Efficient Capacity Allocation in a Collaborative Air Transportation System

by

William D. Hall

S.B. Aeronautical and Astronautical Engineering, MIT, 1991
S.M. Aeronautical and Astronautical Engineering, MIT, 1992

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1999

©1999 by William D. Hall. All rights reserved.

Author

Department of Electrical Engineering and Computer Science

April 24, 1999

Certified by

Amedeo Odoni
Professor of Aeronautical and Astronautical Engineering
Professor of Civil and Environmental Engineering

Certified by

Milton Adams
Division Leader, Draper Laboratory

Certified and accepted by

Cynthia Barnhart
Associate Professor of Civil and Environmental Engineering
Co-director, Operations Research Center

Certified by

Eric Feron
Assistant Professor of Aeronautical and Astronautical Engineering

Certified by

Eugene Gilbo
Operations Research Analyst, Volpe National Transportation Systems Center

Certified by

Stephan Kolitz
Principal Member of the Technical Staff, Draper Laboratory
Efficient Capacity Allocation in a Collaborative Air Transportation System

by

William D. Hall

S.B. Aeronautical and Astronautical Engineering, MIT, 1991
S.M. Aeronautical and Astronautical Engineering, MIT, 1992

Submitted to the Department of Electrical Engineering and Computer Science on May 5, 1999, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Operations Research

Abstract

This thesis proposes methods of allocating airport capacity to the users of the National Airspace System (NAS) during periods in which demand for the resources exceeds capacity. A metric by which the proposed methods are judged is the value that the users of the NAS are able to realize through the allocation. Maximization of this metric produces notably different results from minimization of flight-minutes of delay and similar objectives employed in related works.

The heart of this approach is the treatment of the Federal Aviation Administration (FAA) and the NAS users as solvers of subproblems in a decomposition of the overall problem of determining how to operate the system. The best possible capacity allocation method would allow the users to achieve the same result collectively that a single omniscient entity in control of all decisions in the system could achieve.

The typical approach to decomposition employed in optimization, that of modifying the subproblem objectives through a penalty function determined by a master "dual" problem, is employed in the Objective-Based Allocation Method (OBAM). It is shown that the proper choice of penalty function results in a method that performs well dynamically and, assuming each user operates to maximize its operating objectives through the allocation, achieves the optimal solution that an omniscient single controller would achieve.

OBAM requires complete communication of user objectives and constraints to achieve optimality. It also requires that the penalty functions used to coordinate the subproblem solutions be added to the user objective functions through assessment of fees. The second part of this thesis addresses the improvement of the decomposition by changing the nature of the allocation without these requirements. Rather than allocate airport arrival capacity alone, a more general notion of airport capacity that captures the interactions between arrival and departure processes at an airport
is allocated. This allows the users the flexibility to adjust the operations mix of the airport according to their objectives and improves the ability of the system to match demand to forecast airport capacity. Through simulation, it is shown that this approach could improve significantly on the primary metric of achieving user value. In addition, the approach facilitates side benefits, such as the reduction of fuel consumption, the reduction of harmful emissions into the environment, and the improvement of service reliability for the flying public.

Thesis Supervisor: Amedeo Odoni
Title: Professor of Aeronautical and Astronautical Engineering and of Civil and Environmental Engineering

Thesis Reader: Milton Adams
Title: Division Leader, Draper Laboratory

Thesis Reader: Cynthia Barnhart
Title: Associate Professor of Civil and Environmental Engineering
Co-director, Operations Research Center

Thesis Reader: Eric Feron
Title: Assistant Professor of Aeronautical and Astronautical Engineering

Thesis Reader: Eugene Gilbo
Title: Operations Research Analyst, Volpe National Transportation Systems Center

Thesis Reader: Stephan Kolitz
Title: Principal Member of the Technical Staff, Draper Laboratory
Acknowledgments

I would like to thank my MIT advisor, Amedeo Odoni, and my Draper supervisor, Milt Adams. Each has expended significant time and effort to provide very conscientious and insightful feedback on this work. I am very grateful for the freedom they gave me to research the issues I found interesting, despite the fact that the payoff was at times unclear.

I would like to acknowledge the valuable contributions of the rest of my thesis committee, Cindy Barnhart, Eric Feron, Eugene Gilbo and Stephan Kolitz. These wonderful people have provided encouragement and enthusiasm throughout the development of this thesis. They have diplomatically delivered suggestions that ultimately strengthened this research, and they have provided data, references, and contacts that have helped make this work possible.

I am grateful to Lorenzo Brunetta of the University of Milan for his careful reading of this thesis and his interest in the mathematical programming aspects herein. Charlie Plott at the California Institute of Technology, and Mike Ball and Bob Hoffman at the University of Maryland have provided references and information that have been extremely important, as have numerous members of the CDM community – including Roger Beatty, Rose Hsu, Ken Howard, Rick Oiesen, Tim Byrne and maybe even Mike Wambgsanss.

Several of my colleagues have volunteered invaluable assistance, and for their help I am truly grateful. Ryan “Rif” Rifkin carefully read several early versions of Chapter 3, helping me tremendously with its presentation and encouraging me with the rest of my work. Nicolas Pujet offered valuable insights into the operation of airlines and into some of the techniques used in this thesis to visualize the data and results. Manoj Lohatepanont helped me with some of my data and provided useful reference material. Bertrand Delcaire, Husni Idris, Ioannis Anagnostakis, and Ruben Barocio-Cots have each helped me understand different aspects of the air transportation system and have served willingly as sounding boards to test some of my ideas.

I would like to thank my family for the wonderful start in life they have given me
and for the privilege and support they have so generously provided. Without them I would have accomplished little.

Finally, it is with amazed gratitude that I thank my wife, Melitta King, for her irrational patience, understanding, and love throughout this difficult time.

This thesis was prepared at the Charles Stark Draper Laboratory, Inc., under Internal Research and Development Contract numbers 15916, 13267 and 18623. This research was also supported in part by the Federal Aviation Administration under the National Center of Excellence in Aviation Operations Research (NEXTOR) project “Advanced Concepts for Collaborative Decision Making (CDM),” award number 603607, and by the National Aeronautics and Space Administration under the MIT “Departure Planner” project, NASA Ames Grant Number NAG2-1128.

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

Permission is hereby granted by the author to the Massachusetts Institute of Technology to reproduce any or all of this thesis

______

______
Contents

1 Introduction .................................................. 17
  1.1 Opportunities to Increase System Throughput .......... 19
  1.2 Decomposition of Decision-Making ..................... 21

2 Decision-Making in the Air Transportation System .......... 25
  2.1 Division of Responsibility ................................ 25
  2.2 Air Traffic Management ................................... 26
  2.3 Problems Solved by the Airlines ......................... 28
  2.4 Interaction Between Airline and ATM Problem Solutions .... 30
    2.4.1 Strategic Planning Level ........................... 31
    2.4.2 Tactical Planning Level ............................ 33
    2.4.3 Operations Interface ............................... 34
  2.5 Dynamics of Decision-Making and Implications for System Improvement 35

3 Optimal Allocation of Airport Arrival Slots According to User Objectives 39
  3.1 Assignment of Arrival Slots to Flights ................ 39
  3.2 Arrival Capacity Allocation Today ....................... 40
  3.3 Ghost Flight Cancellation Problem ...................... 41
  3.4 Slot Auction ............................................ 44
  3.5 Objective-Based Allocation Method ...................... 49
  3.6 Observations Concerning OBAM .......................... 55
  3.7 Pros and Cons of OBAM-Like Slot Assignment Methods .... 59
4 Allocation of Arrival-Departure Airport Capacity to Users

4.1 Problems with Arrival Capacity Allocation .................................. 68
4.2 Configuration Planning and Prediction of Arrival-Departure Capacity 70
4.3 Overview of the Arrival-Departure Capacity Allocation Method ...... 72
4.4 Detailed Description of ADCAM .................................................. 73
   4.4.1 Ration By Schedule ......................................................... 75
   4.4.2 Substitution and Cancellation ........................................... 80
   4.4.3 Compression ................................................................. 81
4.5 Potential Benefits to Users of A-D Capacity Allocation Method ..... 83

5 ADCAM Analysis Tools .................................................................. 91

5.1 System Simulation ......................................................................... 93
5.2 Capacity Forecasts ....................................................................... 94
5.3 FAA Flow Management Models .................................................... 97
   5.3.1 Model of CDM GDP ......................................................... 98
   5.3.2 Model of Arrival-Departure GDP .................................... 100
5.4 Airline Decision Models ............................................................... 102
   5.4.1 Mathematical Model of Airline Decision Problem ............... 117
   5.4.2 Mathematical Model Solution Technique .......................... 121
5.5 Airport Model ............................................................................... 123
5.6 Summarizing Results ................................................................... 124
5.7 Data ........................................................................................... 125
   5.7.1 Comparison of Generated Itineraries and Real Itineraries ... 126

6 ADCAM Analysis .......................................................................... 135

6.1 Analysis One: Performance as a Function of Capacity ................. 136
6.2 Analysis Two: Sensitivity to Objective Function .......................... 140
6.3 Analysis Three: Performance as a Function of Arrival-Departure Inter-
   action ...................................................................................... 145
6.4 Analysis Four: Simulated Versus Provided Passenger Data .......... 147
7 Conclusions

7.1 The Capacity Allocation Methods and Their Potential ........ 152
  7.1.1 OBAM: Incorporation of User Objectives into the FAA Problems 152
  7.1.2 ADCAM: Incorporation of Departures into the Capacity Allocation ........ 153

7.2 Implementation of Capacity Allocation Methods ................. 155
  7.2.1 Development of Arrival-Departure Capacity Forecasting Capability ........ 157
  7.2.2 Development Required to Implement ADCAM .................. 160
  7.2.3 Departure Control ...................................... 160

7.3 Interim Solutions ........................................... 162
  7.3.1 Virtual Departure Queue ................................ 162
  7.3.2 Dynamic Arrival Capacity Estimates ........................ 163

7.4 Directions for Further Research ............................ 164
  7.4.1 Further Development of OBAM and Related Methods .......... 165
  7.4.2 Integration of ADCAM with Tactical Tools ................ 165
  7.4.3 Development of Airline Decision Aids ...................... 168
  7.4.4 Improvement of the Simulation .......................... 168
  7.4.5 Sliding Periods ........................................ 170

A Subcapacity Proof ........................................... 173
List of Figures

2-1 Decision-Making Block Diagram ................. 32
3-1 Example showing incentive incompatibility ........ 46
3-2 Time Horizon Constraint ....................... 52
4-1 Flow of Arrivals and Departures in BOS 22L-27, 22R Configuration .... 64
4-2 Typical Arrival-Departure Capacity Curve ........ 65
4-3 Dallas-Fort Worth International Airport .......... 66
4-4 National Radar at 17:00 UTC on 7 March 1998 .......... 67
4-5 Scheduled vs Actual Operations at ATL from 17:00 to 18:00 UTC on 7 March 1998 .......... 67
4-6 Idealized DFW Capacity Graph ................. 69
4-7 Response of Users to Arrival Flow Management ........ 73
4-8 Response of Users to Arrival-Departure Flow Management ........ 74
4-9 Ration by Schedule under Arrival-Departure Capacity Allocation ........ 77
4-10 $\alpha - \delta$ Representation of Airport Capacity ........ 78
4-11 User Airport Subcapacity ...................... 79
4-12 Solution with Binding Subcapacity Constraints but Overall Slack ........ 81
4-13 Connection Matrix Example .................... 85
4-14 Northwest Airlines' Scheduled Banks .............. 86
4-15 Continental's Connection Matrix for EWR .............. 87
4-16 Preserving Connections under CDM (middle) and ADCAM (right) .......... 88
5-1 Simulated Elements of Decision-Making ............ 92
5-2 Capacity Forecast and Scheduled Demand (two periods skipped) . . . 96
5-3 Demand-Capacity Window Pop-Up Menu .......................... 98
5-4 Setting the AAR .................................................. 99
5-5 CDM Ground Delay Program Constraints ......................... 100
5-6 ADCAM Ground Delay Program Constraints ...................... 100
5-7 Route Structure of Major Airlines ................................. 104
5-8 Route Structure of American Airlines .............................. 105
5-9 ORD String Line Diagram ........................................ 107
5-10 Filled Aircraft Seats ............................................. 108
5-11 Connection Targets ............................................... 109
5-12 Arriving Connection Target Chosen .............................. 110
5-13 Pseudo Times Assigned to Departures ........................... 111
5-14 Probability Density Function Used to Choose Departure Connection . 112
5-15 Gamma Distribution Used to Choose Departure Connection ...... 113
5-16 Passenger Connection Made ..................................... 115
5-17 Passenger Connections Generated for NorthWest ................ 116
5-18 Generated Passenger Connections at Airport A .................. 127
5-19 Airline-Provided Passenger Connections at Airport A ............ 128
5-20 Generated Passenger Connections at Airport B .................. 129
5-21 Airline-Provided Passenger Connections at Airport B .......... 130
5-22 Histogram of Passenger Connections Time at Airport A .......... 132
5-23 Histogram of Passenger Connection Time at Airport B .......... 133

6-1 Capacity Forecast Form Used in Analysis One ...................... 137
6-2 Analysis One Results at ORD ..................................... 138
6-3 Analysis One Results at MSP ..................................... 139
6-4 Objective Value of ADCAM Divided by Objective Value of CDM as Delay Multiplier Varies, 50% Capacity at ORD .......................... 144
6-5 Capacity Forecast Forms Used in Analysis Three .................. 146
Objective Value Achieved under CDM versus ADCAM for Simulation-Generated and Airline-Provided Passenger Data
List of Tables

1.1 Increase in average block times of selected flights from Logan Airport 19
1.2 Overview of Problems in the Air Transportation System .......... 20

3.1 Ghost Flight Delay Example. ...................................... 41
3.2 Imaginary Flight Cancellation under CDM. .......................... 42

5.1 Capacity Constraint Input Sequence .................................. 97
5.2 Airline Color Codes .................................................. 103
5.3 Airline-Specific Cost Parameters .................................... 117
5.4 Airline Model Indices ................................................ 120
5.5 Airline Model Decision Variables ................................... 120
5.6 Airline Model Constants ............................................. 120
5.7 Airline Model .......................................................... 121
5.8 Connection Time Statistics .......................................... 131

6.1 Sum of Objective Value Achieved by Airlines in Analysis One .... 137
6.2 Objective Values Achieved by Airlines in Analysis One at ORD .... 141
6.3 Objective Values Achieved by Airlines in Analysis One at MSP .... 141
6.4 Objective Values Achieved by Airlines in Analysis Two at ORD .... 143
6.5 Sum of Objective Value Achieved by Airlines in Analysis Three ... 147
6.6 Objective Value Achieved by Data Provider under CDM and ADCAM with Simulated and Provided Passenger Connections ................. 148
Chapter 1

Introduction

The air transportation system is one of the most logistically complex systems in the world, and its performance is vital to the world economy. According to the National Civil Aviation Review Commission charged by the United States Congress to investigate the safety and funding of the air transportation system,

Unless the FAA and various aviation stakeholders - the Congress, the Executive Branch, and the aviation community - change the status quo, internal and external to the FAA, our nation's aviation system will succumb to gridlock. Delays will skyrocket while we reminisce about the "reliable" flight schedules of the past. This current course will impair our domestic economy, reduce our standing in the global marketplace, and result in a long term deterioration of aviation safety. [8]

The same document reports,

In just the past several decades, ... [aviation has grown] from a minor industrial sector to being 6% of the Gross Domestic Product (GDP). U.S. airline and aerospace industries directly employ approximately 1.5 million people, mostly in highly skilled, high-paying jobs that generate more than $100 billion a year in wages.

According to the 1997 World Development Survey, the world's air travelers are expected to double from one billion to more than two billion over
the next twenty years. The total economic impact of air transport on the world economy was $1.14 trillion in 1994. This is expected to increase to $1.7 trillion by the year 2010. Presently, over $1.5 trillion worth of freight is moved through the air around the globe annually.

Unfortunately, as the volume of air traffic has increased, system capacity has not kept pace. The level of congestion in the system has risen consistently over the years, resulting in increasing delays during peak periods of travel. A clear indication of these delays is the padding the airlines have added to their schedules. For example, Table 1.1 compares scheduled times in 1973 and 1994 for a representative set of important flights out of Boston [12].

Delay in the system is more than an inconvenience; it is costly. As block times grow, an airline needs to use more aircraft to deliver the same frequency of service. Delays cause each flight to burn more fuel, accumulate more wear on the aircraft and require more paid time of airline personnel. Similarly, the longer block times cost the cargo and passengers valuable time; delaying each of the billion or so passengers worldwide by an average of ten minutes amounts to eighty thousand person-years of delay. Perhaps the largest cost associated with the situation is due to the additional uncertainty and lack of reliability caused by unpredicted system delay.

The most readily measured cost of delay is the direct operating cost incurred by the airline industry. The National Civil Aviation Review Commission reports:

In 1995, the FAA estimated that airline delays cost the industry approximately $2.5 billion per year in higher operating expenses. That cost is clearly higher today and will grow. [8]

$2.5 billion amounts to 2.5% of the $100 billion direct economic production of the aviation industry. The significance of this figure becomes even more apparent if one considers that the total profits of the U.S. airline industry in 1998 amounted to approximately $5 billion.

Regardless of the precise costs of delay, the problem is of vital importance. While air traffic is forecast to increase, system capacity is unlikely to keep pace. The capacity
<table>
<thead>
<tr>
<th>Flights from Boston To</th>
<th>Block time 1973 (minutes)</th>
<th>Block time 1994 (minutes)</th>
<th>increase (minutes)</th>
<th>increase (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Guardia</td>
<td>56.6</td>
<td>63.29</td>
<td>6.69</td>
<td>12%</td>
</tr>
<tr>
<td>Chicago O'Hare</td>
<td>149.2</td>
<td>159.68</td>
<td>10.48</td>
<td>7%</td>
</tr>
<tr>
<td>Newark</td>
<td>62.8</td>
<td>79.61</td>
<td>16.81</td>
<td>27%</td>
</tr>
<tr>
<td>Atlanta</td>
<td>149</td>
<td>158.6</td>
<td>9.6</td>
<td>6%</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>69.4</td>
<td>79.31</td>
<td>9.91</td>
<td>14%</td>
</tr>
<tr>
<td>Washington National</td>
<td>78.3</td>
<td>95.64</td>
<td>17.34</td>
<td>22%</td>
</tr>
<tr>
<td>Dallas Fort Worth</td>
<td>227.8 (1974)</td>
<td>244.12</td>
<td>16.32</td>
<td>7%</td>
</tr>
</tbody>
</table>

Table 1.1: Increase in average block times of selected flights from Logan Airport

of the U.S. air transportation system is determined largely by the capacity of the airports located in densely populated metropolitan areas. Addition of capacity to these airports is a long, involved, expensive process that is often blocked by the local politics concerning airport noise, pollution and the fear of air crashes. This thesis focuses on means to increase the economic throughput of the air transportation system given the reality that capacity is insufficient to serve all the demand without delay.

1.1 Opportunities to Increase System Throughput

Operation of the air transportation system requires that many interacting problems be solved by many decision-makers. Some of the problems solved routinely by the airlines and the FAA, loosely ordered from most strategic to most tactical, are shown in Table 1.2. A poor solution to any one of these problems can have major negative consequences for the entire system.

Each of these problems affects the others to some extent. Clearly, the airline's infrastructure and schedule have a large impact on fleet scheduling, crew scheduling, and so forth. The schedule's structure determines the airline's ability to solve the arrival slot assignment problem in such a way as to recover from irregular operations. The schedule should be developed to take into account the effect it has on these other problems. The problems on the right side of Table 1.2 interact with each other similarly.
<table>
<thead>
<tr>
<th>Airline Side</th>
<th>FAA and Airport Side</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fleet acquisition / maintenance</td>
<td>Facility construction / improvement</td>
</tr>
<tr>
<td>Schedule development</td>
<td>Facility maintenance</td>
</tr>
<tr>
<td>Fleet assignment</td>
<td>Development of operating procedures</td>
</tr>
<tr>
<td>Fleet scheduling</td>
<td></td>
</tr>
<tr>
<td>Crew scheduling</td>
<td></td>
</tr>
<tr>
<td>Arrival slot assignment</td>
<td>Strategic flow management</td>
</tr>
<tr>
<td>Flight routing</td>
<td>Tactical flow management</td>
</tr>
<tr>
<td>Gate assignment</td>
<td>Airport configuration planning</td>
</tr>
<tr>
<td>Ground resource allocation</td>
<td>Control of aircraft</td>
</tr>
</tbody>
</table>

Table 1.2: Overview of Problems in the Air Transportation System

While a considerable amount of effort has been expended to implement optimal and near-optimal solutions to the problems on the airline side of Table 1.2, less has been done on the FAA side, partly because of the difficulty of determining a meaningful objective function. Recently, some research has been devoted to solving simultaneously several of the problems on each side of the table. For instance, [22] treats the schedule generation and fleet assignment problems in combination, and [2] solves a combined fleet-assignment and fleet-scheduling problem. It has been found that solutions to these combined problems are often better overall than the combinations of solutions that are, taken separately, optimal to the original problems. Indeed, the combination of solutions optimal to the original problems may not even be feasible to the combined problem.

Until recently, very little attention had been paid to interactions between problems on opposite sides of Table 1.2. However, within the past two years, a modern communications infrastructure between the FAA and the Airline Operations Centers (AOC’s) has been set up under the Collaborative Decision Making (CDM) program. The infrastructure consists of a communications network called the CDMNet which has made possible the use of software such as the Flight Schedule Monitor (FSM) tool that shares information between the FAA and airlines in real time. The CDMNet creates opportunities for new forms of communication between the FAA and airlines and begs such questions as, what information should be shared? How can this techno-
logical capability to change fundamentally the nature of FAA-airline problem-solving be best employed to improve the National Airspace System?

These are the questions this thesis strives to answer, in part. The questions assume a common definition of optimality. A variety of optimality definitions has been used in the past. For instance, [28], [17] and [35] use as a cost function the weighted sum: airborne aircraft-minutes of delay times a multiplier plus ground aircraft-minutes of delay. This objective does not account for the differences between costs at different airlines which may be operating different types of aircraft with different numbers of passengers aboard, trying to make connections with different amounts of slack time between flights. Indeed, any such cost function based on directly measurable quantities such as aircraft-minutes of delay is unlikely to capture all the subtleties of the problem. Such directly measurable cost functions may have a tendency to improve system economic production, but often result in dilemmas in which improvement of a readily measurable objective may run counter to the achievement of more fundamental — and less easily quantified — system objectives.

In recognition of the difficulty of producing an objective function that captures the real system objectives, the objective employed in this thesis is the sum of user objectives\(^1\). In a very real sense, maximizing the sum of user objectives maximizes the economic production of the system. Furthermore, as the air transportation system changes and airlines develop new ways of doing business, the usefulness of any fixed, directly measurable objective function would change, whereas the sum of user objectives automatically adjusts to reflect new system operations through the adjustments the users make to their objective functions.

### 1.2 Decomposition of Decision-Making

No single organization is responsible for operation of the entire air transportation system, nor does any single organization have all the information necessary to operate

\(^1\)The user objectives are, of course, determined by each of the users for himself. They may depend on measures such as operating cost, passenger delay, flight delay, revenue-passerger-minutes of delay, passenger traffic lost to competing carriers, and any other measures important to the user.
the entire system. Thus, the operation of the entire air transportation system may not be directly optimized in a single mathematical model in practice. Rather, operation of the system is determined by a combination of decisions made by separate organizations solving problems with respect to their own objectives. This organizational structure results in a decomposition of the global capacity allocation problem into subproblems solved by the various organizations. The decomposition in use today has evolved into its current form over the years.

This thesis considers methods of allocating airport capacity that respect the practical requirement for decomposition. The two forms of decomposition considered are objective-coordinated decomposition, in which the FAA coordinates the airline subproblems by introducing penalty functions into their objective functions, and constraint-coordinated decomposition, in which the FAA coordinates the airline subproblems through the addition of airline-specific constraints. The first approach studied in this thesis (in Chapter 3) is of the objective-coordinated type, in which the FAA imposes appropriately structured fees on the airlines and solves a dual problem that incorporates the airline objectives. To implement this approach in practice, the airlines would transmit their objectives over the CDMNet to the FAA, which would use them to define and solve an optimization problem. To ensure that the airlines have incentive to report their objectives accurately, they would be charged appropriately structured fees. This approach draws a parallel with decomposition approaches to optimization in which a master problem (i.e., the FAA’s problem) is solved to determine coordination variables that are introduced into the subproblem (airline) objectives.

The second approach studied (in Chapter 4) is a constraint-coordinated decomposition. Because the decomposition is coordinated through constraints, this second approach improves decision-making without requiring that fees be assessed. In this respect, the approach is similar to the method used to allocate capacity during ground delay programs today. However, the family of constraints used to coordinate the subproblem solutions in the decomposition of Chapter 4 represent airport capacity more faithfully than the constraints used today, while allowing the users more flexibility
in the solution of their subproblems. The potential benefit of this better set of constraints to improve system economic performance is investigated in Chapters 5 and 6.

Chapter 2 provides some background on the decisions that are made in the air transportation system and on how the approaches developed in Chapters 3 and 4 fit into the overall decision-making scheme.
Chapter 2

Decision-Making in the Air Transportation System

This chapter provides an introduction to the decision-making processes in the air transportation system, with a focus on the operational decisions made by the airlines and air traffic management (ATM) agencies such as the FAA. ATM agencies decide how to operate airports and airspace in support of the goals of safety and efficiency in air transportation. The airlines, on the other hand, decide how to use their resources of people, equipment, and infrastructure to deliver passengers and cargo to their destinations in a timely, safe and profitable manner.

2.1 Division of Responsibility

The division of responsibility between the FAA and the airlines seen in the problems of Table 1.2 is typical of the division of responsibility between the air traffic management agencies and airlines in most countries. Minor differences may exist: some of the problems under control of the airlines in Table 1.2, such as gate assignment, may be under the jurisdiction of an airport authority or air traffic control agency in other parts of the world. Similarly, the users may take on some of the responsibilities shown under the jurisdiction of the FAA in other parts of the world. In Africa, for instance, there are areas of airspace in which air traffic control services do not include
separation assurance; IFR aircraft must negotiate with each other for use of certain routes using designated radio frequencies. Nevertheless, the areas of the world with high traffic volume employ very similar frameworks for air transportation decision-making with very similar divisions of responsibility between airlines and air traffic management.

Section 2.2 discusses the functions performed by the ATM agencies and the decision problems that the ATM agencies solve in support of these functions, with emphasis on operations in the United States. Section 2.3 discusses the functions and decision problems of the airlines, and Section 2.4 discusses the interaction between the problems faced by the two communities. Finally, Section 2.5 discusses ways in which the performance of the air transportation system may be improved with respect for the dynamics of these interactions and describes several efforts under way to realize such improvement.

2.2 Air Traffic Management

The air traffic management agencies in developed countries around the world, including the FAA, enforce safe capacity limits at each airport and each sector of controlled airspace. The safe capacity of an airport is determined by the capacities of its parts – the runways, the taxiways and the terminal airspace – and by interactions between operations in each part. The capacity of a runway is determined by rules defining minimum allowable spacing (temporal or spatial) between operations on the runway. Yu [38] provides more information on airport capacity. The capacity of a sector is similarly determined by minimum separation requirements between aircraft in the sector and by the structure of defined airways that traverse the sector. Under the Free Flight initiative, it is hoped that these requirements can be reduced to the minimum separation requirement alone.

Tactical decision-making within the FAA is distributed geographically through a structure wherein each part of the agency concerns itself with the safety requirements of a specific airport or a specific part of the airspace. A pilot flying from one airport
to another communicates with the part of the agency responsible for the part of the airspace system in which he is operating. Normally, the communication is with controllers whose primary concern is the satisfaction of the safety constraints for the area they oversee. Occasionally, those controllers will also be concerned with flow management constraints given them by a downstream controller farther along the aircraft’s route. Such flow management constraints typically delay the aircraft’s arrival into the downstream controller’s area, allowing the downstream controller to prevent the simultaneous operation of too many aircraft in his area. In effect, the downstream controller is issuing control to the pilot through an intermediary, and each controller is concerned only with his own portion of the airspace system.

The strategic decisions of the FAA, on the other hand, are made in centralized locations and based on data collected from the entire system. For instance, strategic flow management is coordinated by the FAA’s Air Traffic Control System Command Center (ATCSCC) in Washington, D.C. Information used in flow management and upon which other strategic decisions are based is collected and archived at the Volpe National Transportation Systems Center in Cambridge, Massachusetts.

The functions performed by air traffic management include (listed from most strategic to most tactical):

1. Facility construction / improvement: Investment in air traffic control systems, navigation facilities, airports, and airport facilities for the near-term and far-term future.

2. Facility maintenance: Navigation aids, radars, lighting systems, runways, pavement, and buildings require routine maintenance that must be planned to avoid interrupting airport operations unnecessarily. Unplanned failures of equipment and acts of nature such as snowstorms may require rapid response of personnel.

3. Development of operating procedures: Operating procedures are developed to ensure safety and improve efficiency.

4. Strategic flow management: When capacity in part of the system drops significantly for reasons that are forecast in advance, strategic flow management,
often implemented through a Ground Delay Program (GDP), is used to reduce the traffic to levels that may be satisfied by the capacity.

5. Tactical flow management: Flow management in response to short-term capacity insufficiency may be achieved through Miles-In-Trail (MIT) restrictions or other procedures which space aircraft farther apart en-route, or through procedures that control the flow of aircraft into specific parts of the airspace.

6. Airport configuration planning: Planning the sequence of airport configurations to use, accounting for issues such as noise impact on residential neighborhoods, weather, and air traffic demand.

7. Control of aircraft: Directing aircraft on the ground and in the air to ensure safety and efficiency.

2.3 Problems Solved by the Airlines

Whereas ATM segments tactical control authority geographically, the users, who are concerned with the interactions among their flights, have centralized much of their planning and control. The extra degree of coordination achieved by the airlines’ centralized Airline Operations Centers (AOC’s) is required because the delay of a single flight may affect many other flights in the airline’s schedule through the connections that passengers, flight crew, and equipment must often make.

The decisions that must be made by the airline include:

1. Fleet acquisition / maintenance: The airline needs enough aircraft to serve its schedules. Aircraft orders must often be placed years in advance of delivery. An airline can realize significant cost savings and flexibility by limiting the number of types of aircraft it operates, yet gains operating efficiency by having the aircraft of the right size and range for each flight. New aircraft are more expensive than older aircraft, yet have lower variable operating costs and are more reliable.
2. Schedule development: The published schedule is the airline’s primary means of differentiating itself from its competitors. The schedule encapsulates the marketing strategy of the airline. The schedule also drives the rest of the problems faced by the airlines, determining the logistics and robustness of the airline’s operations.

3. Fleet assignment: Given an airline fleet, the airline must decide which type of aircraft (e.g., 747 or 767) to use for each flight, determining the capacity of the flight and driving the fleet and crew scheduling problems.

4. Fleet scheduling: A fleet schedule specifies which aircraft are to serve each flight. The fleet schedule must include time for each aircraft to spend at maintenance facilities, and must respect physical constraints, such as the fact that aircraft cannot change location without being flown.

5. Crew scheduling: A crew must be rated for the type of aircraft it is to fly. The crew schedule provides a matching between the types and locations of crew available and the fleet schedule. The crew schedule must respect numerous technical, regulatory, and contractual constraints.

6. Arrival slot assignment: During periods when a given airport’s capacity is insufficient to serve the schedule, departures destined for that airport are limited through a ground delay program, or GDP. At such times, arrival capacity is reserved for each affected flight in the form of an arrival slot. Each airline is allowed to assign its arrival slots among its flights as it pleases, subject to some reasonable constraints. An airline must often cancel and delay many flights during a GDP. Choosing an assignment of arrival slots to flights involves re-solving dynamically the fleet assignment, fleet scheduling, and crew scheduling problems to minimize the GDP’s impact on the airline’s passengers and operations.

7. Flight routing: A flight often has many choices of trajectory between origin and destination. Winds, aircraft type and payload all affect the choice of route.
Additionally, the presence of other traffic in the system may cause greater delay on some routes than on others.

8. Gate assignment: The choice of gate determines, in part, the amount of time required for passengers, crew, baggage, and cargo to be transferred to other locations. Some gates can serve only certain types of aircraft, and use of some gates by aircraft with large wingspans blocks adjacent gates. All such considerations must be taken into account when operational changes force last-minute changes in the gate schedule.

9. Ground resource allocation: A wide variety of ground resources is required to service aircraft and passengers at an airport, including support personnel, aircraft tugs (tow vehicles), baggage handlers, catering trucks and fuel trucks. These must be allocated to the tasks involved in turning an aircraft.

2.4 Interaction Between Airline and ATM Problem Solutions

During so-called regular operations, i.e., under normal operating conditions, system capacity allows the airlines to operate their flight networks more or less according to schedule. Thus, the schedule provides reasonably accurate information about flight arrival and departure times to air traffic management for use in planning system operations. Adherence to the schedule also allows the preplanned solutions to the airline’s problems such as flight, crew, and gate assignment to be used with little or no modification.

Unfortunately, regular operations do not occur as routinely as one might hope. Between September of 1998 and April of 1999, there was an average of 1.6 ground delay programs in the United States per day. A ground delay program affecting an airport or airports important to an airline’s operations will cause that airline to

\footnote{Source: Collaborative Decision Making GDP logs.}
experience *irregular operations*. During irregular operations, the airline's schedule cannot be maintained. The airline must find in real time new solutions to many of the problems outlined in Section 2.3.

Figure 2-1 is a diagram of decision-making in the national airspace system during irregular operations. The figure is arranged with the longest-term decisions represented in the top “strategic planning” box, medium-term decisions represented in the “tactical planning” box, and the shortest-term decisions represented in the bottom “operations interface” box.

### 2.4.1 Strategic Planning Level

At the strategic planning level of Figure 2-1, the airlines and the FAA plan ahead as far as their forecasts of system capacity and system operations allow. The results of the FAA’s plans at this strategic level are in the form of ground delay programs issued for those airports at which capacity is forecast to be less than demand. The FAA typically issues a ground delay program an hour or two in advance of the time it becomes effective, although unforeseen changes in weather conditions may cause a ground delay program to become effective simultaneously with its issuance. Ground delay programs are typically issued for a period of six or eight hours, though they may cover longer or shorter periods. Once the ground delay program is issued, it may be canceled, revised, or extended depending on the circumstances.

The objective of the airline at the strategic planning level during irregular operations is to minimize the impact of irregular operations on the airline's passengers and operations. Personnel in the AOC find solutions to the fleet and crew scheduling problems, the arrival slot assignment problem and the flight routing problem in response to the capacity constraints the FAA has given their airline. Their solutions are communicated to the FAA and are taken into account when the FAA plans operations at the airport.

Not only does the need to find new solutions to its problems in real time during irregular operations complicate matters for the airline, it also reduces the accuracy of the FAA’s information concerning airline operations. The FAA needs such informa-
Figure 2-1: Decision-Making Block Diagram
tion to determine the extent to which demand for airport resources exceeds capacity. For instance, if the FAA is not informed sufficiently early that an airline will cancel several flights, it might be unable to allow other flights to use the capacity reserved for the canceled flights.

The FAA accommodates demand uncertainty, such as that caused by unreported cancellations, by maintaining a buffer or reservoir of aircraft at strategic locations in the system. For instance, the FAA maintains a managed arrival reservoir (MAR) of airborne aircraft near each GDP airport by allocating slightly more airport arrival capacity than is forecast to exist. Inaccuracies in the FAA’s estimates of demand lead to higher variance in ground and airborne aircraft queue lengths than could be achieved with perfect demand information. Improvements in the accuracy of demand information could indeed simultaneously reduce the average queue length and reduce the queue-empty time, increasing system throughput.

Just as the accuracy of the information provided to the FAA by the airlines is vital to the quality of the solutions the FAA can produce, the accuracy of information provided to the airlines by the FAA has a profound impact on the quality of the airline solutions. If an airline bases its decisions on inaccurate forecasts of the arrival slots that will be allocated to it by the FAA, those decisions will be unnecessarily poor. Unfortunately, it is quite common for such information to be inaccurate: a very high fraction of ground delay programs are not executed as predicted, but are extended, revised, or canceled instead.

### 2.4.2 Tactical Planning Level

At the tactical planning level of Figure 2-1, the airlines and FAA plan in greater detail than at the strategic planning level. Aircraft arrival and departure times at the strategic planning level are uncertain because strategic plans are formed hours in advance. Plans at the tactical planning level must cover a shorter time horizon because their higher precision requires greater accuracy in predicted arrival and departure times. Much of the information used to plan at the tactical level is based on observation rather than prediction because predictions currently available lack the
accuracy required for tactical planning.

Because the airlines have control over their departures, they have accurate information about the times at which flights will leave their gates. Unfortunately, under current operations, the first accurate indication the FAA receives of an aircraft's departure time usually occurs when the pilot calls the tower for clearance to push back from the gate. Better information about departure times would improve the FAA's ability to plan airport configurations and sequence operations for efficiency.

Similarly, because the FAA controls aircraft enroute, it has better information about arrival times than do the airlines. The airlines could use better arrival time information to plan gate and ramp operations. Better arrival time information might also be useful for an airline in predicting the times at which its arriving passengers could reach the gates of their departure connections.

Tactical planning is normally performed locally at each airport. The airlines operate stations at each airport, where the gate assignment and ground resource allocation problems are solved. The FAA operates the tower and TRACON for each airport, in which the tactical problems related to configuration changes, tactical flow management and other operations are addressed.

2.4.3 Operations Interface

The operations interface level is where the strategic and tactical plans created in the rest of the system are implemented. Several efforts have been undertaken to design automation tools that improve system performance at this level (see, for example, [12], [20], [13], [11], [10], and [25]) including a suite of tools called the Center TRACON Automation System (CTAS) in use today at Dallas-Fort Worth Airport. The design of such automation tools is challenging not only because of the complexity of the tasks the tools must perform, but also because, in order for the tools to be accepted by the aviation community, they must reduce workload. Controllers and airline personnel performing time-critical tasks have little spare time in which to devote attention to automation tools.
2.5 Dynamics of Decision-Making and Implications for System Improvement

The time required for decisions to be made by the FAA in the strategic planning level of Figure 2-1 and the time required for the airlines to make decisions about flight cancellation and delay are significant. Part of the reason these tasks require significant amounts of time is because the decomposition of decision-making between the airlines and the FAA requires each to make commitments to the other in the face of uncertainty. By delaying these decisions, the commitments can be made with more certainty. For instance, an airline might know that it is very likely it will need to cancel a number of its flights because of equipment or crew problems, yet be unable to commit to the specific flights it plans to cancel until it has more information.

The delay in decision-making and communication between the FAA and airlines at the strategic planning level is costly. The delay reduces the quality of information available to flow management, resulting in unnecessary queueing and delay. It reduces the quality of information available to the airlines as well, resulting in higher operating costs and lower performance than could be achieved with better information.

Two approaches can be taken to minimize the cost of these delays in decision-making. First, the delays themselves can be reduced. Second, given that some delay will remain regardless of efforts to reduce it, the structure of decision-making can be changed to accommodate the delays better.

The first approach has been pursued by the Collaborative Decision Making (CDM) effort. The CDM effort has made great strides toward improving the quality of information available to the FAA and the airlines during GDP planning [1]. CDM has provided new procedures and communications infrastructure for the FAA to send GDP coordination information to the airlines and for the airlines to send information about their planned use of airport capacity to the FAA. The new communications infrastructure allows the airline to change its assignment of arrival slots to flights more easily, effectively reducing the airline’s commitment of flight-slot assignment to a commitment of the number of slots to be used.
The second approach, changing the structure of decision-making to accommodate decision-making delays better, is pursued by the capacity allocation methods proposed in Chapters 3 and 4. These allocation methods give the airlines more control over flow management decisions. As a result, the airlines can use the detailed knowledge of the problems they face to improve the aspects of flow management over which they gain control without making unnecessary commitments to the FAA. Additionally, each airline can use its own knowledge of operations and costs and its own tolerance for risk to decide how its share of airport capacity should be used.

The allocation method introduced in Chapter 4 also improves system performance by improving the model of airport capacity on which flow management decisions are based. Reliance on this improved model of airport capacity results in flow management decisions that allow the system to perform better despite the delays present in the feedback of the strategic planning level of Figure 2-1.

Better information and system modeling could be used to improve system performance at the tactical and operational levels of Figure 2-1 as well. Determination of the arrival sequence and runway assignment requires accurate information about the current locations of aircraft and their performance characteristics. Better runway assignment and arrival sequencing could improve airport throughput significantly by avoiding mixing large and small aircraft operations on the same runway. The inefficiency of mixed large and small aircraft operations on a single runway is due in part to the large distance small aircraft must remain behind large aircraft to avoid wake vortex turbulence. The Center TRACON Advisory System (CTAS) is designed to help controllers determine a landing sequence and runway assignment, and in so doing has achieved some improvements in airport throughput [11]. CTAS spans the tactical planning and operations interface levels of Figure 2-1; not only does it plan arrival operations tactically, it also provides sequencing information to controllers on their radar screens.

Determination of the departure sequence could have similar effects on throughput. Departures must be spaced apart to avoid wake vortex turbulence just as arrivals must. Additionally, departures following the same routes out of the TRACON must
be spaced far enough apart to allow departure and en-route controllers to maintain adequate separation between them. One strategy often used by ground controllers sequencing departures at Logan Airport in Boston is to alternate turboprop and jet departures since the turboprops and jets are following different routes. This strategy facilitates achievement of en-route separations and reduces departure and en-route controller workloads, but it may be an inefficient use of the runway. A tool that could plan for the departure sequence ahead of time and help controllers achieve the planned sequence without increasing workload could increase airport efficiency considerably [12] [19] [20].

Finally, planning arrival and departure sequences simultaneously could perhaps provide a greater increase in airport throughput than any tactical tool. Each runway has wake vortex spacing requirements between successive arrivals and successive departures. However, because a departure does not fly through the same part of the airspace as an arrival, there is no wake vortex problem between an arrival and subsequent departure on a single runway. The time required for a departure to pull onto the runway and take off is often less than the time required between successive arrivals. Thus, it may be possible to fit departures between arrivals without affecting the arrivals at all; indeed, this practice is frequently observed when airports are limited to single runway operations. Planning for mixed operations requires additional coordination between the controller managing the runway and the approach and departure controllers, thus it is often used only when made necessary (and simpler) by lack of available runways. A tool that could facilitate such coordination, accounting for all the factors that affect arrival and departure sequencing, could increase the achievable number of operations at an airport considerably. The optimization behind such a tool is under research by Alp Muharremoglu [24].

The information requirements of such tactical decision aids are considerable. Implementation of the strategic tools proposed in this thesis would provide some of the information required by these tactical decision aids, facilitating their implementation.
Chapter 3

Optimal Allocation of Airport Arrival Slots According to User Objectives

3.1 Assignment of Arrival Slots to Flights

During a GDP, the FAA maintains a manageable flow of aircraft arriving at an airport by controlling the departure times of the aircraft from their origins. It does this by assigning arrival slots to the flights and requiring each flight to have an appropriate arrival slot before it may take off. An arrival slot represents an allocation of the destination airport resources necessary for an aircraft to land safely: the sole use of a specific runway and approach airspace for a period of time and the attention of the air traffic control specialists required to ensure a safe landing. The chief identifying characteristics of an arrival slot are the airport and the time at which the aircraft is to land.

As mentioned in Chapter 2, the number of arrival slots predicted to be available at an airport is determined in large part by the weather conditions, the forecasts of which often are not sufficiently accurate for a reliable prediction of the arrival slots available more than a few hours in advance. As a result, arrival slots cannot be assigned to
flights far in advance without either running the risk of wasting considerable amounts of capacity (due to an underestimation of future capacity) or risking large airborne queues of aircraft waiting to land (as a result of overestimation of future capacity). Any practical slot assignment method must therefore be dynamic: its design and implementation must take into account the fact that frequent replanning will be necessary.

3.2 Arrival Capacity Allocation Today

A GDP delays the arrival of a subset of the flights scheduled to arrive at the airport during the period when arrival demand is greater than capacity. International flights are exempt from GDP's, and long-haul flights are often exempt from GDP's of short or uncertain duration since short-haul flights can respond more quickly to changes in predicted capacity. In a GDP, the FAA allocates arrival slots to the flights that are not exempt from the program. The affected flights are assigned arrival slots on a first-scheduled, first-served basis. In practice, the flights do not necessarily land in arrival slot order because the FAA uses the pre-takeoff slot assignments to meter the flow of aircraft into the air only, not to meter the flows of airborne aircraft.

Once the assignments of slots to flights have been made, each user is free within limits to cancel flights and to swap slots among its flights, with the restriction that no swap may assign a flight to a slot that would require it to land more than twenty minutes ahead of schedule. Through swapping, user objectives are accommodated in the solution. This step is important to the users, since it allows them to manage the impact of the capacity shortage on their operations.

One shortcoming of basing the allocation of slots on the schedule is the incentive it gives users to schedule ghost flights – flights they do not intend to fly. Before users were allowed the flexibility provided by CDM in the use of their arrival slots, the advantage to a user of having slots available to be canceled was not great. Now that CDM is in place, a user can take advantage of slots allocated to ghost flights in order to reduce delays on the user's other flights. This additional flexibility gives the users
<table>
<thead>
<tr>
<th>Airline</th>
<th>FLTNO</th>
<th>ETA</th>
<th>CTA</th>
<th>Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>101</td>
<td>0700</td>
<td>0700</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>102</td>
<td>0705</td>
<td>0710</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td>103</td>
<td>0705</td>
<td>0720</td>
<td>15</td>
</tr>
<tr>
<td>B</td>
<td>104</td>
<td>0710</td>
<td>0730</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>105</td>
<td>0710</td>
<td>0740</td>
<td>30</td>
</tr>
<tr>
<td>A</td>
<td>106</td>
<td>0710</td>
<td>0750</td>
<td>40</td>
</tr>
<tr>
<td>C</td>
<td>107</td>
<td>0720</td>
<td>0800</td>
<td>40</td>
</tr>
<tr>
<td>B</td>
<td>108</td>
<td>0740</td>
<td>0810</td>
<td>30</td>
</tr>
<tr>
<td>C</td>
<td>109</td>
<td>0740</td>
<td>0820</td>
<td>40</td>
</tr>
<tr>
<td>A</td>
<td>110</td>
<td>0830</td>
<td>0830</td>
<td>0</td>
</tr>
</tbody>
</table>

| Total   |       |      |      | 220   |

Table 3.1: Ghost Flight Delay Example.

more incentive to "game" the system, as described in the next section.

### 3.3 Ghost Flight Cancellation Problem

The current method of allocating capacity to airlines does so according to each airline's flights in the published schedule. If a user schedules flights it never intends to fly, it gains slots in a ground delay program because it is over-represented in the schedule. Under previous, more restrictive FAA assignment methods, the advantage to a user of having slots available to be canceled was not as great. An example illustrates the problem.

Consider the example given in [37], shown in Tables 3.1 and 3.2. This example was used in [37] to demonstrate Compression, the method used by CDM to remove the disincentives to airlines of reporting cancellations. It works equally well, however, to demonstrate the incentive a user has to add flights to the schedule that it does not intend to use.

The columns in Table 3.1 are airline, flight number, scheduled slot time (ETA provided by the airlines), first-scheduled first-served (FSFS) slot time (controlled time of arrival, or CTA), and amount of delay. The FSFS slot time is the slot that would be assigned by the FAA to the flight in the absence of cancellations and substitutions. Notice that airline A incurs 40 minutes of delay, all on flight 7. However, by adding
<table>
<thead>
<tr>
<th>Airline</th>
<th>FLTNO</th>
<th>ETA</th>
<th>CTA</th>
<th>Delay</th>
<th>Can/Sub</th>
<th>New CTA</th>
<th>Rdly</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100</td>
<td>0700</td>
<td>0700</td>
<td>0</td>
<td>C</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A</td>
<td>101</td>
<td>0700</td>
<td>0710</td>
<td>10</td>
<td>S</td>
<td>0720</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>102</td>
<td>0705</td>
<td>0720</td>
<td>15</td>
<td></td>
<td>0740</td>
<td>15</td>
</tr>
<tr>
<td>B</td>
<td>103</td>
<td>0705</td>
<td>0730</td>
<td>25</td>
<td></td>
<td>0730</td>
<td>25</td>
</tr>
<tr>
<td>B</td>
<td>104</td>
<td>0710</td>
<td>0740</td>
<td>30</td>
<td></td>
<td>0740</td>
<td>30</td>
</tr>
<tr>
<td>B</td>
<td>105</td>
<td>0710</td>
<td>0750</td>
<td>40</td>
<td></td>
<td>0750</td>
<td>40</td>
</tr>
<tr>
<td>A</td>
<td>106</td>
<td>0710</td>
<td>0800</td>
<td>50</td>
<td>S</td>
<td>0710</td>
<td>0</td>
</tr>
<tr>
<td>C</td>
<td>107</td>
<td>0720</td>
<td>0810</td>
<td>50</td>
<td></td>
<td>0800</td>
<td>40</td>
</tr>
<tr>
<td>B</td>
<td>108</td>
<td>0740</td>
<td>0820</td>
<td>40</td>
<td></td>
<td>0810</td>
<td>30</td>
</tr>
<tr>
<td>C</td>
<td>109</td>
<td>0740</td>
<td>0830</td>
<td>50</td>
<td></td>
<td>0820</td>
<td>40</td>
</tr>
<tr>
<td>A</td>
<td>110</td>
<td>0830</td>
<td>0840</td>
<td>10</td>
<td></td>
<td>0830</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>320</td>
<td></td>
<td></td>
<td>220</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Imaginary Flight Cancellation under CDM.

a nonexistent flight to the reported schedule, airline A can effect a change in the allocation to that shown in Table 3.2.

Table 3.2 shows the revised allocation of slots to flights after airline A has introduced imaginary flight 100 and then canceled it. The columns, from left to right, are the airline, flight number, scheduled arrival time, initial allocated arrival time, delay incurred, flight canceled / substituted, arrival time allocated after running the compression algorithm, and resultant delay. Notice that airline A has succeeded in shifting its portion of the delay to its competitors.

This example demonstrates the motivation for a user to schedule and then cancel flights. The CDM working group has emphasized enforcement means to ensure that the users report accurate schedules. Schedule fabrication can be difficult to detect, however. An airline might schedule two flights between the same pair of cities within a relatively short period of time and fly them both on days with adequate demand, while canceling one or the other on days with light passenger demand. If the airline seldom flew both flights on the same day, it would clearly be gaining an advantage from the practice. Yet, it might be argued that the practice of adjusting the set of flights to be used at the last minute to accommodate uncertainties such as airport capacity and passenger demand or to deal with equipment or crew scheduling problems is in everyone’s best interest. The airline can serve more passengers that way when demand
warrants, and when demand does not warrant, the system is saved the expense of serving the extra flight. For instance, imagine that Logan Airport in Boston has a low arrival rate due to foggy conditions and that an airline with two lightly loaded flights destined for Boston during the morning rush hour could consolidate them. Doing so would help to alleviate the capacity problem at Logan and save the airline money.

The task of determining whether this practice is legitimate would inevitably involve a somewhat arbitrary decision that could inflict a cost on the overall system. The users would be compelled by their economic objectives to exploit loopholes in the regulation and in doing so could hurt the airline industry as a whole.

A related shortcoming of the current method of running ground delay programs is that, although each airline may optimize the use of its set of slots once the set has been allocated, no attempt is made to optimize the allocation of sets of slots to users. Optimization-based models for solving the FAA’s decision problems related to air traffic flow management have not fared well despite the efforts of numerous researchers. The models presented in [33, 27, 35, 36, 28, 17] among many others, each solve different aspects of the ground holding problem. The model presented in [7] addresses the ground holding problem combined with the routing problem. Despite the considerable potential of models such as these to improve operations in the air transportation system, none of them has been used in practice to date.

One impediment to the implementation of these models is that they require objective function parameters. As with all optimization problems, these models may improve operation of a system only to the extent that the objective functions accurately reflect the system objectives. When such methods are applied for operations within a single company, the company is motivated by its economic objectives to provide accurate data to the model. Unfortunately, application of optimization models to many of the problems faced by the FAA requires accurate information from organizations with interests that often conflict, complicating the data collection problem. For instance, most of the ground holding problem formulations rely on a parameter specifying the relative cost of airborne holding to ground holding. Gaining industry
acceptance of these formulations would likely require the airlines to agree on a value of this parameter, since allowing the parameter to differ across airlines would give each airline incentive to provide false information in order to gain an advantage in the ground delay program.

The challenge for the FAA is to find a way to gather objective function information from the users of the NAS and to combine the information to find solutions that all users can agree are mutually optimal. The information must be used in a way that motivates the users to provide accurate information; in the language of game theory, the FAA's solutions must exhibit *incentive compatibility* (see, for example, [14]). Milner has proposed a method of assigning arrival slots to flights that treats the incentive compatibility issue [23], discussed in Section 3.4. Section 3.5 presents a new method that extends Milner's method to work dynamically and relaxes an assumption made by Milner.

### 3.4 Slot Auction

In [23], an auction of arrival slots is proposed. The auction is held at the beginning of each day and follows rules described in [29] and outlined here:

1. Each flight \( j \in J \) reports the economic value \( v_{ij} \) it could produce through use of slot \( i \), \( \forall i \in I \).

2. The FAA solves a two-sided matching problem to maximize economic value produced through the assignment of slots to flights.

3. Each flight pays for the use of its slot a fee equal to the marginal price of the slot, i.e., the value of the dual variable associated with the constraints allowing a slot to be assigned to at most one flight.

Clearly, if each user's value information \( v_{ij} \) were accurate, then the problem solved by the FAA would indeed maximize economic value. The question is whether this is possible. There are two aspects to this question. First, the economic objectives may
be significantly more complicated than can be represented in this manner: the value of a flight arriving at a specific time generally is a function of the times when other flights arrive and depart, i.e., the \( v_{ij} \) are not independent. Second, assuming that the \( v_{ij} \) objective structure can be made to represent the true objectives accurately, the users may have incentive to misrepresent the truth. The fees paid by each user in step 3 address this second concern; because of the fees, the users have incentive to provide their objectives accurately through the reported values \( v_{ij} \), as shown in [23].

Consider an example in which several users would each like to use the only slot available. The marginal price for the slot becomes the highest value that any of the users losing the auction could have produced with the slot. Thus, the highest bidder wins the slot and pays the second-highest price. This is the famous second-price auction [34].

In the second-price auction, each user may report a value \( v_{ij} \) that is too low, too high or accurate. A user who reports an accurate value either receives the slot at a price less than its value or does not receive it because its price is greater than its value. If the user reports a value below the true value, the user will not receive the slot for any less money than it could by reporting truthfully, but it runs the risk of losing the slot and associated profit. If the user is able to receive the slot only by reporting a value above its true value, the price of the slot will be above the true value of the slot. The user stands only to lose by reporting inaccurate value information.

Under Milner’s method, this incentive compatibility property holds only if flights do not form coalitions. A coalition of flights is a maximal cardinality group of flights that cooperate to optimize their collective results from the assignment process. A coalition of flights operating under Milner’s allocation method has incentive to report artificially low values \( v_{ij} \) in order to lower the fees it must pay. Since the air transportation system is dominated by users operating multiple flights, the assumption that coalitions of flights do not exist is unrealistic.

Consider the simple example depicted in Figure 3-1. Airline 1 has a single flight which can use any of the three slots available at a value of 90. Airline 2 has three flights, each of which produces a value of 100 through the use of a slot. Under Milner’s
method, truthful reporting would assign all slots to airline 2, albeit at a price of 270 units, for a profit of a mere 30 units. Clearly, airline 2 could do better by understating its slot value, allowing airline 1 to take one of the slots and receiving the other two slots for free for a profit of 200 units. The problem here is that a single coalition may pay multiple times for the loss of productivity it causes to the system. In the example of Figure 3-1, airline 2 paid a fee of 90 three times for causing airline 1 to lose 90 units of productivity once. Indeed, a single coalition must pay for the lost productivity its flights cause each other; had the flight owned by airline 1 been instead a flight of airline 2, truthful reporting would still have cost airline 2 270 units in fees.

Another difficulty with the approach of charging a fee equal to the marginal price for a good is that it fails to provide incentive compatibility in problems exhibiting a duality gap. For instance, consider the problem below in which there are two slots available. User A can produce a value of three through the use of both slots, but has no use for a single slot. Thus, variable \( x_{A12} \) is defined to be one only if both slots are given to user A. User B requires either one of the slots to produce a value of two. Accordingly, variable \( x_{B1} \) takes a value of one if user B receives slot one, and \( x_{B2} \) takes a value of one if user B receives slot two. The optimal LP relaxation of the problem:

Maximize \[ 3x_{A12} + 2x_{B1} + 2x_{B2} \]
Such that:

\[ x_{B1} + x_{B2} \leq 1 \]
\[ x_{A12} + x_{B1} \leq 1 \]
\[ x_{A12} + x_{B2} \leq 1 \]
\[ x_{A12}, x_{B1}, x_{B2} \in \{0, 1\} \]

has an optimal objective value of 3.5 with all variables equal to 1/2, whereas the optimal integral solution has value 3 with \( x_{A12} = 1 \) and the other variables zero. If the optimal integral solution were found and user A were assessed the marginal cost of slots one and two, it would pay a fee of 4 and lose money. In this case, user A would do better to understate its slot values and let user B have a slot, reducing the global economic value produced by one third.

The Groves mechanism [16] avoids these problems essentially by defining a dual for which the fee assessed a user is a function only of the good received by the user and for which there is never any duality gap (there is no magic here computationally: solving this dual requires solving the original IP). The Groves mechanism assesses each user a fee equal to the lost value that user causes everyone else through its presence in the solution. In the slot allocation context, the fee assessed a user is the objective value that would be achieved by the rest of the users if the user in question received no slots, minus the value the rest of the users achieved in the globally optimal solution. For instance, in the example above, user A would pay a fee of 2 for its slots, since the optimal solution to the problem without user A present is 2 and in the globally optimal solution the users other than A produce no value. User A’s objective under the Groves mechanism is to maximize its value produced minus fees paid:

Maximize \[ 3x_{A12} - (V - (2x_{B1} + 2x_{B2})) \]

Such that:

\[ x_{B1} + x_{B2} \leq 1 \]
\[ x_{A12} + x_{B1} \leq 1 \]
\[ x_{A12} + x_{B2} \leq 1 \]
\[ x_{A12}, x_{B1}, x_{B2} \in \{0, 1\} \]
where the decision variables are now functions of the values reported by user A, and 
$V$ is the objective value of the problem with user A removed. Note that the objective 
of this problem can be rewritten:

$$\text{Maximize} \quad 3x_{A12} + 2x_{B1} + 2x_{B2} - V$$

Since $V$ is a constant with respect to user A’s reported values, it can be removed 
from the objective without affecting the set of optimal solutions. With $V$ removed 
from the objective, user A is in fact trying to maximize the global objective. Since 
the FAA has this same objective, user A can achieve its objectives by reporting them 
truthfully.

The technique outlined here to demonstrate the Groves mechanism’s incentive 
compatibility is extremely general. If each user is charged the amount that his 
actions cost the others, his overall objective becomes that of maximizing the pro-
ductivity of all users. One could indeed use the same technique to formulate an 
incentive-compatible mechanism for distributing capacity among the users in a dy-
namic environment under uncertainty through the use of dynamic programming to 
achieve the globally optimal strategy in the expected value sense. Applying such an 
approach to most problems would be impractical, however, because it would require 
the users to communicate to the FAA the expected value of every state in the dynamic 
program’s potentially infinite state space and the state transitions in order to deter-
mine the price to assess each user. Moreover, the dynamic programming problems 
requiring solution in order to implement the method might be intractable.

A mechanism that allows dynamic updates from the users to accommodate un-
certainty without sacrificing computational tractability is presented in Section 3.5. 
It is an application of the Groves mechanism to a dynamic programming problem 
with a limited state space. The auction occurs over time as the characteristics of the 
slots being distributed are determined and as their assignment becomes necessary, as 
would happen under a general dynamic programming approach to the problem. By 
limiting the state space, the information to be collected and the problems to be solved
at each stage are reduced to a very reasonable size, yet the assumptions required for Milner’s method are relaxed.

3.5 Objective-Based Allocation Method

This section introduces the Objective-Based Allocation Method (OBAM) for assigning arrival slots to flights based on the user objective functions. In assigning slots, OBAM attempts to maximize the value produced collectively by the users. Although OBAM does not specifically treat uncertainty, it is a dynamic mechanism that allows the users and FAA to update information as it becomes available in order to accommodate uncertainty.

Let arrival slot with label $i$ be the $i^{th}$ slot in sequence at a given airport. Define $m_i$ to be the time at which slot $i$ would begin were the airport to operate with the highest possible arrival capacity under perfect weather conditions. Let $t_i$ be the time at which the FAA expects slot $i$ to occur. Whenever the FAA updates its estimates of slot times, it provides the $t_i$ to the users.

A flight awarded use of a slot is able to produce a transportation service for its customers using that slot. The value of the slot to the flight may not be known exactly a priori, since the precise time at which the slot will occur is not known ahead of time, and the value of a flight depends to a large extent upon its time of arrival. Slot timing is uncertain both because airport capacity cannot be forecast accurately and because realized demand is uncertain.

OBAM employs a fixed planning horizon, the amount of time it plans into the future. Let $T$ represent the current time, and let $H$ represent the length of the planning horizon. At time $T$, the method considers the assignment of unassigned slots $\{i : T \leq m_i \leq T + H\}$. For instance, if $H$ is 48 hours, all slots that would nominally occur within two days will be considered by the method.

The method, whose steps are presented below in detail, represents the user preferences by a set of values $v_{ij}$, each of which is interpreted as the value that can be produced by an assignment of slot $i$ to flight $j$. Whereas this objective structure is
limited in that it does not allow users to represent combinatorial or stochastic objectives directly, the dynamic nature of the method allows users to compensate for this shortcoming. For instance, a user for whom certain flights have value dependent on other flight-slot assignments could submit the expected value conditioned on the expected flight-slot assignments. Because regular updates would be provided by the FAA under OBAM, the user's prediction of flight-slot assignments should be quite accurate.

The method also uses a \textit{commitment time} for each possible slot-to-flight assignment. The commitment time, $q_{ij}$, is the time by which flight $j$ must be notified of an assignment to slot $i$ if such an assignment is made. It is possible that a user might be able to realize greater value by an earlier commitment time, but still find later commitment times feasible. This information is not used explicitly in this formulation; however, in such an instance the user could submit $q_{ij}$ with corresponding value dynamically.

A characteristic typical of airport demand patterns is that the demand decreases during the night hours. Overnight landing capacity is almost always adequate to handle the demand and to allow any aircraft that were unable to land during the day to do so. The calculation of fees to be paid by the users under OBAM is based on the assumption that such excess capacity exists from time to time. In order to calculate the slot fees, the slot allocation problem is decomposed a posteriori into a series of problems in time. The fees are calculated over one problem at a time.

To be precise, let the term \textit{slack point} refer to a point in time around which there is capacity sufficient so that any assignment of slots to flights that maximizes the sum of values achieved by the flights results in each flight receiving a slot with higher value to it than the value of any slot on the other side of the slack point. For instance, if an assignment of slots to flights that maximizes collective value awards flight $j$ a slot earlier than the slack point, the value to flight $j$ of that slot is higher than the value of any slot after the slack point. The slack points form the boundary points in time between the subproblems into which the overall capacity allocation problem is decomposed. Because excess capacity exists around slack points and because the
temporal extent of each flight's desired slots is limited, the allocation of a user's slots in one period does not affect the other users' costs in the following period, and the fees can be calculated for each period independently. In practice, it is likely that slack points meeting this definition would occur almost every night, since user value functions are roughly unimodal and of duration less than a day. Because the slack points are used to assess fees a posteriori, it is not necessary to predict their occurrence; it is only necessary to recognize slack points after the fact. However, if by some means the regularity of slack points can be determined in advance, that information may be used to determine an upper bound on the length of time horizon that may be desired for OBAM; using a longer time horizon provides no benefit.

Denote the earliest slack point that is after the latest slot time \( t_i \) for which some flight has a commitment time before or at the current time \( T \) as the horizon slack point (see Figure 3-2). By definition of the horizon slack point, no decisions are required at the current time for slots later than the horizon slack point, because all commitment times for those slots are in the future.

If the horizon slack point is within the planning horizon and every flight with nonzero values \( v_{ij} \) for slots \( i \) before the horizon slack point has no nonzero values later than \( T + H \), then OBAM will behave exactly as it would were there an infinite planning horizon. Figure 3-2 depicts this situation. In the figure, each horizontal bar represents the temporal characteristics of a potential slot-flight assignment: the left end of the bar represents the commitment time \( (q_{ij}) \), and the right end of the bar represents the slot time \( (t_i) \). Each flight would have one of these bars for each slot to which it assigned positive value \( v_{ij} \). The definition of the horizon slack point requires that every bar with its left end at or to the left of \( T \) must have its right end to the left of the horizon slack point. If the time horizon is long enough so that every flight with a bar ending to the left of the horizon slack point does not have a bar ending to the right of the horizon end (time \( T + H \)), then the time horizon is long enough to take into account all values of every flight that might be assigned a slot before the horizon slack point. By the definition of a slack point, each flight can be constrained to land before the slack point if its maximal slot value occurs before the slack point,
and it can be constrained to land after the slack point otherwise, without sacrificing any value. The infinite-horizon problem is thus separated into a series of independent problems covering the finite time between slack points.

Let \( A = \{J_1, J_2, \ldots, J_a\} \) represent the set of coalitions. For example, \( J_1 \) might be the flights belonging to American Airlines at the airport, \( J_2 \) might be flights belonging to Continental Airlines, and so on. Let \( J = \bigcup_{i=1}^{a} J_i \), be the set of all flights. Note that by definition, \( J_i \cap J_j = \emptyset, \forall i \neq j \).

OBAM follows the steps given below.

1. Initialize \( I^f \), the set of indices of slots that have been assigned by the method, to \( \emptyset \).

2. Set \( I \) to the indices of all slots to be assigned over the planning horizon, that is,

\[
I = \{i : T \leq m_i \leq T + H\} \setminus I^f
\]

3. Report to the users the estimates of times \( t_i, \forall i \in I \).

4. Accept input from flights. Each flight \( j \) reports the following data:

(a) For each slot \( i \in I \), a value \( v_{ij} \) representing the additional economic value of an assignment of slot \( i \) to flight \( j \) beyond the value of a solution in which flight \( j \) is canceled.
(b) Required commitment time, \( q_{ij} \), for each slot \( i \) and flight \( j \). The commitment time is the time by which the flight must be notified of an assignment to slot \( i \).

There are at most \( 2|I| \) quantities that need to be reported for each flight, a value \( v_{ij} \) and a value \( q_{ij} \) for each slot in \( I \). There could be fewer quantities if unreported values defaulted to zero.

5. Continue accepting input, and wait for the next time \( q_{ij} \) to occur. If, while waiting, a new slot enters the time horizon \((\exists i \notin \{I \cup J\} : T \leq m_i \leq T + H)\), go to step 2.

6. Solve the assignment problem:

\[
IZ(T) = \max_{x \in AP(I,J)} \sum_{j \in J} \sum_{i \in I} v_{ij} x_{ij}
\]

where the set \( AP(I,J) \) is the set of all \( x \) satisfying

\[
\sum_{j \in J} x_{ij} \leq 1 \forall i \in I
\]

\[
\sum_{i \in I} x_{ij} \leq 1 \forall j \in J
\]

\[x_{ij} \in \{0, 1\}\]

\( x \) is a vector of binary decision variables representing the assignment of slot \( i \) to flight \( j \) when \( x_{ij} \) is equal to 1. Let \( x^* \) represent the optimal solution to this problem.

7. Remove slot \( i \) and flight \( j \) from \( I \) and \( J \) respectively, and add \( i \) to \( I' \), for all \( i \) and \( j \) satisfying the conditions:

(a) \( x_{ij}^* = 1 \)

(b) \( q_{ij} = T \)

53
Such assignments will be referred to as final assignments. Assignments that are not final will be referred to as reservations. Final assignments are made at their commitment times and are assigned to the corresponding flight for use. If for some reason a flight $j$ determines that it is unable to use slot $i$ after it has been assigned by a final assignment (as might happen if a mechanical problem with the aircraft were discovered after the assignment was made), slot $i$ is removed from the set $I'$ and returned to the set $I$ so that it may be used by a flight with a shorter commitment time. In this case, slot $i$ is said to have been relinquished, and the coalition to which flight $j$ belongs will be called the relinquisher of slot $i$.

8. Record all $v_{ij}$ for all flights that have received final assignments, and record all $v_{ij}$ for all slots that have been assigned as final assignments. These values will be referred to as final values and denoted $v^f_{ij}$. For slots being assigned in final assignments that were previously assigned and then relinquished, the previously-recorded final values are saved as well. Note that a flight cannot change its value for a slot after it has received a final assignment without relinquishing its assignment, and it cannot change its value for a slot that has already been assigned in a final assignment, because values are only accepted for flights and slots in the sets $I$ and $J$ in step 4.

9. If $T$ is not at a slack point, go to step 2.

10. For each coalition $r \in \{1, 2, \ldots, a\}$ calculate the final assignments to slots with times between $T$ and the previous slack point that would have occurred had coalition $r$ not been present. Let $x^{fr}_{ij}$ represent these final assignments, and let $x^{f0}_{ij}$ represent the final assignments achieved with all coalitions present (i.e., the assignments realized in steps 2 through 9).

11. Define $p(I_r)$ to be the value of $x^{fr}$ according to the final values submitted by the coalitions, minus the value to the coalitions other than $r$ achieved by $x^{f0}$. That is,
\[ p(I_r) = \sum_{j \in J \setminus I_r} \sum_{i \in I} v_{ij}^r x_{ij}^r - v_{ij}^f x_{ij}^f \]  

(3.1)

In the event that a slot is relinquished by a coalition \( r \) and subsequently assigned to another coalition, the \( v_{ij}^r \) used in calculating \( p(I_r) \) in the left hand side of equation (3.1) should be the original final values recorded when the slot was first assigned to \( r \).

12. Assess each coalition \( r \in \{1, 2, \ldots, a\} \) a fee \( p(I_r) \), and go to step 2.

### 3.6 Observations Concerning OBAM

Under the assumption that the values \( v_{ij} \) are an accurate representation of the user preferences and that they are fixed, it is clear that the method maximizes value realized by the community as a whole.

It remains to be shown that the users are motivated to provide accurate information. The proof restricts attention to the times at which assignments are made. Since the flight data reported by each user are not shared with other users, the times at which assignments are made are unknown to the users. There is no motivation to report inaccurate information between assignments, since the values reported have no effect between assignment times. Thus, if there is motivation to report accurate information at assignment times, the users will be motivated to report accurate information at all times. The term *bid history* will be used to refer to the values \( v_{ij} \) in effect at the times at which assignments are made.

The proofs of the following propositions follow the outline described in Section 3.4 for showing incentive compatibility of the Groves mechanism: they show that the objective of each user, when the fee assessed by OBAM is considered in the objective function, is maximized by an accurate report of the user value functions.

**Proposition 1** If each coalition knows its value functions in advance, then the solution in which each coalition accurately reports its value functions forms a Nash
equilibrium\footnote{A Nash equilibrium is defined in [14] as \textit{a profile of strategies such that each player's strategy is an optimal response to the other players' strategies.}}.

\textbf{Proof:} Assume coalition $r$ knows its value functions in advance and that all coalitions other than coalition $r$ report fixed values $v_{ij}$. Consider the result to coalition $r$ of making an accurate bid. Under fixed values $v_{ij}$, OBAM results in an assignment $x^*:\]

$$x^* = \arg \max_{x \in AP(I,J)} \sum_{i \in I} \sum_{j \in J} v_{ij} x_{ij}$$ \hspace{1cm} (3.2)$$

Since all values $v_{ij}$ for all coalitions other than $r$ are fixed in this case, the fee assessed for the set of slots $I_r$, calculated by (3.1), simplifies

$$p(I_r) = \max_{y \in AP(I,J)} \sum_{j \in J} \sum_{i \in I} \left( v_{ij} y_{ij} - v_{ij} x^*_{ij} \right)$$ \hspace{1cm} (3.3)$$

Notice that the fee charged to coalition $r$ is a function only of the assignment $x^*_{ij}$ and of the value functions $v_{ij}$ of other coalitions. It is not a function of the values submitted by the coalition, except indirectly through the assignment. Therefore, the net value to coalition $r$, the value received through the assignment minus the fee assessed for it, is a function only of the assignment.

Assume that there exists an assignment $\hat{x} \in AP(I,J)$ that assigns slots $\hat{I}_r$ to coalition $r$ and is preferable to $x^*$ for coalition $r$, that is,

$$\sum_{i \in I} \sum_{j \in J_r} \hat{v}_{ij} x_{ij} < \sum_{i \in I} \sum_{j \in J_r} v_{ij} x^*_{ij} - p(I_r)$$

Using equation (3.1) to substitute for $p_{I_r}$ and $p(I_r)$ gives:

$$\sum_{i \in I} \sum_{j \in J_r} \hat{v}_{ij} x_{ij} - \max_{y \in AP(I,J)} \sum_{j \in J \setminus J_r} \sum_{i \in I} \left( v_{ij} y_{ij} - v_{ij} \hat{x}_{ij} \right)$$

$$> \sum_{i \in I} \sum_{j \in J_r} v_{ij} x^*_{ij} - \max_{y \in AP(I,J)} \sum_{j \in J \setminus J_r} \sum_{i \in I} \left( v_{ij} y_{ij} - v_{ij} x^*_{ij} \right)$$

\footnote{A Nash equilibrium is defined in [14] as \textit{a profile of strategies such that each player's strategy is an optimal response to the other players' strategies.}}
Since the maximization over \( y \) and over \( \hat{y} \) are on the same cost function and over the same feasible region, the objective values realized are equal and their sums cancel. The equation simplifies to:

\[
\sum_{i \in I} \sum_{j \in J_r} v_{ij} \hat{x}_{ij} + \sum_{j \in J \setminus J_r} \sum_{i \in I} v_{ij} \hat{x}_{ij} > \sum_{i \in I} \sum_{j \in J_r} v_{ij} x^*_{ij} + \sum_{j \in J \setminus J_r} \sum_{i \in I} v_{ij} x^*_{ij}
\]

Or, more simply,

\[
\sum_{i \in I} \sum_{j \in J} v_{ij} \hat{x}_{ij} > \sum_{i \in I} \sum_{j \in J} v_{ij} x^*_{ij}
\]

But this can only be true if \( x^* \) is not the optimal solution, violating (3.2). The contradiction proves that the optimal bid history for the coalition is the accurate one. Therefore, the proposition holds. \( \square \)

**Proposition 2** If all coalitions other than \( r \) hold their bids constant, it is optimal for coalition \( r \) to correct any errors in its bid immediately.

**Proof:** Consider the situation in which all coalitions other than \( r \) submit fixed bids and in which coalition \( r \) has made arbitrary bids until time \( T \). For purposes of this proof, denote the set of slots that have not been assigned in final assignments by time \( T \) as \( I(T) \). Let \( I^f(T) \) denote the set of slots that have been assigned in final assignments by time \( T \), and let \( I = I^f(T) \cup I(T) \). A fixed, accurate bid from time \( T \) onward by coalition \( r \) results in an assignment \( x^* \in AP(I, J) \):

\[
x^* = \arg \max_x \sum_{i \in I^f(T)} \sum_{j \in J} v_{ij} x_{ij} \quad (3.4)
\]

Let \( I_r(T) \) represent the set of slots assigned to coalition \( r \) after time \( T \). Assume there exists a bid strategy that coalition \( r \) could employ at time \( T \) that would have a result preferable in the view of the coalition to the result achieved through an accurate bid. That is, assume that there exists \( \hat{x} \in AP(I, J) \) which awards slots \( \hat{I}_r(T) \) after time \( T \) to the coalition such that

57
\[
\sum_{i \in I} \sum_{j \in J_r} v_{ij} \hat{x}_{ij} - p_{I_r} > \sum_{i \in I} \sum_{j \in J_r} v_{ij} x_{ij}^* - p(I_r)
\]

and

\[
\hat{x}_{ij} = x_{ij}^*, \quad \forall i \in I^I(T)
\]  \hspace{1cm} (3.5)

Equation (3.5) constrains \( \hat{x} \) to be an assignment that respects the assignments made final before time \( T \). Separating the parts of the assignment made final before and after \( T \), and applying equation (3.3) for \( p_{I_r} \) and \( p(I_r) \) gives

\[
\sum_{i \in I^I(T)} \sum_{j \in J_r} v_{ij} \hat{x}_{ij} + \sum_{i \in I(T)} \sum_{j \in J_r} v_{ij} \hat{x}_{ij}
\]

\[
- \max_{y \in AP(I,J)} \sum_{j \in J \setminus J_r} \left( \sum_{i \in I^I(T)} (v_{ij} y_{ij} - v_{ij} \hat{x}_{ij}) + \sum_{i \in I(T)} (v_{ij} \hat{y}_{ij} - v_{ij} \hat{x}_{ij}) \right)
\]

\[
> \sum_{i \in I^I(T)} \sum_{j \in J_r} v_{ij} x_{ij}^* + \sum_{i \in I(T)} \sum_{j \in J_r} v_{ij} x_{ij}^*
\]

\[
- \max_{y \in AP(I,J)} \sum_{j \in J \setminus J_r} \left( \sum_{i \in I^I(T)} (v_{ij} y_{ij} - v_{ij} x_{ij}^*) + \sum_{i \in I(T)} (v_{ij} \hat{y}_{ij} - v_{ij} x_{ij}^*) \right)
\]

By equation (3.5), the sums involving \( \hat{x}_{ij} \) and \( x_{ij}^* \) over the set \( I^I(T) \) cancel, leaving

\[
\sum_{i \in I^I(T)} \sum_{j \in J_r} \hat{y}_{ij} x_{ij} - \max_{y \in AP(I,J)} \sum_{j \in J \setminus J_r} \left( \sum_{i \in I^I(T)} v_{ij} \hat{y}_{ij} + \sum_{i \in I(T)} (v_{ij} \hat{y}_{ij} - v_{ij} \hat{x}_{ij}) \right)
\]

\[
> \sum_{i \in I(T)} \sum_{j \in J_r} v_{ij} x_{ij}^* - \max_{y \in AP(I,J)} \sum_{j \in J \setminus J_r} \left( \sum_{i \in I^I(T)} v_{ij} y_{ij} + \sum_{i \in I(T)} (v_{ij} y_{ij} - v_{ij} x_{ij}^*) \right)
\]

or,

\[
\sum_{i \in I(T)} \sum_{j \in J_r} v_{ij} \hat{x}_{ij} + \max_{y \in AP(I,J)} \sum_{j \in J \setminus J_r} \left( \sum_{i \in I^I(T)} v_{ij} \hat{x}_{ij} - \sum_{i \in I} v_{ij} \hat{y}_{ij} \right)
\]
\[ > \sum_{i \in I(T)} \sum_{j \in J_r} v_{ij} \bar{x}_{ij}^* + \max_{y \in AP(I,J)} \sum_{j \in J \setminus J_r} \left( \sum_{i \in I(T)} v_{ij} \bar{x}_{ij}^* - \sum_{i \in I} v_{ij} y_{ij} \right) \]

Again, the maximizations over \( y \) and \( \bar{y} \) use the same cost functions over the same feasible sets, so they cancel, leaving

\[ \sum_{i \in I(T)} \sum_{j \in J} \bar{v}_{ij} x_{ij} > \sum_{i \in I(T)} \sum_{j \in J} v_{ij} x_{ij}^* \]

But this contradicts equation (3.4). Therefore, coalition \( r \) can do no better than to make an accurate bid immediately, and the proposition holds. \( \square \)

Of course, these results do not mean that the Nash equilibrium solution would occur in practice. For instance, if a coalition knew that future slots would be scarcer than was generally thought, that coalition could improve its solution by lowering the value it placed on the scarce slots. In this way, it would limit the reservations it would hold during the period in which slot prices were likely to increase. Conversely, if it knew that future slots would be more abundant than generally thought, it could lower the value placed on the slots in the near future in order to receive cheaper slots later. In effect, OBAM allows coalitions to use proprietary information to improve the results achieved. If OBAM were implemented, it would encourage the users of the NAS to improve their forecasting capabilities. Notice that by using its proprietary information, if that information is correct, the coalition improves not only its results, but it improves the overall results as well by lowering demand for the congested slots.

3.7 Pros and Cons of OBAM-Like Slot Assignment Methods

There are several benefits to the objective-based approach to allocating arrival slots. The current system of assigning slots to flights rewards a user for scheduling and then canceling flights it never intends to fly; the CDM group has proposed regulation and penalties for such acts, however there is no clear way to differentiate intentional abuse of the system from legitimate operations. Regulations currently in place to prevent
abuse of the landing slot system at the four permanently slotted US airports have the disadvantage that they prevent airlines from making legitimate cancellations, wasting airline and airport resources. None of these problems occurs under the objective-based allocation method: a user who schedules and cancels a ghost flight pays for the lost productivity that act causes other users and achieves no overall improvement in profit from it. Furthermore, a user who realizes that its schedule is not profitable can cancel flights; Proposition 2 indicates that it is in the user's best interest to inform the FAA of the true situation as soon as possible.

Another benefit of the objective-based allocation method is that it takes into account user preference information. Current operations distribute slots among users with the goal of minimizing aircraft-minutes of delay. OBAM allows users to specify their own criteria to be used in part or all of the slot distribution. For instance, an airline that prefers to minimize passenger-minutes of delay, late baggage costs, missed passenger connections, fuel and crew costs and any other important criteria can do so. Measured in these terms, the performance of the NAS could be higher than it is currently, especially during periods of reduced airport capacity. Improving system performance is becoming increasingly important now with growth of demand for use of the NAS far out-pacing increases in its capacity.

OBAM motivates the users to provide preference information, which may have value in and of itself. The preference information could be used to improve the objective functions of other tools such as CTAS and FAST, further improving the performance of the NAS. It could also be used to provide an accurate quantification of the costs of delay and of the economic benefit to capacity-increasing technology. This quantification could be used as a basis for policy and investment decisions, improving the benefit-to-cost ratio of infrastructure and other improvement projects undertaken.

The fundamental principle behind OBAM is the coordination of user subproblem solutions through the introduction of a penalty function into the subproblem objectives, much as is done in decomposition-based optimization methods. In practice, this technique requires the users to pay fees for the slots they receive. Several approaches
could be employed to avoid such fees.

One approach that would avoid fees would be to run both the FAA’s existing allocation method and OBAM as follows. First, the FAA’s existing allocation method would be run, and the slots assigned in the process would be given to the users. Then, OBAM could be run, charging real money for the slots assigned and giving the money collected for each slot to the owner of the slot. This approach has the advantage that a user uncomfortable with OBAM could specify a very high value on each of its FAA-assigned slots and zero value elsewhere and would as a result receive its FAA-assigned slots at no cost (since OBAM would refund any costs to the user). Users desiring to participate in OBAM would be able to do so by submitting bids reflective of their true slot values and would benefit by doing so, according to the analysis of Section 3.6; the resulting allocation might cost the user money, but the cost would be more than offset by the improvement in the value of the slots received.

The benefit of this approach is that it relies on the FAA’s existing slot allocation method, which has been accepted by the users, but allows further optimization of the allocation through a market. The drawback is that the existing allocation method assigns slots to a user in proportion to the user’s representation in the schedule. It would be in a user’s interest to find ways to schedule ghost flights.

Finally, the long-term gaming issues involved in OBAM merit further study. The proofs of Section 3.6 assumed that each user’s objective was to maximize its own production minus the fees it was assessed. It is possible that a user might receive long-term benefit from causing its competitors to pay high fees by bidding for slots for which it had no use. For instance, a wealthy user might be able to drive competitors out of its markets by driving up the cost of slots. The financial resources required to do this would be enormous: the user would have to pay the amount of the lost revenue of its competitor. Furthermore, the strategy would only work if it could be accomplished routinely, yet the behavior could be detected and prevented through regulation.

These characteristics inherent to OBAM may prevent its use. The airline industry is wary of any new means of taxation, and OBAM may be viewed as a means to
introduce new taxes. There have been many attempts recently to introduce user fees that would be charged according to use of air traffic control services, and these attempts have proven unsuccessful. Although OBAM offers a user value greater than the fees charged, it may face obstacles to adoption similar to those seen by user fees.

Rather than focus on the technical improvements that could be made to OBAM, the next chapter takes an approach that does not require the users to pay fees. Instead, the focus of the next chapter is on improving the modeling of capacity in order to improve ground delay program performance while decentralizing additional aspects of the ground delay program decision-making. Through this additional decentralization, the users are able to use internal information to make some decisions that are currently made by the FAA without the benefit of the users' internal information.
Chapter 4

Allocation of Arrival-Departure Airport Capacity to Users

The number of arrivals that an airport can accommodate in a given period of time depends, in general, on the number of aircraft that depart during that period. This characteristic is widely known, demonstrated empirically with historical ETMS data collected by the Volpe Center in [15], reflected by the Engineered Performance Standards used by the FAA to understand the hourly arrival rate achievable as a function of hourly departure rate, and recognized by controllers and airline personnel alike.

The underlying reason for the dependence of arrival capacity on departures is that arrivals and departures compete for some of the same resources (see [18] for a discussion of the constrained resources involved). For instance, Figure 4-1 shows the flow of arrivals and departures at Boston’s Logan Airport (BOS) in the 22L-27, 22L-22R configuration. In this configuration, aircraft arriving on runways 22L and 27 must cross departure runway 22R to reach the terminal. Every time an arrival is to cross 22R, departures on 22R must be stopped until the runway crossing is complete. When the airport is under pressure to operate both arrivals and departures in this configuration, an increase in the number of arrivals requires a decrease in the number of departures and vice versa. If controllers use runway 22L for departures as well as arrivals, the number of departures that can be served is increased at the expense of the arrivals. On the other hand, if departure demand is significantly lower than
the capacity of runway 22R, there is little tradeoff between arrival and departure operations: 22L and 22R are not spaced sufficiently far apart to allow simultaneous instrument approaches, so departure runway 22R cannot be used efficiently for arrivals. Furthermore, with departures using less than the full capacity of 22R, arrivals can cross runway 22R in the naturally-occurring gaps in departure traffic.

Tradeoffs of this type can be represented by an *arrival-departure capacity curve* as shown in Figure 4-2. Each configuration at each airport has a characteristic arrival-departure capacity curve; the degree of the arrival-departure interdependence varies across airports and configurations. For instance, strong northwesterly winds force BOS to operate in the 33L-33R, 33L configuration in which nearly all arrivals and departures must share runway 33L, resulting in a much greater interdependence than is exhibited in the 22L-27, 22R configuration. On the other hand, Dallas-Fort Worth International Airport (DFW), built on over 17,600 acres of land and consisting of seven runways (the figure includes one runway that is planned but that has not yet been built), has enough runway and taxiway capacity in its primary configuration
to prevent significant competition of arrivals and departures for the same resources (Figure 4-3). Each runway typically serves arrival operations or departure operations exclusively. Operating practice, regulation, and the geometry of the runway-taxiway system often obviate the types of tradeoffs common at BOS.

Occasionally even an airport with extensive runway resources does exhibit an interdependence between arrivals and departures from time to time. For instance, Figure 4-4 shows radar returns from the March 7, 1998 storm system that swept across the southern United States. The intensity of this storm was enough to limit airborne capacity. The effect this had on the Atlanta-Hartsfield International Airport (ATL), an airport with independent, dedicated arrival and departure runways, is shown in Figure 4-5. Each graph in that figure shows arrivals (on the vertical axis) and departures (on the horizontal axis) scheduled over the ten-minute period beginning at the time indicated below the graph. The colored blue dot connected by a thin blue line to the origin of each graph represents the arrivals and departures of Delta Airlines, and the grey dot represents arrivals and departures of all airlines combined. The top series of graphs in that figure shows the arrivals and departures scheduled to operate in each of six ten-minute periods from 17:00 to 18:00 UTC. The bottom series of graphs shows the actual counts of arrivals and departures over the same period. The shaded triangular area on each graph serves as a visual reference, and is identical.
Figure 4-3: Dallas-Fort Worth International Airport

on each graph.

Notice that the operations scheduled at ATL from 17:00 to 17:50 UTC extend significantly above the grey region, whereas the observed operations fall very close to the boundary of the grey region. Although the author did not have access to data that might prove the conjecture that the true airport system capacity over the period was that represented by the shaded region on the graphs, this data provides strong anecdotal evidence in support of the conjecture, and it provides a nice illustration of an effect that is widely reported by the operators of the air transportation system.

Another case of arrival-departure interaction in the air transportation system involves arrivals at one airport affecting departures at another, and vice-versa. For instance, this interaction is known to occur between the geographically close, major airports in the New York metropolitan area. Although the primary focus of this thesis is on single airports, the techniques developed herein could be applied to groups of airports for which a joint capacity graph (possibly over many more than two dimensions) could be defined.
Figure 4-4: National Radar at 17:00 UTC on 7 March 1998

Atlanta Scheduled vs Realized Arrivals & Departures
03/07/98 Storm System

17:00 - 18:00 Schedule Counts (top) vs Actual Counts (bottom) in
10-minute intervals.

black hash: 10 operations
blue: Delta operations
grey: All operations

Figure 4-5: Scheduled vs Actual Operations at ATL from 17:00 to 18:00 UTC on 7
March 1998
4.1 Problems with Arrival Capacity Allocation

When running a Ground Delay Program, it is important that accurate estimates of airport arrival capacity be used. Errors in the airport arrival capacity estimates result in wasted airport capacity or increased airborne holding. When airport capacity is underestimated, flights are held on the ground at their origins unnecessarily. When airport capacity is overestimated, aircraft must hold in the air at the destination. Furthermore, in order to reduce wasted capacity in the presence of capacity forecast error, the ATCSCC plans for a small amount of airborne holding, called managed arrival reservoir, or MAR. Reductions in capacity forecast errors have the potential not only to reduce directly the amounts of airborne holding and wasted airport capacity but also to reduce the MAR required to achieve the same level of service.

Since departures have a large effect on airport arrival capacity in some configurations at some airports, it is important that they be accounted for in flow management decisions. Indeed, controllers at BOS rely on projected departure demand in making decisions about the feasibility of the projected arrivals. Unfortunately, this projected departure demand information is based largely on schedule information, since airlines typically do not update their schedules with delay and cancellation information adequately far in advance to be used in flow management decisions. Whereas schedule information reflects true demand accurately during regular operations, flow management decisions are most critical during irregular operations when capacity restrictions are necessary and flights are most likely to be delayed or canceled.

There is a further complication with the approach used today of adjusting arrival capacity estimates according to departure demand estimates. Not only does the schedule not represent the demand during irregular operations, but the departure demand is in fact determined by airlines in response to the arrival capacity allocated to them. If arrivals to an airport are delayed, departures will eventually be affected. An airline without excess inventory of aircraft and crew at an airport must wait for the arrivals before it can send out departures. Even if an airline is capable of releasing departures on schedule, it may be in the best interest of the airline to delay the
departing flights to allow passengers aboard the delayed arriving flights to connect. On the other hand, under some circumstances the airline may be inclined to operate the departing flights as close to schedule as possible to minimize the effect of delays on the rest of the day’s schedule. The factors that go into these decisions are far too complex for ATC to take into account, creating a fundamental restriction on the ability of ATC to account for departures in the arrival capacity forecast.

Finally, the departure demand information used by ATC in predicting arrival capacity consists of hourly counts of expected demand. This information does not distinguish between operations spaced evenly apart over the course of the hour and operations concentrated in one brief period during the hour, although the difference between the two extremes is significant. Consider an airport such as DFW, which has an arrival-departure capacity curve that is nearly rectangular (Figure 4-6). If the airport is operated at the extreme upper-right vertex on the arrival-departure capacity curve, it is capable of roughly twice the rate of operations that it could support if it exclusively served arrivals for one period (upper left-hand corner of the feasible region) and departures for the next (the lower right-hand corner).
4.2 Configuration Planning and Prediction of Arrival-Departure Capacity

Not only is a ground delay program adversely affected by poor demand information, but configuration planning may suffer as well. The configuration planning problem (CPP) is the problem of determining the sequence of configurations in which to operate an airport over the course of a day. The deterministic version of the CPP is to determine a configuration sequence causing minimal delay to flight operations, given the planned times and runway requirements of each arrival and departure, the weather conditions to occur over the course of the day, constraints on airport operations for environmental reasons such as noise impact minimization, and any other peculiarities involved at the airport. The configuration planning problem is tightly coupled with the flow management problem since the configuration planning problem determines capacity in order to suit demand and airport restrictions, while the flow management problem determines demand in response to capacity.

The configuration planning problem is complicated by many airport-specific and configuration-specific constraints. One such constraint is the amount of airport "down time" required to change configurations. For instance, converting operation of runway 22L at BOS from serving arrivals only to serving arrivals and departures requires little time, whereas changing from the configuration in which aircraft depart and arrive on runways 22L and 22R to a configuration using 4L and 4R (the same physical runways used in the opposite direction) requires considerable time and coordination between tower and TRACON controllers.

The configuration planning problem and flow management problems are solved today by controllers, with help at a few airports from computer tools such as PRAS, the Preferential Runway Assignment System. The most readily used source of information that controllers have about arrivals and departures over the course of the day comes from ETMS. The ETMS arrival information is based on radar position reports of aircraft en route. Other information about arrivals and departures in ETMS is based primarily on the schedule. This information reflects actual operations fairly
closely during good weather, but when capacity restrictions are present in the system, the information is often inaccurate, particularly for departures. Furthermore, the information is presented to controllers in the form of a count of the number of aircraft expected per hour into the future, a format which lacks detail that could be useful for configuration planning.

When capacity is sufficient to serve demand, the potential increase in capacity that could be achieved through improvements in configuration planning is of little importance. It is when capacity is insufficient to serve the demand that improvements in configuration planning are most important, for then such improvements may result in the ability to move more aircraft through the airport. Unfortunately, because the demand information is least reliable at such times, configuration planning is largely reactive. For instance, controllers recognize the need for more departure capacity by observing a queue of aircraft building on the taxiways of the airport rather than by knowing ahead of time when the aircraft will depart. In response to the observed queue, controllers in the tower at an airport such as BOS coordinate with TRACON controllers to reconfigure the airport for more departures. By the time controllers are able to effect the changes, the departure queue may have grown or shrunk considerably.

The degree to which flow management and configuration planning must be integrated depends on the amount of down time required to change configuration and on the differences in capacity of the available configurations. For instance, if the capacities of the available configurations are identical, then flow management can be dissociated completely from configuration planning. Similarly, if no down time is required to change configuration, then flow management can be performed based on a forecast of capacity equal in each period to the union of the period capacities under the different configurations. As down time between configurations increases, and as differences between the capacity of different configurations increases, so does the value to an integration of configuration planning and flow management.

Of course, no configuration planner could achieve much improvement without reliable advance information about demand. Sections 4.3 through 4.4 describe a mecha-
nism for allocating arrivals and departures during ground delay programs that would provide better information to a configuration planning function. In addition, the mechanism can provide information about configuration planning options available to the users, allowing them to collaborate with the FAA on configuration planning decisions.

4.3 Overview of the Arrival-Departure Capacity Allocation Method

We propose here the Arrival-Departure Capacity Allocation Method (ADCAM) as a means of mitigating the effects described in Sections 4.1 and 4.2. The differences between the current capacity allocation method and ADCAM are discussed below.

Figure 4-7 illustrates user demand response to flow management as implemented today. Demand is matched to airport arrival capacity by limiting arrivals to be less than the airport acceptance rate (AAR). Scheduled demand for arrivals and departures is indicated by point 1. The FAA requires the users to reduce arrivals to below the dashed line labeled "AAR" during the ground delay program.\(^1\) The users may, however, elect to depart a different number of aircraft, as indicated in the figure by points 3. If the users operate fewer departures than the number of departures at point 2, airport capacity is wasted. If the users operate more departures, a departure queue forms, causing departure delays and additional uncertainty in flight arrival times at other airports.

The proposed method, ADCAM, allows the FAA to specify capacity constraints on the arrival-departure plane rather than being limited to arrival constraints. Figure 4-8 illustrates the concept through a hypothetical scenario identical to that in Figure

\(^1\)For purposes of illustration, the AAR in the figure is set so that operations would be at airport capacity were the users to maintain the same number of departures in the period, at point 2. For short periods of capacity shortage for which the controllers have relatively reliable departure demand information, this strategy may be used in practice. For longer periods, the AAR is usually set to the number of arrivals feasible at the airport if operated with an equal number of departures, i.e., the point at which the 45° line passing through the origin of the figure crosses the capacity constraint.
4-7 except that the FAA has restricted the users to remain within the true airport capacity constraints rather than the AAR constraint. Thus, the users are unable to plan more departures than the system can handle; the details of how this can be achieved are developed in Section 4.4. Not only does ADCAM provide feasibility of user plans, the method gives the users the freedom to trade off arrival capacity and departure capacity according to the airport constraints in order to achieve their business objectives. The ramifications of ADCAM are discussed in Section 4.5.

4.4 Detailed Description of ADCAM

The Arrival-Departure Capacity Allocation Method is based on the same principles as today’s CDM Ground Delay Program, except that it takes into account departures and the tradeoff between airport arrival and departure capacity. These principles are reflected in the components of ADCAM:

1. FAA rations airport capacity among the airlines according to schedule;

2. Airlines assign their shares of available airport capacity to their flights; and
3. FAA runs "Compression" to recover unused capacity and provide an incentive to users to report planned operations accurately.

Each of these components under ADCAM is a generalization of the corresponding CDM GDP component. If one were to employ ADCAM with a model of airport capacity allowing unlimited departures, the CDM GDP would be the result.

In order to run the ground delay program, it is necessary to have determined an airport capacity forecast over the period of time in question. Since capacity is a function of the configuration and operating mode used at the airport, configuration planning is an integral part of the flow management decision process. Before ADCAM is started, airport capacity is determined according to the best estimate of the configuration sequence to be used. Improved demand information revealed during application of ADCAM may be used to improve the configuration plan, producing new estimates of capacity that can be used in the compression component of ADCAM.

The next sections describe the details of the three components of ADCAM.
4.4.1 Ration By Schedule

The first step that must be accomplished in a ground delay program is to ration constrained airport resources to the users. Chapter 3 discusses rationing resources according to the Groves mechanism to achieve optimality according to user objectives. Although that could be attempted in the arrival-departure case, the result would be fairly complicated and difficult to implement. Rather, ADCAM can ration airport resources according to schedule, a paradigm that has been well accepted by the industry and government under the CDM program.

Under CDM, arrival resources are currently rationed first-scheduled, first-served. The flights expected to arrive during the ground delay period are arranged in schedule order and assigned arrival slot times spaced apart according to the AAR. For instance, if the GDP is to begin at 2:00pm, and the AAR is 30 arrivals per hour, then the earliest scheduled arrival expected during that period is given the 2:00 arrival slot, the next earliest is given the 2:02 arrival slot, and so on until the end of the GDP time period is reached.

Under ADCAM, the same approach is used, except that both arrivals and departures are given slots according to scheduled time at the runway. For instance, if the airport is capable of handling 60 operations per hour (perfect tradeoff between arrivals and departures), then the earliest scheduled operation expected during the GDP time gets the 2:00 runway slot, the operation scheduled next is assigned the 2:01 runway slot, and so on.

In the general concept of rationing by schedule, whether under CDM or ADCAM, the order of operations is preserved as scheduled, but the timing of those operations is delayed as necessary to match the modeled airport capacity. One could calculate precisely how long the scheduled sequence of operations would take to be accomplished, taking into account the airport layout in great detail, the aircraft types, and so forth; however, planning in such detail is unlikely to prove worthwhile since the operations realized rarely match the plan to such precision. Instead, ADCAM, like CDM, relies on the average arrival-departure capacity model of the airport.
The average arrival-departure capacity model for an airport indicates the average additional time required to serve each arrival or departure given the operations mix, set of feasible configurations and characteristics of the configuration planning being performed at the airport, averaged over sequences preserving the operations mix. Models of this type could be created by averaging the amount of time required to operate a arrivals and d departure in a given configuration and weather condition and accounting for configuration planning either explicitly or implicitly as discussed in Section 4.2.

The ADCAM ration by schedule process can readily be understood graphically. First, time is divided into periods of, e.g., ten minutes in length. ADCAM then assigns the first n scheduled operations to the first period of the GDP, so that the n\textsuperscript{th} operation either exceeds the capacity curve for the airport or is scheduled to happen after the first period (see Figure 4-9). If configuration planning is performed explicitly in combination with ADCAM, then the configuration planner is involved in determining which configuration's capacity curve is used. The second period is assigned to operations n + 1 through the first operation that is scheduled after or doesn't fit in the second period, and so on.

Once operations have been assigned to periods, each operation is assigned a slot time in the period, in schedule order, spaced evenly throughout the period. This assignment of times to operations is a solution to the problem of determining when each operation is to take place, and will be referred to as the RBS Solution. Although the RBS solution is a feasible solution with respect to the airport capacity constraints, it may not be feasible with respect to other constraints known only to the airlines.

In effect, the RBS Solution rations the resources that are defined by the constraints that are binding in the RBS Solution. The binding resources are those resources that are depleted in the solution, that is, there is no slack in the binding resources. The slack resources, on the other hand, are not depleted in the solution. For example,

\footnote{It is not necessary to use discrete time periods. Section 7.4.5 discusses a continuous version of ADCAM that eliminates some undesirable characteristics of the discrete time period version. Because the mechanics of the method are essentially identical in both cases, this simpler discrete-period method is presented here.}
if the operating point determined by the RBS Solution for a given period were that marked by the asterisk in Figure 4-9, the binding resource would be defined by the constraint on which the asterisk lies, and consists of the ability to move in the direction perpendicular to that constraint in arrival-departure space. The other two constraints, represented in the figure by the adjacent line segments on the capacity curve, are not binding in this solution, so the resources associated with them are slack resources.

Although the slack resources are not binding in the schedule solution, they may well become binding once the users make substitutions in the second component of ADCAM. Therefore, the slack resources must be rationed as well. The combined result of all user resource constraints can be viewed in arrival-departure space just as overall airport capacity can. The set of resources allocated to a user will be called the user's arrival-departure subcapacity.

Graphically, the allocation may be seen on the arrival-departure capacity graph for the airport. Consider a situation in which the RBS Solution allocates $d_r$ departures and $a_r$ arrivals to user $r$ at an airport in some time period $t$. Furthermore let $d$ represent the total number of departures and $a$ the total number of arrivals at the airport in period $t$ according to the RBS Solution. Finally, let the capacity of the
airport be represented by the (in theory possibly infinite number of) breakpoints \( B = L \cup R \) where

\[
L = \{(a + \alpha^L_1, d - \delta^L_1), (a + \alpha^L_2, d - \delta^L_2), \ldots, (a + \alpha^L_{|L|}, d - \delta^L_{|L|})\}
\]

and

\[
R = \{(a - \alpha^R_1, d + \delta^R_1), (a - \alpha^R_2, d + \delta^R_2), \ldots, (a - \alpha^R_{|R|}, d + \delta^R_{|R|})\}
\]

such that the \( \alpha^R_k, \delta^R_k, \alpha^L_k \) and \( \delta^L_k \) are all positive and increasing in \( k \). This representation is shown in Figure 4-10. Note that the set \( L \) consists of the capacity curve to the left of the operating point, and the set \( R \) consists of the curve to the right of the operating point.\(^3\)

\(^3\)This capacity representation restricts us to concave capacity curves such as that drawn in Figure 4-10, resulting in convex feasible regions. Because the curves represent operations possible over a period of time, this convexity of the feasible region of airport capacity is appropriate. For example, if an airport can operate at point \( x_1 = (a_1, d_1) \) and at point \( x_2 = (a_2, d_2) \), then by operating at \( x_1 \) for \( \lambda \) of the time and at \( x_2 \) for \( 1 - \lambda \) of the time, the combined operations will be at \( (\lambda a_1 + (1 - \lambda)a_2, \lambda d_1 + (1 - \lambda)d_2) = \lambda x_1 + (1 - \lambda)x_2 \).
Using this representation, the capacity allocated to user $r$ in period $t$ is given by the breakpoints $B_r = L_r \cup R_r$, where:

\[
L_r = \{(a_r + \frac{d_r}{d} \alpha^L_1, d_r - \frac{d_r}{d} \delta^L_1), (a_r + \frac{d_r}{d} \alpha^L_2, d_r - \frac{d_r}{d} \delta^L_2), \ldots, (a_r + \frac{d_r}{d} \alpha^L_{|L|}, d_r - \frac{d_r}{d} \delta^L_{|L|})\}
\]

and

\[
R_r = \{(a_r - \frac{a_r}{a} \alpha^R_1, d_r + \frac{a_r}{a} \delta^R_1), (a_r - \frac{a_r}{a} \alpha^R_2, d_r + \frac{a_r}{a} \delta^R_2), \ldots, (a_r - \frac{a_r}{a} \alpha^R_{|R|}, d_r + \frac{a_r}{a} \delta^R_{|R|})\}
\]

Graphically, the user subcapacity curve to the left of each user's operating point is a copy of the overall airport capacity curve to the left of the overall operating point scaled to fit between the vertical (arrival) axis and the user's operating point (see Figure 4-11). Likewise, the user capacity curve to the right of the user's operating point is the overall curve scaled to fit between the departure axis and the user's operating point. Note that the scaling used preserves slope.
The determination of user subcapacities and their communication to the users are the products of the ration by schedule step.

4.4.2 Substitution and Cancellation

The substitution and cancellation component allows each user to operate any set of operations feasible to its subcapacity constraints. By design of the subcapacity constraints, the resulting overall operations will be feasible to the overall airport capacity constraints. Furthermore, any operating point feasible to the overall airport capacity can be achieved by a subcapacity-feasible solution, ignoring flight integrality, and of course there exists at least one integral subcapacity-feasible solution (the RBS solution) that uses maximal integral airport capacity. These properties of the subcapacities are proved in Appendix A.

Under CDM, an airline is allowed to rearrange the assignment of slots to its flights as long as the resulting assignment is valid. Validity of an assignment is determined by a set of rules designed to ensure that the airline assigns each slot to a flight that can arrive at the slot time. Since all slots are the same length, any valid assignment of flights to slots should consume the same amount of airport resources as any other and is thus feasible to the capacity constraint imposed by ATFM.

Under ADCAM, an airline is similarly allowed to rearrange the assignment of slots to flights in any way that results in a valid assignment. However, since arrival and departure slots use different amounts of airport resources, all assignments at the capacity limit do not have the same number of slots. When a user adjusts its departures, the subcapacity constraints automatically adjust the number of arrivals the user can operate, whereas, under CDM, adjustment of the arrival rate would take several hours if it were done at all.

The adjustment is not perfect, unfortunately. It is in general possible to reach a solution for which a subcapacity constraint is binding on every user's solution, yet the overall airport capacity constraints are not binding (see Figure 4-12). Furthermore, it is sometimes the case in practice that a user decides not to use all of its allocated subcapacity or is unable to use all of its allocated subcapacity. For instance, a me-
chanical problem may ground an aircraft, crew may be unavailable to fly it, there may be insufficient demand for the flight, or the desired ratio of arrivals to departures may intersect the subcapacity constraint at a non-integral value. For these reasons, the third component of ADCAM, compression, exists.

4.4.3 Compression

Compression under ADCAM is a generalization of CDM Compression. The objective of both forms of compression is to give airport resources that cannot be used by one user (the releasing user) to a user that can use the resources (the bridge user). In exchange, the releasing user receives airport resources that it can use later in the schedule.

Compression occurs on a regular cycle, e.g., every half hour during a ground delay program. Before compression is run, a notification message is sent to the users so that they may ensure their data is accurate. Of particular importance are the accuracies of the data fields specifying the user’s desired times of arrival and
departure: Earliest Runway Time of Arrival (ERTA) under CDM, Desired Runway Time of Arrival (DRTA) under ADCAM and Desired Runway Time of Departure (DRTD) under ADCAM. These times are used in compression to determine which flights are available to serve as bridge flights.

For instance, consider a scenario in which user A canceled its 9:00 arrival to XYZ airport during a ground delay program and that user A’s next arrival to XYZ was at 11:00. Assume that the 9:00 flight had a slot to arrive at 9:30, the 11:00 flight had a slot to arrive at 12:00, and both flights’ ERTA’s or DRTA’s were equal to their original scheduled times. Since the 11:00 flight could not arrive by 9:30 according to its ERTA or DRTA, user A could not use the 9:30 slot, and without compression the user would have no incentive to report the cancellation in time for another user to receive the slot.

Compression would give the 9:30 slot to a bridge flight with an ERTA, DRTA or DRTD at or before 9:30, and give the slot freed by the bridge flight back to user A. If user A could not use this slot either, the process would continue until either user A received a slot it could use or user A had a slot that no other user could use.

Under ADCAM, this process works for departure capacity as well as arrival capacity. Rather than account for individual slots as is done under CDM, ADCAM keeps track of the amount of time used by the slots. Consider a time period of length $t$ minutes for which $a$ arrivals and $d$ departures are planned. If the slope of the binding constraint is such that one additional arrival may be operated at the expense of $m$ (possibly fractional) departures, then the time required for one arrival may be defined as:

\[ t_a = \frac{t}{a + \frac{d}{m}} \]

and similarly,

\[ t_d = \frac{t}{ma + d} \]

Define the difference between the amount of time a user’s set of slots requires and
the amount of time the user's allocated slots require as the time released by the user in the period. Bridge flights capable of using the released time are allowed to do so, and the time freed by the bridge flights is returned to the releasing user.

It is because the time released by a user need not be an integral combination of the time required for arrivals or departures that times are accounted for rather than whole slots. A user whose plans do not require all of its allocated subcapacity is motivated to release the unneeded capacity early enough for another user to take advantage of it. Doing so allows the releasing user the benefit of the time saved by the system.

Many details of compression under both ADCAM and CDM have not been covered here. It has taken the CDM effort several years and person-decades of work to develop the details of compression under CDM. Implementation required addressing many issues involving historical data formats, the physical network, the timing and sources of database updates that occur routinely in the system, partial and complete system failures, and so on. Indeed, the group continues to discover reasons that the implementation of CDM must be modified from time to time. It is important that the objectives of compression described in this section be accomplished as CDM has shown possible. However, the specific details of compression under ADCAM and under CDM are unimportant to the issues investigated in this thesis.

4.5 Potential Benefits to Users of A-D Capacity Allocation Method

This subsection provides some intuition into the benefits users might receive through adoption of ADCAM. In addition to the benefits of increased system throughput and predictability that come through basing airline plans on more accurate models of airport capacity (demonstrated by simulation results in Chapter 6), ADCAM provides to users the benefit of additional flexibility in the times at which they operate their flights. At present, the FAA specifies the arrival times that may be used by an
airline. Under ADCAM, the FAA would specify constraints on arrival and departure operations according to the airport capacity constraints. Since according to airport capacity there exists a tradeoff between arrivals and departures, under ADCAM the users would be able to trade arrivals for departures and vice versa according to this tradeoff. The benefit of this flexibility is not apparent until one considers airline operations in some detail.

A large fraction of airline operations involves banks at a number of critical airports. A graphical representation of passenger connections through banks that is useful for airline tactical planning is the connection matrix. Figure 4-13 shows a part of the connection matrix produced from Northwest Airlines’ schedule for January 21, 1998 at Minneapolis-St. Paul (MSP). The connection matrix for a given period of time lists down the left side the flights arriving during that period, arranged by time of scheduled arrival. Across the top are listed the flights departing during the period, arranged by scheduled departure time. Each row of dots corresponds to the arrival listed at the left of the row, and each column of dots corresponds to the departure listed at the top of the column. At the intersection of the row of arrival A and the column of departure D a colored dot indicates the degree to which the connection from A to D is feasible. For example, if there is enough time for passengers to connect easily from arriving flight A to departing flight D, the dot corresponding to the connection is colored green. If the connection is marginal, the dot is yellow, and if the connection cannot reasonably be made, the dot is red. For purposes of this example, a connection is deemed feasible if the departure leaves the gate at least twenty minutes after the arrival arrives, marginally feasible if the departure leaves between ten and twenty minutes after the arrival, and infeasible otherwise. For instance, Figure 4-13 shows that connections from the 12:45Z arrival from La Crosse, WI (LSE) to the 13:20Z flight to Grand Rapids, MI (GRR) are schedule feasible. Indeed, 13:20 - 12:45 = 35

\[\text{In practice, the feasibility of connection times would be a function of the airport, airline, and specific gates involved. For scheduling purposes, twenty minutes is most likely far too short – it would be inadvisable to schedule connections so tightly. For attempting to achieve connections from a lightly-loaded aircraft that has been delayed on arrival at an airport with a small terminal, twenty minutes might be more reasonable.}\]
Figure 4-13: Connection Matrix Example

minutes, well over the 20 minute feasibility minimum. On the other hand, connecting from the 13:59Z arrival from Grand Forks, ND (GFK) to the 13:20Z flight to GRR is of course infeasible.

Notice how sharp the boundary between red and green is – there are relatively few yellow dots. Notice also how precise a stair-step pattern the boundary makes. This stair-step pattern is clearer in a connection matrix covering a larger period of time, such as that shown in Figure 4-14, generated from the same day in Minneapolis. The outlined square green area in the center is an example of a scheduled bank. The portions of the adjacent banks that fall onto the time period of this connection matrix are outlined as well. The author’s experience has shown that visual delineation of
many airlines’ banks is as clear as it is for Northwest’s banks at MSP. For instance, US Airways’ bank structure at PHL and PIT, TWA’s bank structure at STL, and American’s banks at DFW all have well defined stair-step patterns in the connection matrix diagram. However, the bank structure need not always be so clearly defined. For instance, United’s banks at O’Hare in Chicago (ORD) and Continental’s banks at Newark (EWR) have connection matrix diagrams with much smoother corners (Figure 4-15), and of course some airlines such as Southwest choose not to operate hubs at all.

Next, consider what happens when the FAA issues a GDP under the current system of delaying arrivals only. The hubbing airline is faced with the problem of deciding whether to delay its departing flights so that passengers from late arrivals can connect or to send the departures out as close to schedule as possible to minimize
Figure 4-15: Continental’s Connection Matrix for EWR
Figure 4-16: Preserving Connections under CDM (middle) and ADCAM (right).

the downstream consequences of delay. If the airline elects to hold its departures for late-arriving passengers, the delays its departures must absorb may be significant.

Under ADCAM, an airline would gain the flexibility to use the capacity freed by the late departures for its arrivals. By reallocating some of the early departures’ capacity to arrivals, the airline would be able to land its arrivals earlier, thus requiring the departures to absorb less delay than would be necessary today to achieve the same connection performance. See Figure 4-16, which shows a miniature bank with connections made infeasible by a ground delay program (left) and the delays the airline must incur to partially rectify the situation under CDM (middle) and ADCAM (right), assuming a one-to-one tradeoff between arrivals and departures. Under CDM, the arrival times are fixed, requiring the changed departure times shown. Under ADCAM, the same connection feasibility can be maintained with considerably less delay by trading the 1:07 departure slot for a 1:07 arrival slot (assuming the 1:10 from PIT could land at 1:07 and the 1:15 from DFW at 1:10).

Another benefit to the users of ADCAM would be the ability of each user to determine its own MAR in response to the uncertainty it sees in the weather forecasts. Some airlines have developed advanced weather prediction capabilities. If such an airline concluded that the FAA’s weather forecast were overly pessimistic, it would be able to allocate more of its capacity to arrivals than otherwise, allowing those flights to depart their origins for the affected airport earlier than would be optimal under the FAA’s forecast, with the expectation that improving conditions would relieve the
capacity constraints in time to allow arrivals to land earlier than anticipated by the FAA and departures to leave according to plan. Likewise, if the airline realized that the FAA’s forecast were overly optimistic, the airline could allocate some of its arrival capacity to departures before the arrivals departed their origin airports, effectively operating a self-imposed GDP, so that it would have capacity left for departures when the weather deteriorated.

Another way that an airline could use the flexibility to exchange arrival capacity for departure capacity might prove useful during extremely bad weather. Many airports are able to accommodate more than one departure for each arrival foregone in bad weather. Under ADCAM, an airline with equipment inventory at the airport would be able to cancel fewer operations by using more of its capacity for departures than for arrivals over the short term, letting arrivals catch up once the weather improved.
Chapter 5

ADCAM Analysis Tools

A principal goal of this thesis is to demonstrate the benefit of incorporating departure considerations into ground delay programs in the manner proposed in Chapter 4. The amount of effort necessary to effect such a change is considerable; not only would the FAA need to change operating procedures, the airlines would need to develop new methods of decision making to take advantage of the new flexibilities ADCAM would offer them. Such a change will not be considered unless it can be shown to be worth the effort.

To this end, the simulation and other analysis tools described in this chapter were developed. These tools allow one to run controlled experiments comparing operations under CDM, today's method of operating ground delay programs, to the ADCAM method described in Chapter 4, under a variety of different operating conditions, at different airports, and with different types of airlines. The great advantage of the simulation developed for this study is the ability it provides to operate the system under one set of conditions using one flow management strategy, record the results, and then operate the system under identical conditions using another flow management strategy. Were tests to be performed in the real world rather than in simulation, not only would the costs involved be significant, but comparing the two methods under identical conditions would be impossible.

Of course, simulation is not without its drawbacks. The primary challenge in simulation is ensuring that the simulation models are representative of the real-world
systems they emulate. A number of steps have been taken to achieve this goal. Whenever possible, real system data have been used. Where the quality of data available from the system is not adequate for direct use in the simulation, the simulation data have been generated from models of the processes that create the real world data, and the statistical properties of the simulated and real-world data compared. For data that are not available directly, such as airline cost and objective data, a range of reasonable values that include the values likely to exist in the real world is used in the simulation. Finally, it is important in a simulation to include enough detail of the real world to represent it faithfully, yet there are limits to the amount of detail that can be modeled accurately. Where detail has been omitted from the simulation, that fact is noted and the relevance of the most significant omitted details to the results is discussed.
5.1 System Simulation

Figure 5-1 shows the decisions involved in flow management that are accounted for in the analysis. In that figure, the shaded area contains the elements of the system that are simulated: the FAA’s flow management decisions, the airlines’ plans in response to the flow management decisions made by the FAA, and the performance of the airport as the airlines execute their plans. As in the real system, it is the FAA’s role to determine the airport capacity and to allocate that capacity among the airlines. The FAA does not dictate which flights should be canceled or how much delay specific flights must accept, it only allocates the forecast capacity among the airlines. It is the role of each airline to decide which flights to cancel, which to delay and how to spread the necessary delay among its flights.

Since the purpose of the simulation is to compare the effect of the two fundamentally different methods of allocating airport capacity among the users, ADCAM as proposed in Chapter 4 and the method used in present-day ground delay programs under CDM, the simulation model of the FAA performs the allocation both ways, as described in Sections 5.3.1 and 5.3.2. The resulting allocation is given to the simulated airlines, each of which is modeled by a mathematical decision model representing the decisions of real airlines reacting to ground delay programs as described in Section 5.4. The simulated airlines decide which flights to cancel and how to allocate delay among the remaining flights in order to minimize the delay of passengers and the disruption of the airline’s schedule.

Configuration planning and capacity forecasting are not simulated. Rather, the end result of configuration planning and capacity modeling, the capacity forecast, is an input to the simulation, as described in Section 5.2. The simulation measures, in effect, the open-loop performance of flow management decision-making, since in the simulation the FAA does not automatically re-plan dynamically in response to airline plans or observed simulated airport operations, nor do the airlines re-plan dynamically in response to simulated airport performance. As discussed in Chapter 2, these feedback loops are of limited effectiveness in practice, so the open-loop performance
of the system measured by the simulation is an important indicator of overall system performance. Indeed, in many instances, the real system does operate in an open-loop mode because the delays of information flow in the system preclude use of feedback.

5.2 Capacity Forecasts

In order to produce meaningful flow management results, a capacity forecast that is insufficient to accommodate the scheduled demand must be entered into the simulation. The simulation provides a graphical display of the scheduled demand that is useful in creating such capacity forecasts, and on which the capacity forecast is displayed as well, as shown in Figure 5-2.

The simulation's graphical display of demand and capacity consists of a sequence of subgraphs, each of which shows the demand and capacity within a period of fixed length. The time period shown by a subgraph begins at the time shown below the subgraph\(^1\), and ends immediately before the beginning of the following graph.

In each subgraph, the demand is represented by a solid, colored diamond connected to the origin of the subgraph by a line. The greatest demand shown in a subgraph is colored grey and represents the overall demand within the period. The other demand points shown are demand for specific airlines, colored according to the code in Table 5.2. The vertical axis of each subgraph represents arrivals and the horizontal axis represents departures. The scale on the arrival and departure axes is shown in the form of a white grid in the shaded region of the subgraph on which one grid spacing represents one operation per ten minute period of time (thus, if the subgraphs are of ten minutes length as shown here, one grid space equals one operation, whereas if the subgraphs are spaced an hour apart, one grid space equals six operations, and so forth). The demand is not cropped to fit within the subgraph. For instance, the first subgraph shown in Figure 5-2 shows overall demand extending off the subgraph into the subgraph for the next time period.

---

\(^1\)Subgraph times are in the format DD/HH:MM, where DD represents the day of the month, and HH:MM are the hours and minutes in Universal Coordinated Time (UTC)
The shaded region on each subgraph represents the feasible capacity in effect at the beginning of the subgraph's time period. If the capacity changes in the middle of a subgraph's time period, only that capacity in effect at the beginning of the time period is shown. Unlike the demand drawn on a subgraph, the shaded region representing feasible capacity is cropped if it extends beyond the subgraph; for instance, the capacity shown in the 14:00, 14:10, and 16:10 subgraphs has been cropped because the scale of each subgraph does not allow higher capacity to be drawn without extending beyond the subgraph.

As mentioned above, the demand and capacity display can be viewed on different time scales, e.g., ten, fifteen, thirty, sixty, $60 \times 2^i$, $i \in \{1, 2, 3, \ldots\}$ minutes per period. It can also be scaled to show any number of arrivals and departures on each subgraph. The software can display actual operations, planned operations, or simulated operations in addition to scheduled operations, and it can display the capacity constraints imposed by the FAA on individual airlines.

The capacity forecasts themselves are entered into the simulation as a series of times and coefficients. The times entered are the times at which the capacity of the airport changes. Associated with each time is a series of coefficients, each pair of which represents one constraint in arrival-departure space.

For instance, the series of times and coefficients shown in Table 5.1 define the capacity forecast of Figure 5-2. This sample capacity forecast has unrestricted capacity before 14:20 UTC on the second day of the month. In each ten-minute time period from 14:20 UTC until 16:00 UTC, the following constraints are in effect:

\[
.1a_t \leq 1
\]
\[
.1d_t \leq 1
\]
\[
.07a_t + .07d_t \leq 1
\]

A bug in the drawing software causes a capacity to be drawn one period longer than it is in effect. In Figure 5-2, period 16:00 should show the same capacity as period 16:10. This error does not affect any of the results in any way; it is merely a drawing error that was not deemed important enough to fix, since it shows up only in Figure 5-2.
Figure 5-2: Capacity Forecast and Scheduled Demand (two periods skipped)
Effective time: 02 1420
Coefficients: .1 0
Coefficients: 0 .1
Coefficients: .07 .07
Coefficients: 0 0

Effective time: 2 1600
Coefficients: .05 0
Coefficients: 0 .05
Coefficients: 0 0

Effective time: 99 0
Coefficients: 0 0

Table 5.1: Capacity Constraint Input Sequence

where $a_t$ is the number of arrivals in time period $t$ and $d_t$ is the number of departures in time period $t$. Thus, in any ten-minute time period between 14:20 UTC and 16:00 UTC according to this forecast, no more than ten arrivals or ten departures may operate, nor may more than 14 operations total occur in the period.

At 16:00 UTC, according to the forecast, the following constraints go into effect:

$$0.05a_t \leq 1$$

$$0.05d_t \leq 1$$

Thus no more than twenty arrivals or twenty departures may operate in any given ten-minute time period from 16:00 UTC until the end of the day. The final set of numbers signals the end of the capacity forecast.

5.3 FAA Flow Management Models

Given the capacity forecast at an airport, it is the role of the FAA to restrict traffic at the airport to prevent excessive airborne holding as described in Chapter 1. The simulation models in one of two ways the decisions made by the FAA and the re-
Figure 5-3: Demand-Capacity Window Pop-Up Menu

Restrictions imposed by the FAA on the airlines: according to the procedures developed under the CDM program and in use today, or according to the ADCAM procedures proposed in Chapter 4. The type of ground delay program is selected from a pop-up menu in a demand-capacity simulation window (Figure 5-3).

Under both models, flow management constraints are communicated to each user in the form of constraints on the operations the user can perform in each ten-minute period of time, in a format identical to that used to represent airport capacity. Under the CDM model, the coefficients on departures are zero, meaning user departures are not controlled by flow management. Under ADCAM, both the arrival and departure coefficients may be nonzero, allowing the subcapacity constraints described in Section 4.4.1 to be represented.

5.3.1 Model of CDM GDP

Under the model of CDM ground delay programs, the Airport Acceptance Rate (AAR) is set to the number of arrivals that can be operated while operating the same number of departures. Graphically, this value can be seen as the value on the arrival axis corresponding to the point at which a line passing through the origin at a 45° angle to the departure axis crosses the capacity curve. For example, the capacity of Figure 5-4 would result in an AAR of four arrivals per period since operation of four arrivals and departures per period is feasible, whereas operation of five arrivals and departures is not.

The model then assigns each arrival to a time period according to schedule priority.
Figure 5-4: Setting the AAR.

That is, if the AAR allows \( n \) aircraft to land in a given period, then the \( n \) flights with the earliest scheduled times of arrival within or before that time period that have not received slots in earlier time periods are allocated slots in the period. If fewer than \( n \) such flights exist, then they all are allocated slots within the time period. In effect, this method slows the rate of aircraft arriving at the airport so that it is feasible to the AAR, while maintaining schedule order. This method of assigning arrival slots to arriving flights according to schedule priority is known as "Groover Jack" in the air transportation community; it is widely accepted as fair and is the heart of flow management today.

The constraints imposed on the airlines under CDM consist of arrival slots. In the simulation, the arrival slots in each period are represented as a constraint on the number of arrivals that the airline may operate in the period, shown as horizontal bars color-coded by airline in Figure 5-5. Notice in that figure that four arrivals are allowed in each period, but the airlines to which those four arrivals are allocated change from period to period according to the schedule. Of course, if the capacity forecast changed from period to period, the AAR could change and result in a different total number of arrivals allowed in each period.
5.3.2 Model of Arrival-Departure GDP

Under ADCAM, the AAR is not set explicitly. Rather, the forecast arrival-departure capacity is used to represent the airport capacity and is allocated to the airlines according to schedule as described in Section 4.4.1: rather than allocating arrivals to time periods according to schedule priority, operations are allocated to time periods according to schedule priority, and the resources that are slack in this schedule solution are allocated as well, in the form of airline subcapacity constraints. Figure 5-6 shows the end result of this process.

In Figure 5-6, the subcapacity constraints are color-coded by airline. Notice that the shape and size of the subcapacity allocated to each airline changes from period
to period according to schedule, but that in any period the vector sum of appropriate
points from each airline's subcapacity is equal to a vector to any point within the
overall airport feasible region.

A few details of the simulation implementation of ADCAM ration by schedule
bear mention. When it happens that the overall scheduled demand does not lie
exactly on the boundary of the feasible region, the simulation finds the (often non-
integral) point on the boundary with the same arrival/departure mix, and treats
that point as the overall scheduled operations point. Since this point usually has
more operations than the number of operations scheduled, and each airline's share of
the overall arrival/departure capacity is scaled by its share of the overall operations,
the total capacity allocated to the airlines is generally less than the forecast airport
capacity by some small amount (always less than one aircraft).

An appropriate way to treat this issue in practice would be to recognize that the
sequence of operations planned for the period would be completed in slightly less time
than one standard period and to adjust the time at which the following period began
accordingly. Because the simulation does not do this, it results in operations that do
not completely use the forecast airport capacity and hence slightly underestimates
the advantages of ADCAM. The amount by which this is the case depends on the
capacity forecasts; for example, a capacity constraint such as:

\[ \frac{51}{100} a_t \leq 1 \]

would exhibit this problem quite badly, since for integral \( a_t \) the constraint is equivalent
to \( a_t \leq 1 \). Thus, the scheduled number of arrivals in the period would be 1, which
should require only 51% of the time in the period to achieve. To reduce the magnitude
of this problem in simulation, our attention will be restricted to capacity forecasts
that are close to the convex hull of the integral feasible region of airport capacity.
5.4 Airline Decision Models

The capacity forecast provides one principal part of the simulated flow management scenario. The other principal part of the scenario is the demand for those resources, determined by the airlines as they strive to achieve their business objectives. The airlines base their decisions on information such as the schedule, the aircraft being used to fly the schedule, the passengers booked on each flight, and the connections those passengers are attempting to make.

The airline industry is highly competitive. The airlines run their capital-intensive businesses on very small margins, competing fiercely for passengers to fill their schedules. For instance, the most financially successful airlines are usually about two months from bankruptcy\(^3\) [4]. For this reason, it is vital to an airline’s survival that it make good decisions about flight cancellations, flight delays, and operating strategy particularly during periods of reduced airport capacity. Of course, these decisions are also very important to the passengers whose schedules and business interests depend on the airline’s performance.

The costs of operating an aircraft depend only to a small degree on the number of passengers aboard. An aircraft that is not full weighs less and uses slightly less fuel than a full aircraft of the same type. The same number of crew are required to fly the aircraft and the same number of ground crew are required. The airline might save a very small amount of money on meals and wear on the aircraft [6].

Therefore, the profit an airline makes depends primarily on how full its flights are, how much it can charge for each seat, and how many flights it can operate. Most of the large airlines in operation today operate hub-and-spoke schedules which allow them to fill more seats on many of their aircraft than the older system of flying direct routes allowed. Rather than having a separate flight for each airport pair, the carrier provides flights between many of the airports and a specific hub airport. Passengers can take a flight to the hub airport, change aircraft, and depart for their destinations

\(^3\)Meaning that expenses and obligations would cause the airline to go bankrupt if no revenue were generated for two months.
<table>
<thead>
<tr>
<th>Code</th>
<th>Airline</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>American Airlines</td>
<td>red</td>
</tr>
<tr>
<td>AS</td>
<td>Alaska Airlines</td>
<td>black</td>
</tr>
<tr>
<td>CO</td>
<td>Continental Airlines</td>
<td>green</td>
</tr>
<tr>
<td>DL</td>
<td>Delta Airlines</td>
<td>dark blue</td>
</tr>
<tr>
<td>HP</td>
<td>America West Airlines</td>
<td>grey</td>
</tr>
<tr>
<td>NW</td>
<td>NorthWest Airlines</td>
<td>bright blue</td>
</tr>
<tr>
<td>TW</td>
<td>Trans-World Airlines</td>
<td>purple</td>
</tr>
<tr>
<td>UA</td>
<td>United Airlines</td>
<td>yellow</td>
</tr>
<tr>
<td>US</td>
<td>US Airways</td>
<td>tan</td>
</tr>
<tr>
<td>WN</td>
<td>Southwest Airlines</td>
<td>magenta</td>
</tr>
</tbody>
</table>

Table 5.2: Airline Color Codes

on the new aircraft. One advantage to the airlines of this approach is that small numbers of passengers departing a city to various other cities can be combined into a single flight from that city to the airline’s hub; similarly, small numbers of passengers from various cities to a given city can be combined into a single flight from the hub to that city. As a result, the airlines are able to provide more complete schedules and more frequent service while increasing the load factors on their aircraft.

The prevalence of hub and spoke systems can be seen in Figure 5-7. In that figure, a colored line is drawn between each pair of cities between which a flight was scheduled to fly on January 21, 1998 by one of the ten largest airlines. Table 5.2 shows the assignment of colors to airlines used throughout this thesis. It is clear from the picture that each airline operates many flights in and out of several specific cities. Figure 5-8 shows the flights of American Airlines, clearly showing their principal hubs at DFW, ORD and MIA. These figures are screen dumps from the simulation’s graphical interface.

Another aspect of the airline schedules may be seen by looking more closely at the times of the scheduled flights. Figure 5-9 is a plot of scheduled east-west position of each flight arriving or departing ORD on the horizontal axis versus time on the vertical axis. Lines sloping up as they move toward the center of the figure represent arrivals, whereas lines that slope up away from the center of the figure are departures. Notice
Figure 5-7: Route Structure of Major Airlines
that American and United flights are scheduled to arrive and depart in interleaving banks. The development of interleaved banks allows the two carriers to coexist at the same airport — if the banks were simultaneous, unacceptable delays would result for both carriers.

Notice also that the banks are oriented to the east or west; a bank arriving from the east is followed shortly by a bank departing to the west, and vice-versa. In this manner, the airline reduces the size of the banks — and the resources required by the banks — by almost half without affecting the connections that the majority of passengers want to make.

The schedule information used in the simulation is real schedule information derived from the Airline Service Quality Performance (ASQP) database\(^4\), so it captures those aspects of the real system exhibited in the schedules described above. Another set of data upon which airlines base decisions in response to flow management on a given day is the set of their passengers' itineraries for that day.

Since there is no accessible source for historical passenger itinerary information, these data are generated probabilistically by the simulation based on the schedule. The simulation generates passenger itineraries similarly to the way they are formed in reality. The assumption is made that the schedule has roughly the right capacity for the passenger demand. Each flight is loaded with passengers according to a normal distribution centered at 75% of the number of seats available for the aircraft\(^5\) (roughly the industry-average load factor during months of high travel demand). The standard deviation of the normal distribution was chosen to be 25% of the number of seats of the aircraft, and the distribution is truncated so that flights do not receive negative passengers or more passengers than seats. This choice of distribution results in aircraft that are completely full roughly 16% of the time and less than half full.

\(^4\)ASQP data are compiled by the FAA. The database contains scheduled and actual times of arrival and departure for the jet flights of the top ten domestic air carriers.

\(^5\)The ASQP data also contain aircraft tail numbers for those flights that were not canceled. A separate database is used to look up the type of aircraft and its seating capacity based on the tail number. For those flights that were canceled, the type of aircraft is chosen at random from the types of aircraft used by flights serving the same origin-destination pair such that the probability that the flight is served by an aircraft of type \(X\) is proportional to the number of type \(X\) aircraft known to serve the same route.
East-West String Line Diagram for Chicago O'Hare
2 March 1998

Figure 5-9: ORD String Line Diagram
Figure 5-10: Filled Aircraft Seats

roughly 16% of the time (completely full and half full are each one standard deviation away from the mean of the normal distribution), values which seem reasonable based on the author’s personal experience.

Figure 5-10 shows a hypothetical end result of the passenger load generation on a small schedule. In the figure, arriving aircraft are represented by the left column of boxes, departing aircraft by the boxes on the right. The flights are placed along the time axis which runs down the left side of the figure according to scheduled gate arrival or departure time. When two flights are scheduled to operate at the same time, they are drawn next to each other. The length of the box represents the number of seats in the aircraft serving the flight, and the shaded cyan region represents the
number of those seats that are filled.

Once the number of filled seats has been determined, passenger connections are generated through the hubs in a two-stage process. In the first stage, a “target” for the number of passengers connecting from and to each flight at the hub is generated according to a binomial distribution: each passenger is a “target” to make a connection with probability $pcf$, where $pcf$ is determined based on recent connection statistics for the hub in question. At the end of the first stage, each arriving flight and each departing flight contain a number of connection targets. Figure 5-11 shows the generated connection targets, represented by a hatched box on each flight.

In the second stage of the connection generation process, a passenger arriving at
Figure 5-12: Arriving Connection Target Chosen
the hub is chosen at random from those passengers that are connection "targets", and a departure connection is chosen for that passenger. The red oval in Figure 5-12 represents this chosen arriving passenger. The departure connection for the passenger is chosen as follows: any flight that is earlier than the arrival plus min_schedule_connect minutes is ignored. The departures after that time are assigned pseudo times evenly spaced through the rest of the day. In the example, those departures shown according to schedule times in Figure 5-12 are shown according to pseudo times in Figure 5-13.

A departure flight is chosen by pseudo-departure time according to a truncated gamma distribution with shape parameter 1.2 and mean 105 minutes (see Figure 5-14). This distribution is used to concentrate most of the connections to flights
Figure 5-14: Probability Density Function Used to Choose Departure Connection within the scheduled bank while allowing a few connections to flights outside the bank. In the case illustrated, only one bank of flights exists between the arrival and the end of the day, so the tail of the gamma distribution is not used. Figure 5-15 illustrates the truncated gamma distribution used to pick a candidate departure for the connection from among the departures assigned pseudo times in Figure 5-13. A specific departure is chosen for the connection with probability proportional to the area under the curve as shown in Figure 5-15. Over longer schedules, less of the tail of the gamma distribution would have been truncated and would generate connections outside of the departure bank.

If a direct flight between the resulting origin-destination pair would be less than 80% of the distance traveling through the hub, (including the case in which the seat selected according to this gamma distribution is on an aircraft destined for the passenger’s origin airport), another destination is chosen according to the same procedure. This prevents connections that result in more than 25% extra distance being scheduled. Although such connections are made in practice, they are less frequent than
Figure 5-15: Gamma Distribution Used to Choose Departure Connection
more directly routed connections; the extra distance flown adds enough to the time and cost of the flight that passengers are likely to choose to fly on a carrier with a hub better situated for their travel needs.

As the final step in the selection of a connection, the connection generator checks whether a connection can be made to an earlier flight for the same destination, and if so the earlier flight is chosen. Once the earliest feasible departure is chosen, a later arrival from the same origin is chosen if one exists and has space. This is meant to mimic the behavior of passengers and the scheduling system which try to minimize connection time. Once these checks have been completed, the passenger’s connection is added (see Figure 5-16), and the process continues until no connection targets remain or none of those remaining can be connected.

The result of this passenger generation on the same connection diagram of Figure 4-14 is shown in Figure 5-17. The lightest shade dot represents a single passenger making the connection, the next lightest shade represents two passengers making the connection, the middle shade represents four passengers, the next darker shade represents eight passengers, and finally the darkest shade represents sixteen or more passengers.

In Section 5.7.1, the passenger data thus generated are compared to sample data obtained from a major U.S. airline.

In the simulation, these generated passenger data are used together with the parameters specific to each airline shown in Table 5.3 to form the objective function of the airline. Parameter \( paxval \) is interpreted as the value to the airline of delivering the passenger to his destination on time (which implies that all of his scheduled connections are made), relative to not delivering the passenger at all. Parameter \( paxdlycost \) is the cost to the airline of delaying one passenger by one minute. Thus, it is preferable to cancel a passenger’s flight up front than to delay the passenger by more than \( paxval/paxdlycost \) minutes. In the simulations run in Chapter 6, a value of 90 minutes is used for \( paxval/paxdlycost \).

Passengers can experience delay in one of two ways: either the aircraft delivering the passenger is delayed, or the passenger misses his connection and must wait for the
Figure 5-16: Passenger Connection Made
Figure 5-17: Passenger Connections Generated for NorthWest
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>paxval</td>
<td>The value of fulfilling one passenger's itinerary.</td>
</tr>
<tr>
<td>paxdlycost</td>
<td>Cost of passenger minutes of delay.</td>
</tr>
<tr>
<td>airframefact</td>
<td>Cost of airframe tardiness relative to passenger tardiness.</td>
</tr>
</tbody>
</table>

Table 5.3: Airline-Specific Cost Parameters

next one. The parameter airframefact trades off these two types of delay indirectly. Parameter airframefact is the cost of delaying an aircraft by one minute relative to the cost of delaying the passengers on-board the aircraft one minute. When an aircraft is delayed, not only are those passengers on-board delayed, but the passengers that will be served by the aircraft on its next flight are usually delayed and the passengers to be served by the flight crew of the delayed flight may be delayed as well. A missed connection, on the other hand, delays only those passengers trying to make the connection. A study conducted by Oak Ridge National Laboratories and American Airlines [5] suggests that reasonable values for airframefact fall well within the range, 1-20, tested in Chapter 6.

The model of airline behavior is an optimization model. The objectives of this optimization model are based on the schedule, passenger connection data, and airframe delay costs. The constraints of the optimization model come both from the constraints inherent in the schedule and from the flow management constraints imposed on the airline by the FAA.

5.4.1 Mathematical Model of Airline Decision Problem

The mathematical model of the airline decision problem accounts directly for the decisions at one airport at a time. Network effects, such as the delays caused later in the day throughout the network by an early departure delay of a key flight, are not captured directly. These effects are accounted for only indirectly through the delay costs of the departing flight, in a manner suggested by the results of the analysis in [5], which defines multipliers expressing the relationship between the time of day of a given delay and the additional network delays caused by that delay later on. It is in this spirit that the parameter airframefact listed in Table 5.3 is used.
Table 5.7 shows the integer programming formulation of the airline model, and Tables 5.4, 5.5, and 5.6 describe the indices, variables, and constants used in the formulation. The optimization model relies on a discretization of time into sequential periods of fixed length; for the analyses performed in Chapter 6, the periods are of ten minutes length. Discretization of time into periods allows the flow management constraints placed on the airline to be represented in Equation 5.7 in the form:

\[ c_t^a a_t + c_t^d d_t \leq 1 \]

where \( a_t \) and \( d_t \) are the numbers of arrivals and departures in period \( t \), respectively, \( c_t^a \) is a constant multiplier on the arrivals in time period \( t \) and \( c_t^d \) a multiplier on the departures in period \( t \). For instance, if \( c_3^a = \frac{1}{7} \) and \( c_3^d = 0 \), the airline is restricted to seven arrivals in the third time period. If both \( c_t^a \) and \( c_t^d \) are set to nonzero values, the result is a linear constraint in arrival-departure space. There may be up to eight constraints in this form in any given time period in the software implementation of ADCAM (although increasing this maximum number would be trivial if required). User subcapacity constraints such as those shown in Figure 5-6 are represented through constraints 5.7.

The constraints other than constraint 5.7 in the mathematical model ensure that the variables used in the objective are representative of the physical quantities they model. Constraints 5.2 and 5.3 ensure that each flight is either canceled or assigned to exactly one time period. Constraint 5.4 forces \( z_{ik} \) to be 0 if there is inadequate time between the arrival of flight \( i \) and the departure of flight \( k \) for passengers to connect. Equations 5.5 and 5.6 force \( z_{ik} \) to be 0 if either arrival \( i \) or departure \( k \) is canceled.

The objective (Equation 5.1) of the mathematical model is the sum of three parts, each of which is summed over all arrivals, departures, and time periods:

1. The first sum represents the value of having the arriving flights arrive at the times specified by the \( x_{it} \); \( x_{it} \) is 1 for the time period \( t \) in which arriving flight \( i \) arrives, 0 elsewhere. Each arriving flight that is not canceled contributes \( a_{it} \)
to the airline’s objectives if it arrives at time $t$.

2. The second sum represents the value of having the departing flights depart at the times specified by the $y_{kt}$; $y_{kt}$ is 1 for the time period $t$ in which departing flight $k$ departs, 0 elsewhere. Each departing flight that is not canceled contributes $d_{kt}$ to the airline’s objectives if it arrives at time $t$.

3. The third sum represents the value of achieving the passenger connections specified by the $z_{ik}$; $z_{ik}$ is 1 if passengers have time to connect from arrival $i$ to departure $k$ and 0 otherwise. Hence a connection that can be made between arrival $i$ and departure $k$ contributes $v_{ik}$ to the airline’s objectives.

This objective is extremely flexible; the cost of delay as a function of time may be represented by any cost function that is constant over ten-minute intervals, and the values of individual connections may be set in any way. For purposes of the experiments run in Chapter 6, the values $a_{it}$, $d_{kt}$, and $v_{ik}$ in the objective are determined from the passenger itinerary details and the cost parameters of Table 5.3. The value of operating a flight on time is approximated as the value of serving all passengers aboard that flight plus the value of operating the airframe on time. Thus, the component of value attributed to an on-time arrival, i.e., the $a_{it}$ for $t$ corresponding to flight $i$’s scheduled arrival time, is set to $(airframefact + 1) \times paxvalue \times (terminating\,\, passengers)$. The value for a flight arriving one period in advance of its scheduled time is the same as the value for arriving on schedule. Arrivals are not allowed more than one period ahead of schedule to reflect the fact that early arrivals are difficult or impossible for the airline to achieve. Values of $a_{it}$ for time periods later than the scheduled time period are decreased by $pax\text{dlycost} \times (terminating\,\, passengers) \times (1 + airframefact) \times (minutes\,\, per\,\, time\,\, period)$.

It is assumed that flights do not depart ahead of schedule, thus the variables $y_{kt}$ for $t$ earlier than scheduled departure time are undefined. The value $d_{kt}$ for departure in the scheduled time period is set equal to $(total\,\, passengers\,\, scheduled\,\, aboard) \times paxvalue \times (airframefact + 1)$, and in subsequent time periods $d_{kt}$ is decreased by $(total\,\, passengers\,\, scheduled\,\, aboard) \times (minutes\,\, per\,\, time\,\, period) \times pax\text{dlycost} \times (1+$
<table>
<thead>
<tr>
<th>Index</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t \in T$</td>
<td>time period</td>
</tr>
<tr>
<td>$i \in A$</td>
<td>arriving flights at airport</td>
</tr>
<tr>
<td>$k \in D$</td>
<td>flights departing from airport</td>
</tr>
<tr>
<td>$n \in {1, 2, \ldots, N(t)}$</td>
<td>constraints representing capacity in a time period</td>
</tr>
</tbody>
</table>

Table 5.4: Airline Model Indices

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_{it} \in {0, 1}$</td>
<td>1 if flight $i$ arrives in time period $t$</td>
</tr>
<tr>
<td>$y_{kt} \in {0, 1}$</td>
<td>1 if flight $k$ departs in time period $t$</td>
</tr>
<tr>
<td>$z_{ik} \in {0, 1}$</td>
<td>1 if passengers can connect from arrival $i$ to departure $k$</td>
</tr>
<tr>
<td>$p_i$, $q_k \in {0, 1}$</td>
<td>1 if flight $i$ ($k$) canceled</td>
</tr>
</tbody>
</table>

Table 5.5: Airline Model Decision Variables

Finally, the value of allowing one passenger to make his connection is equal to the cost of the additional time the passenger would have to wait for the next departure to the same destination if he were to miss the connection. Accordingly, each of the $v_{ik}$ is set to $(\text{passengers connecting from } i \text{ to } k) \times (\text{scheduled departure time of next flight} - \text{scheduled departure time of } k) \times \text{paxdlycost}$.

<table>
<thead>
<tr>
<th>Constant</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l$</td>
<td>time periods required for passengers to make a connection</td>
</tr>
<tr>
<td>$a_{it}$</td>
<td>value of flight $i$ arriving in period $t$</td>
</tr>
<tr>
<td>$d_{kt}$</td>
<td>value of flight $k$ departing in period $t$</td>
</tr>
<tr>
<td>$v_{ik}$</td>
<td>value of connecting arrival $i$ to departure $k$</td>
</tr>
<tr>
<td>$c_{t,n}^d$, $c_{t,n}^a$</td>
<td>Capacity constraint constants</td>
</tr>
<tr>
<td>$N(t)$</td>
<td>number of capacity constraints in time period $t$</td>
</tr>
</tbody>
</table>

Table 5.6: Airline Model Constants
maximize \[ \sum_{i \in A} \sum_{t \in T} a_{it} x_{it} + \sum_{k \in D} \sum_{t \in T} d_{kt} y_{kt} + \sum_{i \in A} \sum_{k \in D} v_{ik} z_{ik} \] (5.1)
subject to \[ \sum_{t \in T} x_{it} + p_i = 1, \ \forall i \in A \] (5.2)
\[ \sum_{t \in T} y_{kt} + q_k = 1, \ \forall k \in D \] (5.3)
\[ y_{k,t+1} - \sum_{j=1}^{t} x_{ij} + z_{ik} \leq 1, \ \forall i \in A, \ t \in T, \ k \in D \] (5.4)
\[ z_{ik} \leq 1 - q_k \ \forall i \] (5.5)
\[ z_{ik} \leq 1 - p_i \ \forall k \] (5.6)
\[ \sum_{i \in A} c_{i,n}^a x_{it} + \sum_{k \in D} c_{t,n}^d y_{kt} \leq 1, \ \forall t \in T, \ \forall n \in \{1, \ldots, N(t)\} \] (5.7)
\[ x_{it}, y_{kt}, z_{ik}, p_i, q_k \in \{0, 1\} \] (5.8)

Table 5.7: Airline Model

5.4.2 Mathematical Model Solution Technique

The simulation enumerates those objectives and constraints in the formulation of Table 5.7 necessary to find the optimal solution\(^6\), and solves them with a call to the CPLEX mixed integer programming libraries [21], version 6.0.1. The formulation is highly fractional for some problem instances, particularly those for stringent capacity constraints indicative of a tradeoff between arrival and departure capacities. For reasonable problem sizes, the fractional nature can prevent CPLEX’s stock branch-and-bound algorithm from reaching an integral solution within a reasonable amount of time on a 300 MHz Pentium II with 64MB of memory and 128 MB of swap space running the Linux operating system. For instance, the problem defined using the schedule of United Airlines at Chicago O’Hare International Airport from 1400Z to 2000Z on July 10, 1997 runs until memory is exhausted. The problem involves 168 arrivals, 180 departures and 36 time periods (six hours times six periods per hour),

\(^6\)Complete enumeration is not necessary; for zero-valued \(v_{ik}\), there is no need to define a corresponding \(z_{ik}\). Similarly, for \(a_{it} \leq -\sum_{k} v_{ik}\), the cost of a late arrival cannot be recaptured by the value of passenger connections, thus there is no need for a corresponding \(x_{it}\). Similar reasoning reduces the problem size significantly.
resulting in 6119 variables, 27441 constraints and 269643 nonzero matrix entries.\textsuperscript{7}

Of course, it is not the size of the constraint matrix that causes solution of the problem to take so long; some mathematical programming problems much larger than this one are solved routinely in very little time. The difficulty is that the solutions to the LP relaxation of the user model are often highly fractional. Fractionality is induced by the capacity constraints 5.7. Instances of the problem in which airport capacity is severely reduced suffer the worst fractionality; when airport capacity is higher than required by the demand, constraints 5.7 are not binding and the problem is solved rapidly.

This characteristic motivates a heuristic solution procedure. Rather than solve the problem in one step, it is solved in two stages. In the first stage, integrality of the variables involved in equation 5.7, $x_{it}$ and $y_{kt}$, is relaxed. The resulting solution has integral cancellation variables $p_i$ and $q_k$. In the second stage, those flights which were canceled in the first-stage solution are constrained to be canceled. CPLEX correctly identifies as superfluous the variables and constraints associated with these flights and removes them from the problem. The solution to the second stage is achieved rapidly since the constraints 5.7 are not significantly below the level required to serve the flights that were not canceled in the first stage.

The physical interpretation of the partially integral solution resulting from the first stage is that the times at which flights arrive and depart have not been decided, but that the cancellations required have been. This physical interpretation is reflective of the decision process employed at certain major U.S. airlines, at which one manager decides how many flights must be canceled, and his subordinates decide how best to operate given those cancellation decisions. Of course, the set of cancellations optimal to this partially integral problem is not in general optimal to the fully integral problem.

\textsuperscript{7}The numbers of variables and constraints are reduced to these numbers by the facts that not every flight should be able to land in every time period (limiting the $x_{it}$, $y_{kt}$ and associated constraints) and the matrix of passenger connections between arrivals and departures is sparse (limiting the $z_{ik}$ and associated constraints).
This two-stage procedure produces integral solutions in reasonable time\(^8\). Furthermore, since the optimal solution found in the first phase is less constrained than the fully integral problem, its objective value provides an upper bound to the true optimal value that may be used to compute an upper bound on the distance from optimality of the fully integral solution found in phase two. Often when solving problems with arrival-only constraints, the result of phase two is an optimal solution to the fully integral problem. It is while solving problems with arrival-departure constraints that a gap most frequently exists.

### 5.5 Airport Model

The airport model measures the physical airborne and departure queues that would form and the periods of capacity underutilization that would occur in the absence of FAA and user feedback. The airport is modeled as a single-queue, single-server queueing system whose capacity has been reached in any period once any of the arrival-departure constraints for the period has been violated.

Simulated airport capacity follows the prediction of the capacity forecast exactly. Thus, not only does the airport operate deterministically, but there is no additional detail in the model not present in the forecast. For these reasons, the simulation’s results are interpreted as a measure of the ability of the system under each allocation method to comply with a forecast.

Under reasonable assumptions about the quality of the capacity forecasts, it follows that the better the system can comply with a deterministic forecast, the better the system should operate in the presence of uncertainty and model errors. Since control during a ground delay program is exercised before the flights depart their origins, control must generally be effected several hours in advance of the time of the capacity shortage (more or less time depending on the en-route travel time of the

---

\(^8\)Some hard problems may take an hour or more to reach optimality. However, the first few solutions are usually reached by CPLEX within ten minutes, and the objective values of these solutions are usually within a percent or two of the optimal solution objective if not equal to it. For rapid analysis of different scenarios, the simulation may be run in a suboptimal mode in which the best solution found within 1000 branch-and-bound nodes of the first solution is used.
flights under control). In this sense, operation of the air transportation system fits the open-loop feedback paradigm [30] under which the ability to control the system in open loop, i.e., comply with the forecast, is vital to system performance.

5.6 Summarizing Results

The results of the simulation are displayed in a number of ways. The simulated passenger connection data for a single carrier at a time are displayed on a connection matrix showing passenger connections, passenger misconnections and flight delay, similar to the display of passenger connection data based on actual flights times in Figure 5-17. Airport throughput data are displayed on arrival-departure demand and capacity graphs such as Figure 5-6, with scheduled demand replaced by either planned operations or simulated operations. Finally, a number of statistics can be calculated and printed out.

The sets of flights on which statistics can be calculated are the sets of flights with contiguous simulated operating times at an airport. For instance, the flights simulated to arrive or depart ORD between 1400Z and 2000Z may be examined. The statistics are calculated over sets of flights identified by simulated time of operation so that the value produced within the time horizon of a ground delay program may be examined.

These printed statistics are measurements of the number of flights simulated to arrive or depart in the time period, the number of passengers simulated to have flown during the time period, the number of passenger connections made between flights within the time period, the total number of passenger connections scheduled, the combined number of minutes of flight delay experienced by the flights in the set, the number of minutes of passenger delay incurred (including delay caused by misconnections), and finally, the value of the user’s objective function realized by operations within the period (which is often different from the value planned because the simulated events do not necessarily follow the plan). The statistics are calculated individually for each airline and are also totalled for the combined group of airlines.
5.7 Data

The data used in the simulation for the study are derived from several sources. One source is the Airline Service Quality Performance (ASQP) database, assembled by the Department of Transportation for assessing on-time performance of all airlines that carry more than one percent of ticketed passengers in the USA. Each record of ASQP data contains the flight number and the times the flight was scheduled to push back and arrive according to the Official Airline Guide (OAG) and according to the Computer Reservation System (CRS). If the record corresponds to a flight that was not canceled, it also contains the aircraft tail number, the times at which the aircraft pushed back from the gate, lifted off from the runway, landed and arrived at the gate. These times are referred to as the out, off, on and in times, respectively.

The simulation also maintains a database of equipment types and available seats by tail number. This database allows the simulation to produce passenger itineraries for the entire schedule based on the seats available on each flight, since the ASQP data contains the tail number of aircraft that flew. For aircraft without tail numbers in the database (typically associated with canceled flights), the simulation assigns the flight an aircraft type randomly from the types used in the flight's origin-destination market. The probability that a flight will be assigned a specific aircraft type is proportional to the number of aircraft of that type flown between the flight's origin and destination. Passenger itinerary generation is covered in detail in Section 5.4.

The simulation maintains all times as days since a reference day and hours and minutes in Coordinated Universal Time (UTC). This time format allows direct comparisons of time in different time zones and of times before midnight to those after. Since ASQP data are in local time and explicitly contain the day of scheduled departure only, some conversions and inferences are made to convert to UTC and add the day to each time explicitly. To assist in making these inferences and for graphical display purposes, the simulation maintains a database of information on airports, including the time zone, latitude and longitude of each airport. When the simulation needs information about an airport that is not contained in the database, it looks up
the latitude and longitude from the web site http://www.airnav.com/, guesses the
time zone by geographic location and by scheduled travel times to and from other
locations with known time zones, and suggests that the user add the inferred airport
information to the airports database.

The simulation is written in C++. It was developed using the free egcs develop-
ment environment on a Linux workstation. It makes extensive use of the HP
Standard Template Libraries for data storage and manipulation. The user interface
is built using the Mesa GL libraries for three dimensional graphics and the GLUT
user interface and menu libraries. Robert Davies’ Newran02 random variable libraries
[9] were used to generate random numbers according to various distributions, and the
CPLEX Callable Libraries are used to solve the airline optimization models.

5.7.1 Comparison of Generated Itineraries and Real Itineraries

One major US airline generously provided passenger connection data it uses for re-
search in fleet assignment. These data were used to validate the passenger connection
generation algorithm. The provided data consist of a listing of passenger itineraries
and for each itinerary the number of passengers desiring that itinerary averaged over
a week in July 1997. By itinerary, we mean the sequence of flight legs serving the
passenger’s travel needs from origin to ultimate destination.

The capability to load this passenger data was added to the simulation. This
capability allows comparison of passenger connection data generated by the simulation
with the airline-supplied data. Figures 5-18 through 5-21 show the provided and
generated passenger data in connection matrices for two hub airports\textsuperscript{9}. Of course,
the data do not agree exactly; the simulation has no knowledge of many factors
that determine demand. In addition, the airline-provided data do not represent the
number of passengers that boarded each flight or that booked tickets. Rather, they
represent uncapacitated demand data: there is greater demand for some itineraries
than capacity. Finally, in the data supplied by the airline, passengers are present only

\textsuperscript{9}In the generated passenger connection data displayed in the connection matrices, the load factor
has been matched to the load factor in the airline-provided data.
on a large subset of the flights scheduled in the OAG; that is, there are significant numbers of flights and passenger itineraries without any demand in the data. As a result, the load factor calculated as demand measured in the airline-supplied data divided by seats available according to the schedule is small: the average load factor for the airline’s flights through airport A was only 45%, and the average load factor for the flights through airport B was 48%. True numbers for July, one of the times of the year with the highest load factors, are almost certainly over 80%, perhaps as high as 95%.

Another comparison between the airline-provided data and the simulation-generated data is given in Figures 5-22 and 5-23, which show histograms of the scheduled connection time desired by passengers in the airline-provided data and booked by the passengers in the simulation-generated data for the two airports. The means and
Figure 5-19: Airline-Provided Passenger Connections at Airport A
Figure 5-20: Generated Passenger Connections at Airport B
Figure 5-21: Airline-Provided Passenger Connections at Airport B
<table>
<thead>
<tr>
<th>Simulated Connections Airport A</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>57.8</td>
<td>53.2</td>
<td>26.2</td>
</tr>
<tr>
<td>Simulated Connections Airport B</td>
<td>63.7</td>
<td>24.2</td>
</tr>
<tr>
<td>Airline-Provided Connections Airport A</td>
<td>64.8</td>
<td>29.2</td>
</tr>
<tr>
<td>Airline-Provided Connections Airport B</td>
<td>35.2</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.8: Connection Time Statistics

standard deviations of scheduled passenger connection time are shown in Table 5.8.

Note that the simulated passenger connection time mean at airport B is a bit higher than the provided connection time data for the airline. As can be seen in Figure 5-23, much of the discrepancy occurs below thirty minutes connection time, where the provided data have more passenger connections than the simulated data do, and in the vicinity of 100 minutes, where the simulated data have more connections. Some of the discrepancy in the 100-minute region may be due to the fact that the provided data are uncapacitated demand data rather than actual booking data. Had the demand been forced to be within capacity, some passengers would have been required to make longer connection times to book flights with space available.

Whereas the generated data do not match the provided data perfectly, the fit is quite good. Notice, for instance, that the histogram of simulated connection time for each airport matches the histogram of provided connection time better than either matches the simulated or provided histogram for the other airport. Furthermore, the purpose of the simulation is not to model the airlines in perfect accuracy and detail as they behave today. Rather, the purpose of the simulation is to generate a range of plausible airline objectives and constraints and to compare the performance of the system under ADCAM with the performance of the system under CDM across this range.

In this spirit, perhaps the most appropriate test one can perform is to see whether the relative performances of ADCAM and today's CDM Ground Delay Program are similar using real data and provided data. This test was carried out, producing further evidence that the results of Chapter 6 based on simulation-generated passenger data
are similar to the results that would be realized using airline-provided data. The results of this test are presented in Section 6.4, so that they can be described in relation to the analysis results of Chapter 6.
Figure 5-23: Histogram of Passenger Connection Time at Airport B
Chapter 6

ADCAM Analysis

The description of ADCAM in Chapter 4 explains the benefits of ADCAM relative to CDM qualitatively. This chapter provides a quantitative comparison of the two methods using the analysis infrastructure described in Chapter 5.

Performance of the air transportation system operating under the rules of ground delay programs today is compared to performance under the ADCAM ground delay program and measured in terms of the metrics described in Section 5.6. The metric of primary interest is the objective value achieved by the airlines according to their own objective functions. The other metrics – passenger connections achieved, passenger and flight delay, and so forth – are used to understand the way in which the airline objectives were achieved.

These metrics are used to compare performance of the two methods in three analyses. Analysis One compares the methods' performance as a function of the degree of capacity insufficiency, using schedule data from Chicago O'Hare (ORD) and Minneapolis-St. Paul (MSP). Analysis Two examines their relative performance as a function of the degree to which arrivals and departures at the airport interact, using the schedule for ORD. Analysis Three examines performance as the airlines' objectives are varied, also using the ORD schedule. Schedule data from July 10, 1997 are used for all analyses.

A final test of the applicability of using simulation-generated passenger itinerary data is performed in Section 6.4. That section tests the hypothesis that the simulation-
generated passenger connection data provide results representative of those that would be achieved with actual passenger connection data, by running ADCAM and CDM with airline-provided passenger data and comparing the results to those achieved at the same airport and under the same capacity scenarios using simulation-generated passenger data.

6.1 Analysis One: Performance as a Function of Capacity

Analysis One uses the demand scheduled through Chicago O'Hare and Minneapolis-St Paul on July 10, 1997 with a series of capacity forecasts exhibiting some arrival-departure interaction. The forecasts are for unlimited capacity until 1400Z, for limited capacity from 1400Z to 2000Z, and for the airport to cease operations after 2000Z. Four different capacity forecasts are used at each airport, differing only by the amount of capacity available between 1400Z and 2000Z. Between 1400Z and 2000Z in each forecast the capacity is in the form shown in Figure 6-1, but differs in magnitude from forecast to forecast. The forecast with the greatest capacity during this period at O'Hare predicts that no more than sixteen arrivals, sixteen departures, or twenty-three operations in total will be able to operate in any ten-minute period between 1400Z and 2000Z. The forecast with greatest capacity at Minneapolis-St Paul predicts that no more than 10 arrivals, 10 departures, or 14 operations in total may be operated in any ten-minute period between 1400Z and 2000Z. The other three forecasts are scaled down to roughly 70%, 50% and 30% of that capacity during the reduced-capacity period (the scaling is selected so that the capacity limits are an integral number of arrivals or departures).

The results of these runs are summarized in Table 6.1 which shows the total objective value achieved by all airlines under ADCAM and under CDM at the two airports under the four capacity profiles. Figures 6-2 and 6-3 show these same data at ORD and MSP respectively, with the objective value achieved on the vertical axis
Figure 6-1: Capacity Forecast Form Used in Analysis One

<table>
<thead>
<tr>
<th>ORD Capacity</th>
<th>ORD ADCAM GDP</th>
<th>ORD CDM GDP</th>
<th>MSP Capacity</th>
<th>MSP ADCAM GDP</th>
<th>MSP CDM GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>27,902,700</td>
<td>27,764,500</td>
<td>100%</td>
<td>12,913,700</td>
<td>12,757,400</td>
</tr>
<tr>
<td>68.75%</td>
<td>22,513,300</td>
<td>18,159,600</td>
<td>70%</td>
<td>12,066,100</td>
<td>11,165,400</td>
</tr>
<tr>
<td>50%</td>
<td>18,746,200</td>
<td>7,944,310</td>
<td>50%</td>
<td>9,543,540</td>
<td>6,093,200</td>
</tr>
<tr>
<td>31.25%</td>
<td>10,678,400</td>
<td>-1,289,780</td>
<td>30%</td>
<td>6,331,120</td>
<td>1,581,810</td>
</tr>
</tbody>
</table>

Table 6.1: Sum of Objective Value Achieved by Airlines in Analysis One

and the airport capacity on the horizontal. Notice that as airport capacity decreases, the objective value achieved by the airlines decreases. The bend in these capacity curves appears at the point where cancellations, mis-connections, and large amounts of delay become necessary; to the left of the bend, the capacity reduction has cost the airlines small amounts of delay, whereas to the right, capacity reduction results in the relatively higher cost of flight cancellation. Notice also how objective value achieved by the airlines decreases more slowly under ADCAM than under CDM as airport capacity decreases. This difference is noticeable in Figures 6-2 and 6-3 as the gap between the CDM curve and the ADCAM curve.

To the left of the bend in the capacity curves, ADCAM allows the airlines to achieve a greater fraction of their passenger connections than CDM for the same level of delay, a lower level of delay for the same connectivity, or a mixture of the two
Figure 6-2: Analysis One Results at ORD

depending on the airlines' objective functions. This is a direct result of the airlines' ability to trade arrival capacity for departure capacity according to the overall airport capacity curve as outlined in Section 4.5; maintaining the relative timing of the arrival and departure banks allows each airline to achieve better connectivity without using more airport capacity.

To the right of the bend in the capacity curves, the dominant factors causing the difference between ADCAM and CDM realized objective values are the airlines' improved model of overall airport capacity under ADCAM and the fact that departure capacity has been allocated. Under the model of CDM, one airline might overload the airport with a delayed departure push, resulting in delays to all airlines' operations that might have been avoided through relatively inexpensive alternative operational strategies. Under ADCAM, the same situation can be approached through cancellation more in line with the capacity forecast and through a redistribution of delay. For instance, the airline with the poorly timed departure push might sacrifice some
of the lightly loaded passenger connections, sending certain departures out closer to schedule to take advantage of departure capacity that would otherwise go unused. Under CDM, not only does the airline operating the delayed departure push have little incentive to avoid causing trouble for other airlines’ operations, it has no information about the impact its choice of timing might have. Under ADCAM, the information and incentive are provided through the arrival-departure allocation of airport capacity.

The ability of one airline to have a large impact on the operations of the others under the current allocation scheme adds uncertainty to every airline’s operations. The airlines plan operations according to their assumptions of available capacity. In the simulation, those assumptions are very nearly correct under ADCAM but may be far off under CDM, resulting in instances where there is a big discrepancy between the objective value the airline expects to achieve through its plans and that realized. The 31.25% capacity scenario at ORD under CDM shows clear evidence of this – the
airlines would not plan to operate a negative-valued solution when the zero-valued solution of canceling all flights is available!

In practice, the airlines’ assumptions of available capacity would probably be better than those in the simulation. Recall from Section 5.4 that the airlines’ models of airport capacity in the simulation are those given them by the FAA as flow management constraints. Thus, under CDM, the airline plans operations assuming an unlimited departure capacity, resulting in long departure queues (and because the airport is modeled as a single server with arrival-departure interactions, airborne delays result as well). On the other hand, sacrificing one’s own departure timing for the good of one’s competitors does not make business sense. Without ADCAM or some form of departure capacity allocation, an airline has incentive to push its departures to get into the departure queue just before the queue gets long.

Tables 6.2 and 6.3 show the breakdown of the values of Table 6.1 according to airline. The trends identified for the overall industry hold, for the most part, within each airline’s results. The exceptions are with those airlines with sufficiently few flights that their performance numbers are dependent on specific periods of time within the ground delay program. For instance, America West had only one operation during the ground delay program, a departure scheduled for 16:55 (the middle of the ground delay program period). As capacity was reduced under CDM, the departure spent more and more time in the departure queue for the same push-back time. Under ADCAM, the allocated departure time was pushed late enough that the airline decided to cancel its flight in the 50% and 31.25% scenarios.

6.2 Analysis Two: Sensitivity to Objective Function

Analysis Two is performed on schedule data for Chicago O’Hare from July 10, 1997. The 50% capacity forecast from Analysis One is examined. Both ADCAM and CDM are run using each of four objective functions for the airlines.
<table>
<thead>
<tr>
<th>Airline</th>
<th>100% ADCAM</th>
<th>100% CDM</th>
<th>68.75% ADCAM</th>
<th>68.75% CDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>American</td>
<td>9,925,440</td>
<td>9,842,960</td>
<td>7,502,300</td>
<td>6,425,730</td>
</tr>
<tr>
<td>Continental</td>
<td>476,612</td>
<td>476,810</td>
<td>247,120</td>
<td>292,610</td>
</tr>
<tr>
<td>Delta</td>
<td>1,097,310</td>
<td>1,093,920</td>
<td>794,356</td>
<td>652,376</td>
</tr>
<tr>
<td>America West</td>
<td>69,000</td>
<td>69,000</td>
<td>48,300</td>
<td>56,580</td>
</tr>
<tr>
<td>Northwest</td>
<td>690,964</td>
<td>696,343</td>
<td>486,177</td>
<td>407,838</td>
</tr>
<tr>
<td>TWA</td>
<td>376,796</td>
<td>375,968</td>
<td>207,566</td>
<td>216,658</td>
</tr>
<tr>
<td>United</td>
<td>14,567,600</td>
<td>14,513,200</td>
<td>12,764,100</td>
<td>9,730,950</td>
</tr>
<tr>
<td>US Airways</td>
<td>699,000</td>
<td>696,324</td>
<td>463,362</td>
<td>376,838</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Airline</th>
<th>50% ADCAM</th>
<th>50% CDM</th>
<th>31.25% ADCAM</th>
<th>31.25% CDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>American</td>
<td>6,379,470</td>
<td>2,632,660</td>
<td>3,447,750</td>
<td>-330,133</td>
</tr>
<tr>
<td>Continental</td>
<td>273,864</td>
<td>99,374</td>
<td>141,031</td>
<td>23,414</td>
</tr>
<tr>
<td>Delta</td>
<td>547,810</td>
<td>220,456</td>
<td>227,406</td>
<td>-104,654</td>
</tr>
<tr>
<td>America West</td>
<td>0</td>
<td>23,460</td>
<td>0</td>
<td>-75,900</td>
</tr>
<tr>
<td>Northwest</td>
<td>333,058</td>
<td>154,289</td>
<td>185,506</td>
<td>-75,564</td>
</tr>
<tr>
<td>TWA</td>
<td>121,312</td>
<td>110,980</td>
<td>82,134</td>
<td>33,300</td>
</tr>
<tr>
<td>United</td>
<td>10,750,000</td>
<td>4,579,600</td>
<td>6,393,820</td>
<td>-713,026</td>
</tr>
<tr>
<td>US Airways</td>
<td>340,666</td>
<td>123,498</td>
<td>200,776</td>
<td>-47,220</td>
</tr>
</tbody>
</table>

Table 6.2: Objective Values Achieved by Airlines in Analysis One at ORD

<table>
<thead>
<tr>
<th>Airline</th>
<th>100% ADCAM</th>
<th>100% CDM</th>
<th>70% ADCAM</th>
<th>70% CDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>American</td>
<td>535,546</td>
<td>531,514</td>
<td>471,386</td>
<td>456,530</td>
</tr>
<tr>
<td>Continental</td>
<td>178,500</td>
<td>176,160</td>
<td>171,360</td>
<td>166,680</td>
</tr>
<tr>
<td>Delta</td>
<td>359,964</td>
<td>356,904</td>
<td>310,290</td>
<td>311,430</td>
</tr>
<tr>
<td>America West</td>
<td>50,000</td>
<td>50,000</td>
<td>50,000</td>
<td>50,000</td>
</tr>
<tr>
<td>Northwest</td>
<td>10,455,200</td>
<td>10,321,800</td>
<td>9,924,080</td>
<td>9,075,420</td>
</tr>
<tr>
<td>TWA</td>
<td>263,100</td>
<td>257,520</td>
<td>203,094</td>
<td>198,714</td>
</tr>
<tr>
<td>United</td>
<td>753,893</td>
<td>746,939</td>
<td>703,455</td>
<td>682,581</td>
</tr>
<tr>
<td>US Airways</td>
<td>317,500</td>
<td>316,612</td>
<td>232,472</td>
<td>224,018</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Airline</th>
<th>50% ADCAM</th>
<th>50% CDM</th>
<th>30% ADCAM</th>
<th>30% CDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>American</td>
<td>323,244</td>
<td>306,336</td>
<td>173,418</td>
<td>198,878</td>
</tr>
<tr>
<td>Continental</td>
<td>94,740</td>
<td>93,060</td>
<td>8,580</td>
<td>4,000</td>
</tr>
<tr>
<td>Delta</td>
<td>149,400</td>
<td>210,450</td>
<td>33,400</td>
<td>85,580</td>
</tr>
<tr>
<td>America West</td>
<td>38,000</td>
<td>50,000</td>
<td>14,000</td>
<td>50,000</td>
</tr>
<tr>
<td>Northwest</td>
<td>8,230,340</td>
<td>4,793,930</td>
<td>5,616,610</td>
<td>1,009,760</td>
</tr>
<tr>
<td>TWA</td>
<td>86,900</td>
<td>115,074</td>
<td>119,574</td>
<td>9,014</td>
</tr>
<tr>
<td>United</td>
<td>520,908</td>
<td>404,436</td>
<td>338,616</td>
<td>182,918</td>
</tr>
<tr>
<td>US Airways</td>
<td>100,000</td>
<td>119,918</td>
<td>26,920</td>
<td>41,660</td>
</tr>
</tbody>
</table>

Table 6.3: Objective Values Achieved by Airlines in Analysis One at MSP
Under objective function 1, delay of an airframe costs the airline only the amount that the resulting passenger-minutes of delay cost. Thus, under objective function 1, a solution in which a departure carrying one hundred passengers is delayed ten minutes is equally desirable to a solution in which the flight operates on time and leaves ten passengers to wait one hundred minutes for the next connection – both cost one thousand passenger-minutes of delay.

Under objective functions 2, 3 and 4, the delay of the airframe costs twice as much, five times as much, and twenty times as much, respectively, as the resulting passenger-minutes of delay. The justification for this variation in cost function is that delays on the flight may cause additional downstream delays; depending on the time of day of the flight, the downstream delays may be significant. In [5], such multiplying factors are calculated on the network of American Airlines assuming that crew and equipment changes do not occur.

Table 6.4 summarizes the results of Analysis Two, showing the delay multiplier, objective achieved under ADCAM and objective achieved under CDM. Figure 6-4 shows a plot of the fraction of the ADCAM solution achieved under CDM as a function of the delay cost multiplier. According to the plot and the table, one can see that ADCAM offers the greatest improvement over CDM at moderate values of the delay multiplier, where the airline’s objective is to operate on time and connect passengers as well as possible. At very large delay multiplier values, the main objective is to operate to schedule; although ADCAM does improve on CDM through better modeling of airport capacity, the improvement is not as pronounced as the improvement ADCAM offers the airline in its ability to maintain passenger connections. As the delay multiplier values decrease further, the relative benefit of ADCAM decreases somewhat again as the cost of the additional delay incurred by the airlines under CDM to achieve the same passenger connections achieved under ADCAM is not as great. ADCAM reduces the delay required to achieve connections and the mis-connections required to achieve a given level of on-time performance.

142
<table>
<thead>
<tr>
<th>Airline</th>
<th>1x ADCAM</th>
<th>1x CDM</th>
<th>2x ADCAM</th>
<th>2x CDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>1,482 (32%)</td>
<td>595 (29%)</td>
<td>2,769 (32%)</td>
<td>1,178 (32%)</td>
</tr>
<tr>
<td>CO</td>
<td>71 (1.6%)</td>
<td>10 (0.5%)</td>
<td>132 (1.6%)</td>
<td>18 (0.5%)</td>
</tr>
<tr>
<td>DL</td>
<td>115 (2.5%)</td>
<td>63 (3%)</td>
<td>220 (2.6%)</td>
<td>131 (3.6%)</td>
</tr>
<tr>
<td>HP</td>
<td>0 (0%)</td>
<td>6 (0.3%)</td>
<td>0 (0%)</td>
<td>16 (0.4%)</td>
</tr>
<tr>
<td>NW</td>
<td>94 (2.1%)</td>
<td>52 (2.5%)</td>
<td>173 (2.0%)</td>
<td>114 (3.1%)</td>
</tr>
<tr>
<td>TW</td>
<td>28 (0.6%)</td>
<td>20 (1.0%)</td>
<td>50 (0.6%)</td>
<td>38 (1.0%)</td>
</tr>
<tr>
<td>UA</td>
<td>2,691 (59%)</td>
<td>1,305 (63%)</td>
<td>5,033 (59%)</td>
<td>2,157 (58%)</td>
</tr>
<tr>
<td>US</td>
<td>79 (1.7%)</td>
<td>19 (0.9%)</td>
<td>159 (1.9%)</td>
<td>49 (1.3%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>4,563</strong></td>
<td><strong>2,071</strong></td>
<td><strong>8,525</strong></td>
<td><strong>3,704</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Airline</th>
<th>5x ADCAM</th>
<th>5x CDM</th>
<th>20x ADCAM</th>
<th>20x CDM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>6,391 (33%)</td>
<td>2,605 (32%)</td>
<td>24,763 (33%)</td>
<td>22,630 (36%)</td>
</tr>
<tr>
<td>CO</td>
<td>267 (1.4%)</td>
<td>131 (1.6%)</td>
<td>1,327 (1.8%)</td>
<td>992 (1.6%)</td>
</tr>
<tr>
<td>DL</td>
<td>536 (2.8%)</td>
<td>298 (3.7%)</td>
<td>2,114 (2.8%)</td>
<td>2,299 (3.7%)</td>
</tr>
<tr>
<td>HP</td>
<td>0 (0%)</td>
<td>29 (0.4%)</td>
<td>0 (0%)</td>
<td>329 (0.5%)</td>
</tr>
<tr>
<td>NW</td>
<td>407 (2.1%)</td>
<td>223 (2.7%)</td>
<td>1,587 (2.1%)</td>
<td>1,878 (3.0%)</td>
</tr>
<tr>
<td>TW</td>
<td>115 (0.6%)</td>
<td>72 (0.9%)</td>
<td>439 (0.6%)</td>
<td>783 (1.2%)</td>
</tr>
<tr>
<td>UA</td>
<td>11,303 (58%)</td>
<td>4,680 (57%)</td>
<td>43,295 (58%)</td>
<td>32,125 (51%)</td>
</tr>
<tr>
<td>US</td>
<td>397 (2.0%)</td>
<td>101 (1.2%)</td>
<td>1,591 (2.1%)</td>
<td>1,790 (2.8%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>19,419</strong></td>
<td><strong>8,142</strong></td>
<td><strong>74,766</strong></td>
<td><strong>62,828</strong></td>
</tr>
</tbody>
</table>

Table 6.4: Objective Values (in 1000's) Achieved by Airlines in Analysis Two at ORD
Figure 6-4: Objective Value of ADCAM Divided by Objective Value of CDM as Delay Multiplier Varies, 50% Capacity at ORD
6.3 Analysis Three: Performance as a Function of Arrival-Departure Interaction

Analysis Three compares ADCAM to CDM under a set of capacity forecasts that differ in the degree to which arrivals and departures interact. The analysis is performed on the schedule at O'Hare on July 10, 1997. The forecasts are similar to those used in Analyses One and Two in that capacity is unrestricted until 1400Z, at which time it is reduced below demand, and operations occurring after 2000Z are not considered to produce value. The forms of the capacity forecasts between 1400Z and 2000Z are shown in Figure 6-5. The first one exhibits a direct tradeoff between arrival and departure capacity, as might an airport operating only a single runway in visual conditions. The second forecast employs the form used in the previous analyses. The remaining three forecasts allow progressively less direct exchange of arrivals for departures. Forecasts 4 and 5 are representative of capacity at an airport with separate runways for arrivals and departures and little or no interaction between these operations.

Table 6.5 summarizes the results of Analysis Three. Notice that ADCAM provides better performance even in the case of no arrival-departure interaction. Part of the performance improvement is due to the fact that ADCAM rations departures as well as arrivals, thus ADCAM would allow an airline to reserve departure slots and spend time at the gate normally spent waiting in a departure queue. Not only would this save fuel and harmful environmental emissions, but it would also allow passengers on delayed arrivals to connect to their departure flights.

Some of this performance improvement in capacity form 5 is illusory, however. Under ADCAM, the airlines ensure that available arrival and departure capacities are used in the best possible way. Under CDM, the FAA must perform this function, but in the simulation model of the airport the FAA merely serves operations in a first-planned, first-served manner until operations are infeasible. The consequence of this is that the simulation of an airport with independent arrivals and departures may allow fewer operations in a period than the real system would under similar circumstances.
Figure 6-5: Capacity Forecast Forms Used in Analysis Three

In the simulation, the capacity of the airport for a time period is considered to have been reached when the arrival limit, the departure limit, or both have been achieved. At a real airport with independent arrivals and departures, arrivals and departures are served by separate servers, each of which reaches its capacity limit if demand is sufficient.

Another issue complicating these results is that the performance of ADCAM relative to CDM is dependent upon the amount of capacity available according to the relationship shown in Section 6.1. The capacity forms 1 through 4 in this section support at most five arrivals and six departures simultaneously or six arrivals and five departures simultaneously, but differ greatly in their arrival-only and departure-only capacities. Thus, some of the difference in performance across capacity forms seen in Table 6.5 is likely due to the difference in amount of capacity available rather than due to the shape of the arrival-departure capacity curve. Nevertheless, these results do suggest that ADCAM outperforms CDM in even the case in which arrivals and departures are independent.
### Table 6.5: Sum of Objective Value Achieved by Airlines in Analysis Three

<table>
<thead>
<tr>
<th>Scenario</th>
<th>ADCAM Objective</th>
<th>CDM Objective</th>
<th>Percent Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21,441,000</td>
<td>11,408,200</td>
<td>87.9%</td>
</tr>
<tr>
<td>2</td>
<td>19,779,300</td>
<td>8,496,970</td>
<td>133 %</td>
</tr>
<tr>
<td>3</td>
<td>18,403,600</td>
<td>7,650,160</td>
<td>140 %</td>
</tr>
<tr>
<td>4</td>
<td>17,206,100</td>
<td>5,047,220</td>
<td>240 %</td>
</tr>
<tr>
<td>5</td>
<td>19,172,400</td>
<td>6,649,540</td>
<td>188 %</td>
</tr>
</tbody>
</table>

#### 6.4 Analysis Four: Simulated Versus Provided Passenger Data

As explained in Section 5.7.1, the applicability of the results presented in this section rely on the assumption that the performance of ADCAM relative to CDM under simulation-generated passenger demand data is similar to that using real data. To this end, two series of simulation experiments were conducted at a hub airport dominated by the airline-provided data. In the first series, simulation-generated passenger data were used for all airlines at the hub to run ground delay programs with four capacity forecasts ranging from one that is adequate to serve the demand to one that could handle only about half of the demand. In the second series, the same experiments were performed using airline-provided data. The results, summarized in Table 6.6 and Figure 6-6, require some explanation.

Table 6.6 shows, for both series of experiments, the objective values realized by the airline that provided the data normalized by the full-capacity ADCAM GDP objective value of the series. The two left columns present the results for the simulated passenger series of experiments, whereas the two right columns are the results using the provided passenger data.

Recall that the airline-provided passenger data appears to be missing some passengers. Many of the scheduled flights show no passenger bookings in this data. As a result, there are fewer flights operated under the carrier-provided passenger data set than under the simulation-generated passenger data set since there was no benefit to operating many of the flights carrying no passengers.
<table>
<thead>
<tr>
<th>Capacity</th>
<th>Simulated Pax ADCAM GDP</th>
<th>Simulated Pax CDM GDP</th>
<th>Provided Pax ADCAM GDP</th>
<th>Provided Pax CDM GDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>100%</td>
<td>97.3%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>70%</td>
<td>97.3%</td>
<td>85.3%</td>
<td>98.3%</td>
<td>95.3%</td>
</tr>
<tr>
<td>50%</td>
<td>82.1%</td>
<td>57.2%</td>
<td>89.7%</td>
<td>78.7%</td>
</tr>
<tr>
<td>30%</td>
<td>56.3%</td>
<td>22.9%</td>
<td>63.2%</td>
<td>25.1%</td>
</tr>
</tbody>
</table>

Table 6.6: Objective Value Achieved by Data Provider under CDM and ADCAM with Simulated and Provided Passenger Connections

As a result of there being fewer flights in the carrier-provided experiment set, the airline is able to achieve a greater fraction of its overall objectives with a smaller amount of capacity (see Table 6.6). If one compares the relative ability of CDM and ADCAM to achieve value within limited capacity, one sees that in both cases as the ADCAM value decreases from 100%, the CDM value decreases faster, and that in both cases the rate at which the gap grows between ADCAM objective value and CDM objective value is similar. Figure 6-6 shows this trait using the data from Table 6.6 by plotting the objective value achieved under CDM versus the objective value achieved under ADCAM for both series of experiments. The similarity of the performance using simulation-generated data to that using airline-provided data suggests that the results of this chapter are indeed representative of results that would be seen using real passenger data.
Figure 6-6: Objective Value Achieved under CDM versus ADCAM for Simulation-Generated and Airline-Provided Passenger Data
Chapter 7

Conclusions

This thesis addresses the simultaneous solution of problems faced by the FAA and airlines during ground delay programs. The FAA and airlines solve very different problems during ground delay programs, yet the solution of each one's problems has significant impact on the problems faced by the others. The airline and FAA problems cannot be solved together as a single, combined optimization problem in practice because the FAA and airlines have separate objectives, information and control authority. The FAA lacks the information necessary to discern the impact of arrival and departure timing on an airline's network of flights and service to passengers, while the airlines lack the global perspective to restrict their operations to meet system capacity when necessary. These fundamental differences necessitate a decomposition of the combined problem of inter-airline and intra-airline capacity allocation. The airlines must apply their objectives, information, and control authority to some aspects of the problem while the FAA focuses on other aspects.

Although decomposition is necessary in the solution of these problems, there is flexibility in the form of the decomposition employed. In this thesis, we measure the quality of decomposition in the same way that mathematical programming decomposition quality is measured: by the ability of the decomposed solution technique to achieve the objectives of the non-decomposed problem. The objective of the combined problem within the framework of this thesis is the maximization of the economic production of the air transportation system, measured as the sum of the objectives of
the users. The use of this objective is a novel contribution of this thesis.

The thesis also contributes two new methods of allocating airport capacity, each of which is based on a different decomposition of the combined capacity allocation problem. These methods are judged according to their ability to achieve the objective of maximizing the sum of user objectives. Section 7.1 describes the capacity allocation methods developed in this thesis and discusses their ability to maximize the sum of user objectives. Section 7.2 explores practical issues concerning the implementation of these decomposition methods in the air transportation system, and Section 7.3 discusses interim changes to the air traffic management system that could improve operations during development of the permanent system improvements in this thesis. Finally, Section 7.4 outlines areas for further research.

7.1 The Capacity Allocation Methods and Their Potential

The two capacity allocation methods developed in this thesis are the Objective-Based Allocation Method (OBAM), developed in Chapter 3, and the Arrival-Departure Capacity Allocation Method (ADCAM), developed in Chapter 4. Section 7.1.1 describes OBAM and discusses its potential effectiveness. Section 7.1.2 treats ADCAM.

7.1.1 OBAM: Incorporation of User Objectives into the FAA Problems

OBAM is a capacity allocation method in which the FAA incorporates user objectives into the problem it solves of allocating capacity to airlines. OBAM theoretically achieves complete efficiency in the allocation of arrival slots. It is based on the application of the Groves mechanism to the problem of determining an allocation of arrival capacity to users. It broadens the conditions under which incentive compatibility is reached by the mechanism proposed in [23].

In theory, the objective of maximizing the sum of user objectives through the allo-
cation of airport capacity cannot be accomplished better than through an application of the Groves mechanism. Through the Groves mechanism, it has been shown here that the airlines have incentive to communicate their objectives and constraints to the FAA as accurately as possible. With this information, the FAA solves the combined problem of allocating capacity to maximize the combined objectives of the airlines such that their constraints are satisfied. Because the problem is solved centrally\(^1\) with the best possible information, it will produce the best possible allocations.

The potential benefit of optimizing the allocation in this way is enormous. However, it would be very difficult to implement such a mechanism. Achieving the full benefit of OBAM would require each airline to develop the capability to assign numerical value to each flight-slot assignment, in some cases requiring a radical change in the current way of doing business. Furthermore, since the mechanism requires that the airlines pay for the airport capacity they use, its acceptance is likely to face a fate similar to the U.S. Government's many fruitless recent efforts to introduce air traffic management user fees.

### 7.1.2 ADCAM: Incorporation of Departures into the Capacity Allocation

ADCAM, developed in Chapter 4, represents an improvement to the decomposition of decision-making during ground delay programs that does not require the airlines to pay fees or to communicate their complete set of objectives and constraints to the FAA. It improves on the decomposition in current use in several ways: (1) by allocating departure capacity as well as arrival capacity; (2) by relying on a higher-fidelity model of airport capacity than that in use today; and, (3) by allowing the airlines to determine the dynamic distribution of airport capacity between arrivals and departures.

\(^1\)Note that although the problem is solved centrally, the overall problem-solving structure remains a decomposition: the centrally-solved solution is in addition to the optimization problems solved by the airlines. The Groves mechanism structures the decomposition such that it is optimal for the airlines to provide complete, accurate information to the FAA, yet the airlines continue to solve their own optimization problems.
Allocation of Departure Capacity

Currently, departure capacity is allocated on a first-come, first-served basis. For a flight to receive its share of departure capacity, it must often wait in a queue that forms on the airport’s taxiways. The time spent waiting in queue is costly not only because the engines are running and burning fuel, but also because it requires the use of a deprecating aircraft without serving any real purpose and erodes the good will of the passengers.

ADCAM allocates departure capacity among the users, eliminating the first-come, first-served discipline that leads to physical departure queues. Under ADCAM, an airline gains the flexibility to use its aircraft productively during the times its aircraft must currently spend in departure queues. For instance, the airline might keep the aircraft at the gate longer so that passengers on late-arriving flights could make their connections. Similarly, in some cases the airline could provide the same service with fewer aircraft. The airlines would probably invent unforeseen ways to take advantage of the flexibility departure capacity allocation would provide.

Increased Model Fidelity

ADCAM allocates airport capacity according to a model of airport capacity that captures the interactions between arrivals and departures. The precision of this capacity model allows flow management to restrict demand so that it matches the true airport capacity more accurately. Reduction of model error in any control system improves its performance guarantees; in air traffic flow management, the reduction of model error allows a simultaneous reduction of wasted airport capacity and MAR.

Airline Control over Distribution of Capacity Between Arrivals and Departures

The distribution of airport capacity between arrivals and departures is currently made by the FAA according to an arbitrary method. Rather than adjust the planned operations mix dynamically to meet the requirements of demand during flow management,
the FAA allocates half the capacity to arrivals and half to departures. As a result, it can be difficult for the airlines to maintain bank integrity; at a hub airport with interchangeable arrival and departure capacity, an airline could maintain better connectivity if all of its capacity were used for arrivals during its arrival banks and all of its capacity were used for departures during its departure banks. The best way for the airline to allocate capacity between arrivals and departures depends on many factors about which the FAA may have little or no information. Giving each airline control over the dynamic distribution of its share of capacity among arrivals and departures allows it to increase the economic benefit derived from the use of its share of capacity.

The results of applying ADCAM discussed in Chapter 6 suggest that these improvements may be significant. The simulated performance of the air transportation system during ground delay programs is improved remarkably through the use of ADCAM, particularly in cases of ground delay programs with severely reduced capacity. For instance, in simulation of a specific ground delay program modeled after current operations, the airlines collectively achieved only 29% of the objective value they would have achieved on a day with excess capacity and no ground delay program. When the same ground delay program was simulated under ADCAM, the airlines achieved 67% of their collective excess-capacity objective value (Table 6.1). The economic impact of such an improvement, were it realized in the real air transportation system, would be significant indeed.

7.2 Implementation of Capacity Allocation Methods

The airlines and the FAA collaborate on solutions to air traffic management problems routinely. Collaboration is vital since the problems solved by the FAA and the airlines are interdependent. CDM has formalized this collaboration and developed the computer and communications infrastructure necessary for effective collaboration.

Because the novel capacity allocation methods developed and analyzed in this the-
sis represent new decompositions of the combined capacity allocation problem, they involve new forms of collaboration. OBAM and similar methods that could be designed based on the Groves mechanism involve a very intensive form of collaboration. In OBAM, the FAA assesses fees from the users to coordinate their solutions; the fees modify the user objectives so that they are consistent with each other. Under this consistent set of objectives, each of the users has incentive to cooperate with the others to maximize the sum of user objectives. Competition between the users is reduced to a competition in which each user strives to achieve the greatest efficiency possible with its share of the system resources. This form of competition is economically desirable.

As mentioned above, implementation of OBAM or a related method would require considerable changes to the current way of doing business. Each airline would require the capability to express its objectives numerically in a specified format. Under OBAM, this format consists of a value function specifying the value for each flight’s arrival as a function of the time at which the flight arrives. It might be impossible to capture some aspects of the user’s objectives in such a format. For example, under OBAM one cannot express a dependence of value on the amount of time between the arrivals of two flights. Increasing the complexity of the objective function format to capture such characteristics is possible, however the amount of information that must be communicated as a result and the complexity of the problem the FAA is required to solve can increase dramatically. Furthermore, the implementation of OBAM or a similar allocation method would require the community to accept the fees the FAA would need to assess, and, as mentioned before, the U.S. government has recently tried numerous times without success to assess new forms of user fees.

Implementation of ADCAM, on the other hand, does not entail the difficulties associated with implementation of OBAM. It requires a form of collaboration that is less intensive and less invasive than OBAM; many of the system changes required to implement ADCAM have already been put in place under the existing CDM project, however, a number of additional issues remain. Many of these issues are significant. Nevertheless, the benefits that ADCAM would provide may justify the effort required.
to address these issues.

The following sections outline the technical issues that must be addressed to implement ADCAM. They are organized according to their functional relationship to ADCAM. The first two sections address the development of the inputs required by ADCAM: Section 7.2.1 discusses the capability to produce arrival-departure capacity forecasts required by ADCAM, and Section 7.2.2 discusses the input ADCAM requires from the airlines. The output of ADCAM is the desired allocation of airport arrival and departure capacity to flights. Section 7.2.3 addresses changes in FAA control procedures that would be helpful and changes that would be necessary to implement the allocation produced by ADCAM.

While the technical aspects of the issues outlined below would not require much time to address were significant effort brought to bear on them, some of the political, administrative, and procedural aspects of these issues might prove considerably more complicated. Section 7.3 discusses interim solutions that could achieve some of the benefit of ADCAM without much of this complexity.

7.2.1 Development of Arrival-Departure Capacity Forecasting Capability

The primary inputs to ADCAM are the capacity forecast and the user preferences. The capacity forecasts under ADCAM differ from the ones used in flow management today in that they represent constraints on both arrivals and departures.

In current GDP’s, officials from the ATCSCC and the airport confer to predict what the arrival capacity will be. The result of their conference is a forecast of capacity and an airport acceptance rate (AAR) appropriate for the forecast. A typical forecast might call for the AAR of the airport to be 32 aircraft per hour for four hours, followed by 48 aircraft for an hour and unrestricted thereafter. The ability to develop appropriate AAR’s has evolved from years of experience; ground delay program experts learn through trial and error what arrival rates work well under each weather condition they encounter.
Any significant change in flow management may require ATC to develop new guidelines for determining the appropriate acceptance rates. A recent instance of such a change was the implementation of CDM. When CDM began prototype operations at San Francisco International Airport, the controllers remarked that CDM was sending too many aircraft to the airport, resulting in excessive airborne holding times. In fact, CDM was sending the number of arrivals the controllers specified through the AAR. The problem was that the controllers had learned to inflate the AAR slightly above what the airport could accommodate, since before CDM was implemented unfilled arrival slots usually resulted in actual arrival rates lower than those specified by the AAR.

Similar issues will be faced should ADCAM be implemented. Under ADCAM, the format of the forecast is completely different; rather than an AAR, it consists of a sequence of arrival-departure capacity constraints similar to that in Figure 5-2. The FAA has developed the Engineered Performance Standards (EPS) which specify the arrival rate an airport can accommodate given the departure rate and vice-versa. However, EPS is used as a rough guide by the controllers when deciding whether a ground delay program is necessary. EPS is not designed to be a predictor of airport capacity.

The development of accurate arrival-departure capacity forecasts required under ADCAM would require significant effort by ATC. It would likely involve a continual learning process similar to the one in use today. A conservative approach to development of operational arrival-departure capacity forecasts would be to use an initial arrival-departure capacity forecast consisting of a rectangular feasible region in arrival-departure space with the upper-right hand corner set at the AAR used in current operations. The area of this forecast could be increased with experience to allow more arrivals than the current AAR at the expense of departures and vice-versa. Eventually the true airport capacity would be learned.

An alternative approach to develop arrival-departure capacity forecast capability that would yield near-term results would be to analyze historical airport operational data in a manner similar to that used in [15]. Whereas the emphasis of [15] was the
determination of the capacity limits of an airport, the work required to implement ADCAM would emphasize development of predictors of airport arrival-departure capacity. Capacity predictors would differ from estimates of capacity limits in that they would be designed to determine a reliably achievable capacity forecast, with emphasis on measuring the probability that the forecast will be achieved.

Development of such arrival-departure capacity predictors is not trivial. It requires analyzing the number of operations that were achieved at the airport in various weather conditions and airport configurations. It also requires analyzing the demand that lead to that number of operations. For instance, if one observes that an average of 30 arrivals and departures per hour are operated in the current system under certain conditions, that observation may be the result of a capacity limitation (demand exceeded capacity), or it may be the result of restrictive air traffic flow management constraints preventing more than 30 aircraft per hour from reaching the airport.

One indicator that an airport is accepting all the operations that it can accommodate is the presence of departure queues and terminal-area airborne delay. If one could measure the number of operations performed per unit of time while queues were present, that measurement would provide a better sample of data from which to develop predictors of airport capacity. The presence of departure queues can be inferred readily from ASQP data as described in [12]. A method developed with the help of Ryan "Rif" Rifkin and described by the author in [1] provides a measure of airborne delay in the system. An analysis could be performed with these measurements of queueing, ETMS data of arrivals and departures, and the configuration and weather conditions at an airport to develop capacity predictors. Of course, no matter how carefully such predictors were developed in advance, they would represent the capacity of the system operating as it did on the days the data were collected. Manual tuning of the sort performed today will be required to account for extraordinary circumstances and to improve the capacity estimates as the air transportation system evolves over time.
7.2.2 Development Required to Implement ADCAM

Development of new software and decision aids would be required to implement ADCAM. For ADCAM to function, the FAA must provide arrival-departure capacity forecasts to a computer that has access to current airline schedules. The computer would run ADCAM using these capacity forecasts and schedules, generating arrival-departure subcapacities for each of the airlines. The arrival-departure subcapacities and the ration-by-schedule assignment of flights to arrival and departure times would then be sent over the CDMNet to the airlines.

Simultaneously, the arrival times would be converted into departure slots at the origin airports of the arriving aircraft, and the departure times would be converted into departure slots at the GDP airport. These slots would be communicated to the appropriate air traffic control towers the same way that they are communicated today.

The airlines would need tools to form subcapacity-feasible reassignments of arrival and departure times to flights. These reassignments would be transmitted back across the CDMNet to the ATCSCC, generating revised arrival and departure times and slots for the affected aircraft.

Development of the software and decision aids required to implement ADCAM would build on the infrastructure already provided by CDM. Indeed, FSM runs on computers in the ATCSCC and at the airlines, performing precisely those communication functions that software implementing ADCAM would perform. ADCAM could be implemented by upgrading FSM with the ADCAM logic. Embodiment of the ADCAM logic in software would begin with the tools developed for this thesis.

7.2.3 Departure Control

Accomplishment of the capacity allocations determined through ADCAM requires some changes to the control exercised over departures in the system. Because ADCAM rations departure capacity as well as arrival capacity at an airport for which a GDP has been issued, it would be necessary to control the times of aircraft departures from the GDP airport as well as the times of departure of aircraft destined to arrive
at the GDP airport. Of course, in the current system the departure times of arriving aircraft are already controlled at their origin airports during GDP’s. However, control of the departures from the GDP airport is not practiced in the current system (except when those departures are bound for another airport experiencing a GDP).

The same procedures that are used to release aircraft for departure to meet arrival slots at other airports could be used to release aircraft for departure from a GDP airport. In fact, such procedures are in use at many airports in Europe today in all weather conditions. One caveat is that the arrival and departure times specified under ADCAM are times of runway arrival and departure. The ability to achieve accurate control of runway departure times may be improved through efforts to model and control the taxi process such as those reported in [19], [20] and [26]. These efforts should also contribute to arrival time control accuracy, since the bulk of arrival time uncertainty is due to uncertainty in the departure time, not uncertainty in the en-route flight time.

Another development that would improve system performance during ground delay programs run under the current system as well as those under ADCAM would be ATC enforcement of controlled departure times. Currently, departure times issued by the ATCSCC are treated as earliest allowed times of departure – flights are allowed to leave later than their controlled time of departure. Researchers of CDM have studied the compliance of flights with their controlled departure times and have found that many depart long after their assigned times. Late departures often cause the arrival rate at the airport to be low during the time period in which they are intended to arrive and too high during the time period in which they actually arrive.

Requiring compliance with the controlled departure times would not prevent a flight from departing late if unable to make its controlled time. Rather, it would require the flight to notify the FAA of its tardiness through a request for a departure slot at an appropriately delayed time. The notification of the delay would allow the FAA to plan accordingly; the FAA could allow other flights to use the capacity originally reserved for the delayed flight and could reserve capacity appropriate to the delay. The European air traffic control system, which enforces departure times in
precisely this way, is proof that the concept could be made to work.

7.3 Interim Solutions

The technical implementation issues associated with ADCAM may require several years to address, and non-technical concerns could add to this time. In the meantime, some of the benefits of ADCAM could be realized through relatively minor adjustments to current operating procedures. The most likely candidate interim solutions are implementation of a *virtual departure queue*, defined and described in Section 7.3.1, and adoption of dynamic arrival capacity estimates, described in Section 7.3.2. While these interim solutions may provide significant benefit, each is a local solution without the benefit of the global perspective of ADCAM. ADCAM integrates many more aspects of system operation in one consistent framework, allowing a unification of the treatment of arrivals and departures and an integration of strategic planning with tactical operations and tools, as discussed further in Section 7.4.2.

7.3.1 Virtual Departure Queue

Part of the benefit of ADCAM arises from the reduction of time aircraft must spend in departure taxiway queues. ADCAM allows a reduction in taxiway queues because it allows aircraft to reserve departure capacity before entering the taxiway queues. Other mechanisms that allow the reservation of departure capacity without requiring the aircraft to be in a physical departure queue could be implemented.

The term *virtual departure queue* has been coined to capture the concept of mechanisms, other than physical queues, that allow the reservation of departure capacity. In its simplest implementation, a basic virtual departure queue could be realized through a simple change in control procedure. The controller responsible for clearance of departing aircraft to taxi to the runway would provide that clearance only when the physical departure queue was small enough. While this scheme would reduce the amount of fuel wasted in taxi, it would not provide the airline with additional predictability and flexibility the way ADCAM does.
A more advanced type of virtual departure queue exists in Europe, where a flight is allocated departure capacity in the form of a departure slot before it enters the physical departure queue. Many of the European departure slots are reserved weeks or months in advance in a manner similar to the way in which slots at the four permanently-slotted airports in the United States are allocated. This approach to building a virtual queue provides the airline with more predictable departure times than the basic virtual departure queue outlined above. However, the procedure is not as flexible for the airline as it could be – a departure slot is specific to the flight and cannot be readily exchanged for the slot of another flight during periods of congestion except under special conditions.

A virtual departure queue that allocates departure slots in advance and allows an airline to choose the assignment of flights to departure slots might work very well in the United States. In Europe, where the departure slot is coordinated with en-route capacity constraints, swapping slots among flights might not work as well. Fortunately, en-route capacity constraints in the United States are less restrictive than in Europe.

This approach would provide the airlines more flexibility than the current system does. It would also reduce fuel burn, airframe use and environmental emissions by reducing physical departure queue length. However, the approach would still not allow the airlines the flexibility that ADCAM allows. Nor would it allow the airport arrival-departure mix to be adjusted dynamically as efficiently as ADCAM would. Section 7.3.2 suggests an interim solution that could help adjust arrival-departure mix dynamically.

7.3.2 Dynamic Arrival Capacity Estimates

An interim solution that may provide some of the ability of ADCAM to adjust the airport arrival-departure mix dynamically without requiring changes to departure control procedures is to employ dynamic arrival capacity estimates that account for the expected departure demand. Currently, the airport acceptance rate (AAR) during a ground delay program is set to a constant number of arrivals per hour over periods
during which the airport arrival-departure capacity is unchanged. Air Traffic Flow Management attempts to spread the arrivals evenly over each hour, as is explained in Chapter 5.

Under ADCAM the airport can provide extra arrival capacity at the expense of departure capacity or vice-versa. Arrival-departure capacity tradeoff is responsible for some of the benefit attributable to ADCAM. It allows the airport to accommodate the demand surges caused by the airlines' arrival and departure of banks.

A similar capability could be provided if the FAA had a good estimate of the airlines' desired response to the ground delay program. The FAA could predict how the airlines would allocate arrivals and departures if ADCAM were in use. Then, the FAA would set the arrival acceptance rate as a function of time to match the estimate of the airlines' demand for arrival capacity.

If the FAA's estimate of the airlines' demand were accurate, the AAR would be reduced in those periods of high demand for departure capacity, and the AAR would be increased in those periods of low demand for departure capacity. Current ATCSCC tools are capable of running ground delay programs based on a dynamically changing AAR. However, the FAA would need tools to determine how the AAR should vary. The tools could consist of optimization-based airline behavior models similar to the one developed in this thesis. Alternatively, simpler models could be used. Shumsky [31] found that aircraft turnaround times provide a reasonable predictor of the number of departures over time at Logan airport. Even a model that estimated departure demand by accounting only for average aircraft turnaround times would be far better than the current lack of any model.

7.4 Directions for Further Research

There are many directions future research could take. The ideas of Chapter 3 could be developed into practical methods for resource allocation. The implications of ADCAM for the development of new tactical decision aids in air traffic management, and the impact of new tactical decision aids on an air transportation system operating under
ADCAM merit further research. The simulation developed to compare ADCAM to current operations could be improved to take into account more of the detail of the air transportation system, and it could be used to study the effects of other changes to air traffic management. Finally, ADCAM could be improved to model airport capacity more accurately.

7.4.1 Further Development of OBAM and Related Methods

The potential benefit of methods related to OBAM is great. Not only do they allow the airlines the flexibility of operating to make the best possible use of their shares of constrained system resources, they also optimize the allocation of system resources to airlines.

The difficulty with these methods is that in order to optimize the allocation, they require the airlines to disclose their objectives and constraints to the FAA in a format that can be used in the FAA’s optimization. If the chosen format is too simple, it may not be capable of representing important aspects of the airlines’ objectives and constraints. If the chosen format represents too much detail, the airlines may be required to communicate to the FAA information of a combinatorial nature. The amount of such information may become unmanageable, as may the optimization problems faced by the FAA.

Many different compromises may be taken between the complexity of the airline information and the quality of the optimized allocation of resources determined by the FAA. Determination of suitable compromises requires an understanding of the objectives and constraints of the airlines, an understanding of the difficulty an airline would face in determining and communicating that information, and research into solution techniques of the optimization problems the FAA would need to solve.

7.4.2 Integration of ADCAM with Tactical Tools

Although the main emphasis of this thesis is strategic planning in the air transportation system during GDP’s, ADCAM could be integrated with tactical planning in the
system. An implementation of ADCAM would include interfaces between the airlines and the FAA for the exchange of data concerning each airport. These interfaces would allow the airline to specify how it wanted to allocate its arrival-departure capacity among its flights. They would also allow the FAA to update the arrival-departure capacity allocations when forecast weather conditions and forecast demand failed to materialize.

There is no reason that the same technique could not be used tactically as well as strategically. The arrival-departure subcapacity allocations would change to indicate operational inflexibilities arising as time progressed. For example, once a departure is queued on the taxiway connected to the departure runway at some airports, there is no way for that departure to allow departures of other airlines ahead of it in queue. Similarly, once an arrival is in the terminal airspace, there may be no safe way to delay it without delaying the arrivals behind it. These limitations would prevent an airline from trading arrival slots for departure slots at the last minute. Such limitations would be communicated to each airline through adjustments to its subcapacity constraints.

Before a departure enters the taxiway system and before an arrival enters terminal airspace, there may be significantly more freedom to change position in the arrival and departure queues, allowing tradeoffs of arrival capacity for departure capacity and vice-versa. ADCAM would provide an effective means of communicating the airline preferences to ATC.

Such communication is important to the tactical planning and operations of the FAA. Advance knowledge of airline departure times would allow the FAA to sequence departures better, coordinate arrival and departure traffic together and anticipate the need for configuration changes, each of which offers the potential to reduce delays. Advance knowledge of departures, for instance, would simplify the issues faced by a departure planner discussed in [20].

Similarly, the added flexibility could benefit an airline greatly. Consider, for example, a situation in which the airline’s gates will all be occupied at the time that a given arrival is due to arrive. Rather than allow that arrival to use valuable runway resources only to wait on the ramp for a free gate, the airline could delay that
arrival en-route and reduce the delay the airline's departures experienced by swapping the arrival slot for a departure slot. Subsequently, when a gate was expected to be available, the airline could swap a departure slot for an arrival slot. Similar considerations are motivating development of the Collaborative Arrival Planner tool of CTAS which allows airline-specified preferences to be taken into account in arrival sequencing. ADCAM offers an integrated means of accomplishing the same end for arrivals and departures simultaneously.

Another benefit of using ADCAM tactically as well as strategically is the improved situational awareness it would afford the airline. Northwest Airlines has recently reported that having an ARTS feed (which provides information about aircraft positions in the terminal area) allows them to allocate ground resources better because it provides advance information about the order in which aircraft will arrive. Similarly, American Airlines at DFW has a CTAS feed, which they find extremely helpful in gauging the likelihood that diversions will be necessary and in managing situations in which diversions are necessary [4].

Several efforts are under way to develop tactical or operational tools such as the Center TRACON Automation System (CTAS), the Surface Movement Advisor (SMA), the Surface Management System (SMS), Expedited Departure Planner (EDP), the Final Approach Spacing Tool (FAST), and the Departure Planner (DP). ADCAM would serve to integrate these presently uncoordinated tools in one consistent set of interfaces to the airlines. Not only would this offer benefit in terms of the tools' ease of use, but more importantly, it would allow these interrelated problems to be solved in a coordinated fashion. The current approach to development of these tools does not readily support such coordination. The result of an uncoordinated set of tools may well be operations that accomplish each tool's local objectives, but which could be improved significantly had they been coordinated.

Just as the airlines have discovered that explicit consideration of the interactions between the problems they solve improves their profits, a coordinated approach to the solution of arrival, departure, and surface operations promises to improve system operations. If an integrated solution to air traffic flow management, airport con-
figuration planning, runway operations sequencing, arrival planning and departure planning were implemented, the air transportation system could benefit enormously. The simulation developed in this thesis to test ADCAM could be used to test such approaches to integrating tactical and strategic planning.

7.4.3 Development of Airline Decision Aids

Because the system does not currently allow an airline control over the tradeoff between arrivals and departures allowed by ADCAM, the airlines have developed neither the decision-making processes nor the decision support tools required to make those decisions well. Development of such processes and tools is required before the full potential of ADCAM may be realized. For example, many decisions made at the airport, such as details pertaining to push-back of aircraft from their gates, would need to be coordinated with decisions made in the airlines' operating centers. Of course, ADCAM could be implemented, giving the airlines the flexibility to trade arrival and departure slots, before the airlines developed the means to capitalize on this flexibility. Indeed, practical experience in trading arrival capacity for departure capacity may be necessary to learn how best to use the capability.

7.4.4 Improvement of the Simulation

The results presented in Chapter 6 that have been generated using the simulation described in Chapter 5 are promising. However, they are also preliminary. Many aspects of the simulation could be improved. These improvements can be classified according to the part of the simulation they affect – the forecast model, the airport model, or the airline model. These improvements are outlined below.

The Forecast Model

Because the simulation does not currently model forecast uncertainty, the simulated airlines solve deterministic optimization problems, whereas real airlines do not. Furthermore, the model of the airport used in the optimization problems solved by the
airlines under ADCAM matches the simulation airport model perfectly.

Use of stochastic airport capacity forecasts would add to the realism of the simulation. Development of realistic stochastic capacity forecasts would entail some of the work outlined in Section 7.2.1. Realizations of airport capacity from the stochastic forecast could be used in a Monte Carlo simulation of airport operations to generate efficiently a distribution of value produced from the airline plans.

The Airport Model

The simulation models the airport as a single-server queueing system in which airline operations are served according to a strict first-planned, first-served policy. For instance, if one airline had planned the departure of its flight 101 from BOS at 10:00 and another airline planned the arrival of flight 102 to BOS at 10:10, the simulation would maintain the order of flight 101 ahead of flight 102.

This model probably represents operations at airports with strong arrival-departure interactions such as BOS, LGA, EWR and SFO quite well. It is probably not a very accurate representation of operations at airports like DFW or DEN which handle arrivals and departures almost completely separately. The variation in operating practice from airport to airport complicates the task of modeling the airport, and is an area in which significant research could be done.

The Airline Model

The airline model employed in this thesis could be enhanced in a number of ways. The addition of crew and equipment connection constraints in some form could add considerable realism to the simulation. Including these constraints may require significant additional research, since fleet planning and crew scheduling alone are significant research problems (see [3], [2] and [32], for instance).

The simulation model of the airlines would be strengthened through better modeling of network effects. The model used in this thesis approximates network effects solely through the use of delay multipliers as suggested in [5]. These multipliers make the delay of a flight early in the day more costly than the delay of a flight late in
the day. Such an approach does capture to first order the downstream cost of delay, however it does not capture any of the detail of the propagation of delay through an airline's network.

Finally, the models could be improved to allow planning under uncertainty. The optimization model currently representing airline behavior in the simulation solves a deterministic optimization problem. Development of stochastic optimization models to treat the problems faced by the airlines would enhance the realism of the simulation considerably.

7.4.5 Sliding Periods

The description of ADCAM in Chapter 4 relies on a division of time into a series of ten-minute intervals. For a number of reasons, it may be preferable to implement a version of ADCAM that does not discretize time. The discretization of time in practice would lead to some undesirable effects. For instance, the additional time required to operate an arrival or departure at an airport depends on the flights landing and departing at the airport at close to the same time. A model using discretized time periods, however, would miss the effect on a 9:59pm flight of the flights planned to operate after it. This section considers an alternate method of allocating airport arrival-departure capacity that does not rely on disjoint time periods.

Consider a sequence of operations on a runway. The additional time required to operate that sequence due to the insertion of an extra operation into the sequence depends primarily on those operations immediately preceding and following the inserted operation. If one knew the precise sequence of runway operations, one could calculate the amount of additional time required to serve each operation.

Of course, the air transportation system contains too much uncertainty to allow the accurate prediction of the sequence of runway operations hours in advance. Rather than calculate the amount of additional time required to serve an operation given a precise ordering of flights, it would be more appropriate to calculate the amount of additional time the operation would be expected to require, accounting for the reordering likely to occur. This expected additional time to operate a flight depends
on the relative numbers of different types of operations expected at about the same time as the flight in question.

The farther in advance of a flight’s expected time of operation one calculates expected operations mix, the more operations around its expected time of operation should be factored into the calculation. For example, five minutes in advance of a flight’s arrival at the runway, the operating mix is known quite well and need be calculated over only a small range of time surrounding the flight’s expected arrival time. Five hours in advance of arrival, however, the operating mix might better be taken as the relative numbers of arrivals and departures in a half-hour period around the flight’s expected arrival time.

In addition, the flights in the window of time over which the operating mix is calculated need not have equal weight in the calculation. It might make sense to give more weight to a flight expected to operate very near in time to the given flight than to others, representing the relatively higher probability of the two flights operating sequentially.

ADCAM rations airport resources according to an implicit assumption that the operations mix of the ten-minute period in which an operation occurs determines the time required for that operation. The amount of additional time required to serve the airline’s scheduled flights according to this implicit assumption determines the ration-by-schedule (RBS) allocation of resources to the airline. ADCAM is constructed so that the airline is allowed to rearrange its operations in any way that does not delay the flights of the other airlines beyond the delay that would result from operation of the RBS solution.

The calculation of operating mix in discrete ten-minute periods upon which ADCAM is based could be replaced by a calculation of the expected operating mix at the time each operation is expected. The other aspects of ADCAM would remain the same: the amount of additional time required to serve all the scheduled flights of an airline could be considered to be the airline’s RBS allocation, and any rearrangement of an airline’s operations that did not delay the flights of other airlines more than its RBS allocation did would be allowed. Complete development and implementation of
this idea is an area for further research.
Appendix A

Subcapacity Proof

**Proposition 3** Any operating point feasible to the overall airport capacity constraints can be achieved by a subcapacity-feasible solution, ignoring flight integrality, and any operation that is feasible to all subcapacity constraints also satisfies the overall airport capacity constraints.

*Proof:* The RBS solution contains $a$ arrivals and $d$ departures and lies on the boundary of the feasible region of airport arrival-departure capacity. This proof shows by construction that any point in that part of the feasible region of airport arrival-departure capacity bounded by the origin, the point $(a, d)$, and the set of points in $L$ is achievable by subcapacity-feasible points. The same construction can be used in the other half of the feasible region, defined by the origin, the point $(a, d)$, and the set of points in $R$. Because the airport feasible region is the union of these two regions, the fact that the property holds in both regions proves that it holds for the entire airport feasible region.

Consider the point $(a, d)$. The RBS solution, in which user $r$ operates at the point $(a_r, d_r)$, is feasible to the user subcapacities and meets the point $(a, d)$ exactly.

Consider breakpoint $b = (a + \alpha_b^L, d - \delta_b^L)$. If each user $r$ operates at its breakpoint $b_r = (a_r + \frac{a_r}{d} \alpha_b^L, d_r - \frac{d_r}{d} \delta_b^L)$, the result is that the airport operates at breakpoint $b$: 173
\[ \sum_r b_r = \sum_r (a_r + \frac{d_r}{d} \alpha_b^L, d_r - \frac{d_r}{d} \delta_b^L) \]  
(A.1)

\[ = \sum_r (a_r, d_r) + \sum_r \left( \frac{d_r}{d} \alpha_b^L, -\frac{d_r}{d} \delta_b^L \right) \]  
(A.2)

\[ = (a, d) + (\alpha_b^L, -\delta_b^L) \]  
(A.3)

\[ = (a + \alpha_b^L, d - \delta_b^L) \]  
(A.4)

The origin is a feasible point to every subcapacity and to the airport capacity constraints. Because the origin, \((a, d)\), and the breakpoints in \(L\) comprise the only extreme points of the feasible region, any extreme point of the airport feasible region may be achieved through a combination of points feasible to the user subcapacities. By convexity of all regions involved, any point feasible to the overall airport capacity constraints is the sum of points feasible to the user subcapacities.

Furthermore, the origin, the \((a_r, d_r)\), and the breakpoints in the \(L_r\) are the only extreme points of the user subcapacities. Because the user subcapacities are convex regions, every feasible point in the subcapacities is a convex combination of these extreme points. If the overall airport operating point \(p\) is the sum of subcapacity-feasible points, it may be written:

\[ p = \sum_r p_r \]  
(A.5)

\[ = \sum_r \left( \sum_b \lambda_{b,r} (a_r + \frac{d_r}{d} \alpha_b^L, d_r - \frac{d_r}{d} \delta_b^L) \right) \]  
(A.6)

\[ = \sum_b \left( \sum_r \lambda_{b,r} (a_r + \frac{d_r}{d} \alpha_b^L, d_r - \frac{d_r}{d} \delta_b^L) \right) \]  
(A.7)

\[ = \sum_b \lambda_b (a_r + \alpha_b^L, d_r - \delta_b^L) \]  
(A.8)

where \(0 \leq \sum_b \lambda_{b,r} \leq 1, \forall r\) and \(\lambda_b = \sum_r \lambda_{b,r} \frac{d_r}{d}, \forall b\). Step A.6 is possible from step A.5 by convexity of the user subcapacities. Because \(d = \sum_r d_r, 0 \leq \lambda_b \leq 1\), thus Equation A.8 shows that the combination of points feasible to the user subcapacities is a convex combination of the extreme points of the overall airport capacity, and by
convexity of the overall airport capacity it must therefore be feasible.
Bibliography


177


