2%</td>
<td>16</td>
<td>3.46</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>7,448</td>
<td>5,400</td>
<td>5.93%</td>
<td>4.63%</td>
<td>19</td>
<td>5.58</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>17,986</td>
<td>10,604</td>
<td>20.48%</td>
<td>12.27%</td>
<td>15</td>
<td>6.46</td>
</tr>
</tbody>
</table>

Table 6.2: CDM Passenger Delay Reduced Using Auction, for Various GDPs
Overall, the results show that the auction appears to be very effective at reducing passenger delay, even though there is a great deal of variation in its effectiveness. Under each of the scenarios tested, overall passenger delay was reduced by at least 5-7%, and the auction captured as much as 75% of the reducible delay.

### 6.3 Sensitivity of Results to Delay Cutoff

It was assumed in this thesis that the delay to a single passenger from cancelling that passenger's flight is bounded above by a fixed amount \( \rho \), as described in 4.3.2. The sensitivity of the results to this delay cutoff were tested by running the simulation for a few different values of \( \rho \).

Fig. 6-1 shows the total passenger delay resulting from Ration-by-Schedule slot allocation, for various values of the delay cutoff. The results are for a single simulation of a two-hour reduction in the arrival rate at BOS to zero arrivals per hour. Other than the delay cutoff, all other factors are constant for all points on the plot—including the number of passengers on each flight, and the passenger itineraries.

The graph indicates that the passenger delay calculation is in fact highly sensitive to \( \rho \); increasing \( \rho \) from three hours to six hours increased the total calculated delay from 560,000 passenger minutes to 640,000 passenger minutes. On the other hand, for values of \( \rho \) above 12 hours, the total passenger delay seems to become less sensitive to \( \rho \), and levels off at approximately 750,000 passenger minutes of delay.

Since the calculations are highly sensitive to \( \rho \), the simulations from 6.2 were repeated for various values of \( \rho \). Tables 6.3 and 6.4 show how changing \( \rho \) affects the amount of delay reduced through the auction method. In all cases, the GDP scenario simulated was for a two-hour arrival rate reduction at BOS, to zero arrivals per hour. The simulation was run at least fifteen times for each value of \( \rho \), and the results of all runs for a particular value of \( \rho \) were averaged.

Table 6.3 shows, for each value of the delay cutoff \( \rho \), the average percentage of reducible passenger delay captured by the auction. For a delay cutoff of six hours, the average percentage of reducible delay captured was 74.51%. For a cutoff of nine
hours, the average percentage captured was 77.28%, and for a cutoff of 12 hours, the average percentage captured was 68.55%. However, since the standard deviations are very large, the figures are not statistically different.

Table 6.4 shows, for each value of $p$, the average amount of passenger delay reduced under the auction, and also the reduced delay as a percentage of the total CDM passenger delay. For a delay cutoff of six hours, 6.84% of total delay was reduced. The percentage reduction was 9.15% for a delay cutoff of nine hours, and 5.56% for a cutoff of 12 hours. However, as before, the standard deviations are large, and the figures are not statistically different.

Thus, the calculation of passenger delays is highly sensitive to the cutoff value for the maximum delay to a single passenger from cancelling his flight. However, increasing the maximum delay cutoff does not significantly affect the results described in 6.2. In particular, for the GDP scenario of shutting down BOS for two hours, approximately 70-80% of the reducible passenger delay is captured through the auction, and at least 5-7% of total passenger delay is reduced, for values of $p$ between six and 12 hours.

### 6.4 Analysis of Alternative Airline Behaviors

As described in 3.4.3, a few different airline behaviors were tested to determine whether airlines could benefit from alternative strategies, and to determine the resulting impact on the overall amount of passenger delay.

Specifically, airlines denoted as “cautious” were simulated as being less likely to release an unused slot under CDM, and more likely to set their reservation price to infinity under the auction. The caution level $p$ of an airline is the probability that it will withhold an unused slot in CDM, and the probability that it will set the reservation price of a slot to infinity in the auction.

Airlines designated as “predictive” were assigned a predicted final slot allocation based on the result of an initial auction. A predictive airline determines the value of a slot relative to the predicted final slot allocation, rather than to the slot allocation
Figure 6-1: Effect of Increasing Maximum Individual Passenger Delay on Total RBS Airline Delay

<table>
<thead>
<tr>
<th>Maximum Delay Cutoff (hrs)</th>
<th>% of Reducible Delay Captured by Auction</th>
<th>Std. Dev.</th>
<th>Number of Runs</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>74.51%</td>
<td>9.54%</td>
<td>23</td>
<td>37.46</td>
</tr>
<tr>
<td>9</td>
<td>77.28%</td>
<td>13.00%</td>
<td>20</td>
<td>26.59</td>
</tr>
<tr>
<td>12</td>
<td>68.55%</td>
<td>11.81%</td>
<td>15</td>
<td>22.48</td>
</tr>
</tbody>
</table>

Table 6.3: Percentage of Reducible Delay Captured by Auction, for Various Delay Cutoffs

<table>
<thead>
<tr>
<th>Maximum Delay Cutoff (hrs)</th>
<th>Delay Reduced (passenger min.)</th>
<th>Std. Dev.</th>
<th>Percent Reduced</th>
<th>Std. Dev.</th>
<th>Number of Runs</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>42,757</td>
<td>30,518</td>
<td>6.84%</td>
<td>4.30%</td>
<td>23</td>
<td>7.63</td>
</tr>
<tr>
<td>9</td>
<td>55,212</td>
<td>31,697</td>
<td>9.15%</td>
<td>4.67%</td>
<td>20</td>
<td>8.76</td>
</tr>
<tr>
<td>12</td>
<td>36,977</td>
<td>18,437</td>
<td>5.56%</td>
<td>2.27%</td>
<td>15</td>
<td>9.49</td>
</tr>
</tbody>
</table>

Table 6.4: CDM Passenger Delay Reduced Using Auction, for Various Delay Cutoffs
at the time the bid is placed.

Two types of tests were performed. In the first type, the behavior of a single airline was altered, and the effect of the airline’s altered behavior on its total passenger delay was determined. In the second type, the number of airlines with a specific type of behavior was varied, and the impact of increasing the number of airlines with that behavior was determined.

A representative airline, US Airways, was chosen to illustrate the effect of varying airline behavior. The results were verified by re-running the tests for other airlines.

6.4.1 Effect of Increasing Airline Cautiousness in the Auction

Under the auction mechanism proposed in this thesis, an airline assigned a slot through RBS has the option of keeping the slot under all circumstances; to do so, it can simply set its reservation price to a very high value. A relevant concern is that the auction may not be effective at reducing delay, if every airline decides to hold onto all of its slots in this way.

Fig. 6-2 shows the effect of increasing the caution probability \( p \) on the delay of a single airline. The results are for a single simulation run of a two-hour reduction in the arrival rate at BOS to zero arrivals per hour, where all factors other than the caution probability of the single airline were held constant. For each probability \( p \), the difference in the single airline’s delay from the baseline case \( (p = 0) \) is plotted.

The graph shows that as the airline becomes more cautious, its passenger delay actually falls. In this case, the reason for the reduced delay is that if the airline is more cautious, it ends up not selling certain slots that would increase its delay, but on which it would be able to make a profit.

A more accurate measure of the effect of increasing cautiousness would also take into account the income received by the airline through the auction. Since all bid and sell values in this thesis were determined in terms of passenger minutes of delay, income is defined as the total number of passenger minutes received through the
auction (from selling slots to winning bidders), minus the total number of passenger
minutes paid through the auction (from winning a slot from another airline). The
net passenger delay of the airline is equal to the passenger delay of the airline, minus
the income received through the auction.

Fig. 6-3 shows the results of the same simulation shown in Fig. 6-2, but in terms
of net passenger delays. The graph shows that increasing the caution level of the
airline in most cases increases the net passenger delay, but sometimes reduces it. The
net effect is therefore ambiguous, and actually somewhat random.

The reason for the randomness is that if an airline probabilistically decides whether
to withhold a slot $s$, the result of auctioning $s$ is probabilistic. Then, the results of
the individual auctions for subsequent slots are also probabilistic, because airlines bid
the marginal value of a slot relative to their current slot holdings—the net result may
or may not benefit the airline. Therefore, there is no clear benefit to holding onto a
slot in the auction, and the results for the simulation run indicate that the airline is
more likely to lose by being cautious.

Fig. 6-4 shows the result of increasing the number of cautious airlines on the overall
passenger delay in the auction. The results are for a single simulation of a two-hour
reduction in the arrival rate at BOS to zero arrivals per hour, where all factors other
than the number of cautious airlines were held constant. The caution probability $p$
was fixed at 0.3 for each cautious airline.

The labels indicate which airline was added to the set of cautious airlines at any
given step in the simulation. (Table 6.5 lists the airlines corresponding to the two-
letter airline codes.) For each set of cautious airlines, the difference in the overall
passenger delay from the baseline case (zero cautious airlines) is plotted.

The graph shows that although there is a great deal of noise, the overall trend
is that increasing the number of cautious airlines in the auction increases the total
amount of passenger delay.

Thus, increasing the number of cautious airlines appears to increase the overall
amount of passenger delay. However, individual airlines do not gain by withholding
slots, so in practice there would be no cautious airlines. A necessary qualification
Figure 6-2: Effect of Increasing One Airline's Caution Level on its Passenger Delay in the Auction

Figure 6-3: Effect of Increasing One Airline's Caution Level on its Net Passenger Delay in the Auction
Figure 6-4: Effect of Increasing Number of Cautious Airlines on Overall Passenger Delay in the Auction

to this statement is that revisions to the GDP were not simulated in this thesis. Depending on how slots are reauctioned after a GDP revision, there may or may not be an additional incentive for airlines to withhold slots.

6.4.2 Effect of Increasing Airline Cautiousness in CDM

The effect of increasing airline cautiousness was also tested for the CDM slot allocation method.

To test the effect of increasing cautiousness on a single airline’s delay, the CDM delay of a single airline was determined for various values of the caution probability $p$, where all other factors were held constant. Again, the GDP scenario tested was that of reducing the arrival rate at BOS to two hours.

It was found that increasing the level of cautiousness for a single airline in most cases did not change the resulting passenger delay for that airline. In a few cases, the delay increased.

The effect of increasing the number of cautious airlines (all with $p$ fixed at 0.3) on the overall passenger delay was also tested. The results shown in Fig. 6-5 are for
a single simulation of a two-hour reduction at BOS, to zero arrivals per hour, where all factors other than the number of cautious airlines were held constant. Although there is again a great deal of noise, the overall trend is that increasing the number of cautious airlines in CDM increases the total amount of passenger delay.

Thus, passenger delay appears to increase in CDM if more airlines are cautious, and increasing cautiousness for a single airline seems to have little or no effect on that airline’s delay. However, since GDP revisions were not simulated, it remains unclear whether an airline benefits from withholding slots under CDM.

6.4.3 Effect of Predictive Bidding in the Auction

It is possible that alternative bidding strategies in the auction may lead to better outcomes for individual airlines. In this thesis, one such strategy was simulated—bidding with respect to a predicted final slot allocation, rather than with respect to the slot allocation at the time of the bid.

An attempt was made to determine whether predictive bidding could successfully reduce a single airline’s total passenger delay. Simulations were run in which the
passenger delay of a single airline was first determined with "naive" bidding, and then with predictive bidding. However, the results were inconclusive—predictive bidding was neither consistently advantageous nor consistently disadvantageous. In fact, there is no clear benefit from bidding predictively; this is likely due to the stochastic nature of the slot auction.

The auction was also simulated to determine the effect of increasing the number of predictive airlines on the overall passenger delay. Fig. 6-6 shows the result of a single simulation of a two-hour reduction in the arrival rate at BOS to zero arrivals per hour, where all factors other than the number of predictive airlines were held constant. For each set of predictive airlines, the difference in the overall passenger delay from the baseline case (zero predictive airlines) is plotted.

Again, the results display a high degree of randomness, and the graph represents only one set of simulation results, so it is not possible to draw any firm conclusions. However, other than the points representing three and four predictive airlines, the overall amount of passenger delay appears higher when the number of predictive airlines is increased above zero.

Overall, the effect of predictive bidding is unclear. It is likely that this is due
to the random nature of the auction, where bidders are unsure how the result of bidding on a particular slot will affect subsequent slot auctions—predicting the actual final slot allocation is very difficult, and slight changes in airline behavior can have unpredictable effects on the final slot allocation.

### 6.5 Running Time of the Optimization

The key to a practical auction of airport arrival slots is that the bids for each slot can be determined and processed quickly. The speed of the airline recovery optimizer was tested to determine whether it is a viable tool for this purpose.

Since the set of unslotted route alternatives for each aircraft is pregenerated and remains fixed throughout the auction, the relevant measure of optimizer speed is the time \( t \) to generate slotted routes and solve the integer program. Table 6.5 lists the average value of \( t \) for each airline, for a two-hour arrival rate reduction at BOS to zero arrivals per hour. For each airline, the number of aircraft disrupted by the GDP is listed. Also listed is the total number of unslotted route alternatives generated for all of the airline’s disrupted aircraft. The optimizer running time listed is the average of 100 optimizer runs for that airline.

Most of the airlines have relatively few disrupted aircraft and generate few route alternatives; these airlines were able to generate slotted routes and solve the integer program in well under a second on the 2.20GHz processor. The sole exception, Business Express, generated over 1800 route alternatives, but was still able to solve the airline recovery problem in under four seconds. (Business Express generated so many route possibilities because it operated back-and-forth shuttles several times a day between Logan and nearby airports, and so each aircraft could be assigned to many different combinations of flights.)

The running times show that for a GDP of reasonable length, the airline recovery optimization model proposed in this thesis is extremely fast and adequate to the purpose of calculating the value of a slot in any potential auction.

A caveat is that for extended GDPs or smaller GDPs at hub airports, memory
requirements may become increasingly large, since the number of generated routes can quickly multiply. For example, 600MB was not enough memory to simulate a three-hour shutdown of BOS—in particular, the number of routes generated for Business Express was too large.
<table>
<thead>
<tr>
<th>Airline Code</th>
<th>Airline Name</th>
<th>Num. Disrupted Aircraft</th>
<th>Num. Route Alternatives</th>
<th>Running Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>American Airlines</td>
<td>3</td>
<td>5</td>
<td>290</td>
</tr>
<tr>
<td>AC</td>
<td>Air Canada</td>
<td>3</td>
<td>8</td>
<td>350</td>
</tr>
<tr>
<td>CO</td>
<td>Continental Airlines</td>
<td>4</td>
<td>15</td>
<td>300</td>
</tr>
<tr>
<td>DL</td>
<td>Delta Airlines</td>
<td>9</td>
<td>35</td>
<td>301</td>
</tr>
<tr>
<td>HP</td>
<td>America West Airlines</td>
<td>1</td>
<td>2</td>
<td>280</td>
</tr>
<tr>
<td>HQ</td>
<td>Business Express</td>
<td>23</td>
<td>1809</td>
<td>3750</td>
</tr>
<tr>
<td>NW</td>
<td>Northwest Airlines</td>
<td>2</td>
<td>3</td>
<td>289</td>
</tr>
<tr>
<td>OH</td>
<td>Comair Inc.</td>
<td>2</td>
<td>8</td>
<td>350</td>
</tr>
<tr>
<td>QK</td>
<td>Air Nova</td>
<td>2</td>
<td>7</td>
<td>350</td>
</tr>
<tr>
<td>UA</td>
<td>United Airlines</td>
<td>4</td>
<td>6</td>
<td>290</td>
</tr>
<tr>
<td>US</td>
<td>US Airways</td>
<td>13</td>
<td>36</td>
<td>332</td>
</tr>
<tr>
<td>W9</td>
<td>Eastwind Airlines</td>
<td>2</td>
<td>5</td>
<td>290</td>
</tr>
<tr>
<td>9A</td>
<td>Air Atlantic</td>
<td>1</td>
<td>1</td>
<td>280</td>
</tr>
<tr>
<td>9L</td>
<td>Colgan Air</td>
<td>3</td>
<td>7</td>
<td>323</td>
</tr>
</tbody>
</table>

Table 6.5: Airline Codes and Optimization Running Times
Chapter 7

Conclusions

In this thesis, an auction was proposed as a mechanism for efficiently allocating airport arrival slots during Ground Delay Programs (GDPs), and for thereby reducing overall passenger delay. The auction mechanism, a sequential, sealed-bid Vickrey auction without package bidding, provides airlines with a direct monetary incentive to release arrival slots to other airlines that can use them more efficiently.

The auction was simulated on historical flight data and compared to two other slot allocation methods—the current method of slot allocation, known as Collaborative Decision Making (CDM), and a global optimization method, which simulates the slot allocation that would result if one airline owned every flight and every slot. The CDM and global optimization methods determine upper and lower bounds, respectively, on the amount of achievable passenger delay, and the difference between these bounds is referred to in this thesis as the amount of reducible delay.

An integer programming formulation of the airline recovery problem was developed to determine how an airline could optimize the assignment of flights to arrival slots, and simultaneously assign reroute each of its disrupted aircraft. This formulation was used extensively in all three slot allocation methods simulated, to determine the optimal allocation of flights to slots, to calculate the overall amount of passenger delay, and to determine the value of an individual slot in the auction. The optimization runs in at most a few seconds for problem instances as large as a few thousand route alternatives, although it requires a significant amount of memory for generating
routes, and the memory requirements may be prohibitive for extended GDPs with many route possibilities.

For the GDP scenarios tested, the auction method was successful in reducing the passenger delay resulting under CDM by at least 5-7%, and in some scenarios, by as much as 29%. The percentage of reducible delay captured by the auction varied widely depending on the GDP scenario tested, but was in general higher for more severe, longer-lasting GDPs. As much as 75% of the reducible delay was captured for a simulated scenario where BOS was shut down for two hours.

Alternative airline behaviors were also tested to determine their effect on individual airline passenger delays and on the overall passenger delay. These results were largely inconclusive, although it appeared that increasing the number of cautious and predictive airlines increased the overall delay in the auction, and that increasing the cautiousness of a single airline in the auction increased its delay.

7.1 Future Research

Several areas of research remain to be explored.

As pointed out in Chapter 2, many different types of auction mechanisms can be designed. Any potential auction mechanism could be simulated on actual flight data to determine how effective it is at reducing passenger delay. Also, any auction could be simulated with human participants (perhaps from actual Airline Operations Centers), in order to determine how airlines would actually behave.

One promising mechanism is the combinatorial auction proposed by Rassenti and described in 2.2.1. While the Rassenti auction did not necessarily favor a fair distribution of slots, it may be possible to add additional constraints that would do so. For example, certain constraints in the integer program formulation might guarantee small airlines a minimum number of slots in the resulting allocation. Also, the complexity of the combinatorial auction could be reduced by limiting the size and composition of available slot packages.

The auction described in this thesis serves as a market-clearing mechanism. One
alternative would be simply to assign slots to the airlines, and then allow them to buy and sell the slots on their own. The main problem with this system is that the FAA would lose control over slot allocation; for example, airlines might attempt to exchange slots without notifying the FAA. Furthermore, anonymity might be difficult to preserve in such a trading system, and groups of airlines might be able to collude to deny other airlines choice slots.

A few other issues regarding the implementation of any auction need to be addressed. For example, it is necessary to impose penalties for bidders who default on their bids, in order to prevent false bidding [13]. Also, it is necessary to impose a limit on the time for bidding on slots, in order to prevent an arrival slot from being sold to a different airline after the flight assigned to it has taken off.

There are also a number of possible extensions with respect to the model of the airline recovery problem presented in Chapter 4. In this thesis, several assumptions were made regarding how airlines determine the value of an arrival slot. For example, costs other than the cost of delay were ignored, and the cost of a single passenger minute of delay was assumed to be constant. A more complete model of an airline's costs could be developed for the optimization. Also, improvements could be made to the passenger rerouting model used to determine individual passenger delays.

Also, the airline recovery integer program could be made more accurate through adding back in constraints, such as those for the gate and crew assignment. And, although in this thesis all GDPs are assumed to occur at only one airport, a more comprehensive simulation would account for multiple GDP airports.

It was also assumed in this thesis that the GDP is fixed. However, in practice, GDPs are often revised as more information about weather and air traffic conditions becomes available. Future work might involve determining the impact of GDP revisions on the effectiveness of the auction. In order to compare the auction to CDM, where released slots are redistributed at regular intervals, it would be necessary to hold a new auction of arrival slots at the time the GDP is revised.

Finally, more work could be done to test alternative airline behaviors. In particular, more tests could be run to conclusively determine whether or not "predictive
bidding” and “cautious” behavior lead to improved passenger delay.

7.2 Summary of Thesis Contributions

In this thesis, it was shown that a simple auction mechanism can significantly reduce the amount of passenger delay that results under the current method of slot allocation. It is the first known attempt to quantify the impact of an auction mechanism as an alternative slot allocation method by actually simulating the mechanism on historical flight data.

The other major contribution of this thesis is a new model for solving the airline recovery problem, which allows airlines to assign flights to slots, reroute aircraft and determine how to delay or cancel flights in response to a GDP. The key feature of this airline recovery optimization model is that it runs within a few seconds, and can be used to determine the value of an arrival slot if time is very limited, as it would be in a real-time slot auction.

Through more efficient use of airport resources, it is possible to significantly reduce passenger delays, and therefore, airline costs. This thesis has presented a more efficient method of allocating one particular type of scarce resource, the arrival slot. Hopefully, this thesis and future research will lead to greater use of airport resources, and to lower passenger delays systemwide.
Bibliography


http://www.bepress.com/bejte/frontiers/vol1/iss1/art1/.


