Itinerary-Based Airline Fleet Assignment

by

Timothy S. Kniker

B.S.E. Department of Mechanical and Aerospace Engineering, Princeton University (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

June 1998

© 1998, Massachusetts Institute of Technology
All rights reserved

Signature of Author

Department of Electrical Engineering and Computer Science
May 12, 1998

Certified by

Cynthia Barnhart
Mitsui Career Development Associate Professor
Thesis Supervisor

Accepted by

Robert M. Freund
Seley Professor of Operations Research
Co-director, Operations Research Center
Itinerary-Based Airline Fleet Assignment

by

Timothy S. Kniker

Submitted to the Department of Electrical Engineering and Computer Science
on May 12, 1998, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

Abstract

The fleet assignment problem is to determine an optimal assignment of aircraft types to flight legs in a specified flight network. The assignment process must trade off the operating costs of large capacity aircraft with the lower operating costs, but potentially large amount of lost revenue, associated with small capacity aircraft. In this thesis, we investigate the relationship between this trade-off and the multi-leg nature of passenger itineraries.

We detail limitations of the current fleet assignment practices. In an attempt to alleviate these limitations, we present a new paradigm for fleet assignment that better estimates the quantity of revenue lost (called spill costs) when inadequate capacity is assigned to flight legs. Our paradigm also includes a procedure that allows a basis of comparison for different assignments.

We develop a deterministic linear program that routes passengers over a flight network which is the first to include the partial recapture of passengers. This model has a variety of uses in other airline decision processes including recovery operations, revenue management, and the fleet assignment problem.

To estimate spill costs, airlines typically pro-rate multi-leg passenger fares based on the flying miles of each flight leg in the itinerary. We present different spill models, coupled with various fare allocation schemes, that can yield better fleeting decisions. Using data from a major U.S. airline, we calculate that the annual savings from our approach are on the order of $30M.

We combine the fleet assignment model and the passenger mix model and show that our models result in near optimal fleeting. In fact, there are special conditions when the optimal solution of the conventional fleet assignment model is also the optimal solution for this combined model. Finally, we present evidence that suggests that capturing the network effects of multi-leg passenger itineraries is as important as capturing the stochastic nature of airline passenger demand in the fleet assignment model.

Thesis Supervisor: Cynthia Barnhart
Title: Mitsui Career Development Associate Professor
Contents

1 Introduction .................................................. 15
   1.1 Overview of the Thesis ................................. 16
   1.2 The Airline Scheduling Process ...................... 17
   1.3 Computational Information ............................ 19
   1.4 Conclusion ............................................ 20

2 The Fleet Assignment Model ................................. 23
   2.1 Overview .............................................. 23
   2.2 Model Description ..................................... 23
      2.2.1 An Illustrative Example ......................... 24
      2.2.2 The Set of Constraints ......................... 25
      2.2.3 Spill and Recapture ............................. 29
      2.2.4 The Objective Function ........................... 29
      2.2.5 The Attainable Contribution Problem .......... 32
      2.2.6 Basic Spill Model ............................... 34
   2.3 A New Paradigm for Fleet Assignment ................ 37
      2.3.1 An Iterative Methodology for Fleet Assignment 37
      2.3.2 Spill Calculation ................................ 37
      2.3.3 Problem Modification ............................ 39
      2.3.4 Fleet Assignment Component .................... 39
   2.4 A Mathematical Formulation of the Fleet Assignment Model 40
      2.4.1 Notation ........................................... 40
2.4.2 The Formulation ................................................. 41
2.4.3 The Solution Procedure ....................................... 41
2.5 Analysis of the Fleet Assignment Model ....................... 45
2.6 Extensions of the Fleet Assignment Model ..................... 48
  2.6.1 Other Formulations ......................................... 48
  2.6.2 Warm Start Fleet Assignment ............................... 51
  2.6.3 Applications ................................................ 52
2.7 Conclusion ..................................................... 53

3 The Passenger Mix Model with Partial Recapture ............... 54
  3.1 Overview ..................................................... 54
  3.2 Problem Description and Contributions ...................... 55
  3.3 Previous Research ........................................... 56
    3.3.1 The Stages of the Itinerary Selection Process .......... 56
    3.3.2 Deterministic Demand: Heuristics ...................... 57
    3.3.3 Deterministic Demand: Optimization .................... 58
    3.3.4 Stochastic Demand: Revenue Management ............... 59
    3.3.5 Stochastic Demand: Simulation and Heuristics .......... 59
    3.3.6 Stochastic Demand: Analytical Methods ............... 59
  3.4 The Mathematical Model and Formulation .................... 60
    3.4.1 Notation ................................................ 60
    3.4.2 The Basic Formulation ................................ 61
  3.5 The Keypath Formulation ..................................... 62
  3.6 Column and Constraint Generation ........................... 65
    3.6.1 An Overview of Column Generation ...................... 65
    3.6.2 The Solution Procedure ................................ 66
    3.6.3 Computational Experience with Implementation ........ 67
  3.7 Assumptions and Data Issues ................................ 70
    3.7.1 Airline Competition .................................... 70
    3.7.2 Congestion ............................................. 70
    3.7.3 Passenger Connectivity ................................ 72
3.7.4 Demand Data .................................................. 73
3.7.5 Recovery Rates ............................................. 77
3.7.6 Fare Data ................................................... 82
3.7.7 Parallel Markets ........................................... 82
3.8 The Passenger Mix Model with Stochasticity .................. 83
  3.8.1 Demand Distribution ..................................... 83
  3.8.2 A Formulation with Stochastic Demand ................. 84
3.9 Conclusion .................................................... 86

4 Applications of the Passenger Mix Model ....................... 87
  4.1 Overview .................................................... 87
  4.2 Irregular Operations ........................................ 87
    4.2.1 Implementation ...................................... 88
    4.2.2 Different Objective Functions ...................... 90
  4.3 Revenue Management ....................................... 90
    4.3.1 A Revenue Management Primer ....................... 90
    4.3.2 Leg-Independent Seat Inventory Control ........... 92
    4.3.3 Network Effects in Revenue Management ............ 93
    4.3.4 Itinerary Blocking .................................. 96
    4.3.5 Dual Variables as Bid Prices ....................... 96
    4.3.6 Bounds for Revenue Management .................... 96
  4.4 Fleet Assignment ........................................... 101
    4.4.1 Attainable Contribution Problem .................... 101
    4.4.2 A Combined Fleet Assignment Model ................. 102
  4.5 Conclusions ................................................ 102

5 Improving the Traditional Fleet Assignment Model .............. 104
  5.1 Overview .................................................... 104
  5.2 The Current State of the Practice ......................... 105
    5.2.1 A Sequential Framework ............................. 106
    5.2.2 The Current Methodology ............................. 107
5.2.3 Spill Estimation ........................................ 110
5.2.4 Mileage-Based Pro-Rated Fare Versus Full Fare Allocation .......... 113
5.3 Simple Improvements to the State of the Practice ....................... 115
  5.3.1 Equal Fare Allocation .................................. 116
  5.3.2 A Categorization of Passenger Itineraries ....................... 118
  5.3.3 Another Illustrative Example ................................ 121
  5.3.4 Capacity-Based Fare Allocation ................................ 122
  5.3.5 Probabilistic Allocation Schemes ................................ 125
  5.3.6 Spill Integration ........................................ 126
5.4 Analysis of Solution Techniques ...................................... 130
  5.4.1 Performance Trends ...................................... 130
  5.4.2 Performance as Demand Characteristics Change ...................... 131
5.5 The Iterative Methodology .......................................... 133
5.6 Cost Coefficient Modification Approach .................................. 135
  5.6.1 Spill Fare to Average Fare Ratio Modification ..................... 135
  5.6.2 Capacity-Based Fare Allocation Modification ....................... 137
  5.6.3 Probabilistic Allocation Scheme Modification ....................... 138
5.7 Multiple Flight Column Approach .................................... 139
  5.7.1 The Illustrative Example Revisited ................................ 139
  5.7.2 New Decision Variables and Assignment Combinations ............... 140
  5.7.3 Developing Partitions: A Set Partitioning Problem .................... 141
  5.7.4 The Multiple Flight Column Formulation ........................... 144
  5.7.5 Problem Modification Routine .................................. 147
5.8 Generalizations to Demand Uncertainty and Recapture ................... 147
  5.8.1 Fare Allocation and Recapture .................................. 147
  5.8.2 Flight Leg Classification .................................... 148
  5.8.3 Integrated Spill Model ....................................... 148
5.9 Conclusions .................................................. 148

6 The Combined Fleet Assignment and Passenger Mix Model .................. 151
  6.1 Overview .................................................. 151
6.2 Itinerary-Based Airline Fleet Assignment ........................................... 151
6.3 The Formulation ................................................................................. 153
  6.3.1 An Upper Bound ........................................................................ 153
  6.3.2 Solution Methods ....................................................................... 154
6.4 The Direct Solution Approach ............................................................ 155
  6.4.1 The LP Relaxation ....................................................................... 155
  6.4.2 Causes of Fractionality ................................................................. 156
  6.4.3 Achieving an Integer Solution ....................................................... 161
  6.4.4 Analysis of the Optimal Solution and Contribution ...................... 167
6.5 Lagrangean Methods ......................................................................... 168
6.6 Special Case Optimality .................................................................... 171
  6.6.1 One Capacitated Leg Passenger Itineraries ..................................... 172
  6.6.2 The Optimality of the Multiple Flight Column Approach ............. 175
6.7 The Relative Importance of Network Effects and Demand Uncertainty .... 177
  6.7.1 The Experiment Description ......................................................... 178
  6.7.2 Computational Experience .......................................................... 181
6.8 Conclusion ....................................................................................... 183

7 Conclusions ....................................................................................... 185
7.1 Contributions .................................................................................. 185
  7.1.1 The Passenger Mix Model ............................................................ 185
  7.1.2 The Basic Fleet Assignment Model .............................................. 186
  7.1.3 The Multiple Flight Column Approach ......................................... 187
  7.1.4 The Combined Model ................................................................ 187
7.2 Some Caveats .................................................................................. 188
  7.2.1 Itinerary Demand ....................................................................... 188
  7.2.2 Recapture .................................................................................. 188
  7.2.3 The Attainable Contribution Solver ............................................. 189
7.3 Areas of Future Research .................................................................. 189
  7.3.1 Further Validation ..................................................................... 189
  7.3.2 Extensions ................................................................................ 190
7.4 Final Thoughts ......................................................... 191

A Simulation Methods .................................................. 198
  A.1 Assumptions ......................................................... 198
  A.2 The Gamma Distribution .......................................... 199

B Glossary ................................................................. 201
  B.1 Definitions ......................................................... 201
  B.2 Notations .......................................................... 210
    B.2.1 Roman Symbols ............................................... 210
    B.2.2 Greek Symbols ............................................... 213
List of Figures

1-1 The airline scheduling process ........................................... 18
1-2 An example of a point-to-point airline network: El Al Airlines based in Israel . . 21
1-3 An example of a hub-and-spoke airline network: America West Airlines . . . . 22

2-1 The geographic network (a) and the time-space network (b) for the illustrative example ................................................................. 24
2-2 The timeline network associated with the illustrative example ......................... 26
2-3 Spill for a typical demand pattern .............................................. 34
2-4 A new fleet assignment paradigm .................................................. 38
2-5 A timeline segment (a) and node consolidation (b) ................................... 42
2-6 Arcs that would be removed if the airport is considered an island ................... 43
2-7 The network structure for the $MIN - Y$ formulation .............................. 47
2-8 Evenly-spaced flight copies within a flight leg's time window ......................... 50
2-9 A swap opportunity (taken from Berge and Hopperstad) ............................ 51

3-1 The airline passenger itinerary selection process .................................... 56
3-2 A typical market demand curve as a function of the time of day ....................... 57
3-3 The column and constraint generation solution procedure flowchart .................. 67
3-4 A discretized Gaussian unconstrained demand distribution .......................... 74
3-5 The idealized historical constrained demand function .................................. 75
3-6 The actual historical demand function for constrained demand ....................... 76
3-7 An expected marginal seat revenue curve .............................................. 84

5-1 The sequential framework to compare different fleet assignment methods .......... 107
5-2 The contribution as a function of the spill fare to average fare ratio 111
5-3 A comparison of the estimated spill and the actual spill when using the full fare allocation scheme 112
5-4 A comparison of the estimated spill and the actual spill when using the mileage-based fare allocation scheme 113
5-5 The quality of the contribution as a function of spill inaccuracy 114
5-6 A comparison of the equal fare and the full fare allocation scheme 117
5-7 Flight network for the illustrative example 121
5-8 The performance of various solution techniques as the demand factor is varied 132
5-9 The performance of various solution techniques while the passenger connectivity ratio is varied (keeping the demand factor constant) 133
5-10 An iterative methodology for the fleet assignment model 134

6-1 The optimal aircraft frontier 158
6-2 The direct solution approach for the combined fleet assignment and passenger mix model 162
6-3 The optimal aircraft frontier after coefficient reduction 164
6-4 The Lagrangean Primal-Dual Solution Technique for itinerary-based airline fleet assignment 170
6-5 The Lagrangean Primal-Dual method on Data Set A97-4A 171
1.1 The characteristics of the different problems ........................................ 20
2.1 Data for aircraft types in the illustrative example ................................. 25
2.2 Flight schedule information for the illustrative example ........................ 25
2.3 Demand data ....................................................................................... 30
2.4 Operating costs ................................................................................... 30
2.5 The four possible fleetinggs ............................................................... 30
2.6 The optimal spill costs and contributions for each fleeting ........................ 31
2.7 The contribution using a myopic greedy algorithm ............................... 31
2.8 The expected spill costs calculated by the different allocation schemes compared to the optimal solution .......................................................... 36
2.9 The total costs for the different allocation schemes ............................... 36
3.1 The effects of the solution procedure on the number of iterations, generated rows and columns, and run times ......................................................... 68
3.2 The number of rows and columns that are added compared with the total number of rows and columns for the Best Column approach .................... 69
3.3 The effects of congestion on system load factor, contribution, and run times ................................................................. 71
3.4 The effects of passenger connectivity on system load factors, contribution and run times ........................................................... 73
3.5 The effect of recapture on contribution .................................................. 79
3.6 The effect of recapture on the dual variables .......................................... 80
3.7 The sensitivity of the passenger mix model with random perturbations in the recovery rates .................................................. 81
3.8 The sensitivity of the passenger mix model with increased perturbations in the recovery rates ............................................. 81
3.9 The sensitivity of the passenger mix model with decreased perturbations in recovery rates ................................................ 82

4.1 The different restrictions for each fare class. The fare is for a one-way ticket from Boston(BOS) to Chicago(ORD) for March 16, 1998 ........................................ 94
4.2 Tightness of bound as a function of passenger connectivity ratio(network effects) and the demand factor (congestion) .............................. 100
4.3 Tightness of bound as a function of the fare class spread (segmentation) and the Z-factor (demand variance) ........................................... 100

5.1 A comparison of the full fare and partial fare allocation schemes ................................................................. 114
5.2 A comparison of the equal fare allocation scheme compared to the methods of full fare and cost-based pro-rated fare allocation ........................................ 117
5.3 A classification of flight legs based on the passenger demand and potential capacity119
5.4 A classification of passenger demand .............................................................. 120
5.5 Demand data .................................................................................. 122
5.6 Fleet type data ..................................................................... 123
5.7 A comparison of the capacity-based allocation schemes and the other fare allocation schemes .............................................. 124
5.8 A comparison of the probabilistic allocation schemes ......................... 126
5.9 The performance of spill integration compared to the representative approach ............................................. 129
5.10 The demand characteristics of the different data sets .................... 132

6.1 The number of integral flights and the LP relaxation solution time for traditional FAM model and the itinerary-based FAM .................................... 156
6.2 The operation cost and the number of seats for a specific flight leg .................. 157
6.3 The number of integral flights after the initial LP relaxation with and without coefficient reduction .................................................. 163
6.4 The number of integral flights after the initial LP relaxation with both pre-processing and post-processing cuts ................................................................. 166
6.5 The performance of the branch and bound procedure ......................................................... 166
6.6 A comparison of the contribution achieved with the state of the practice and the combined model ....................................................................................... 167
6.7 The values used for the different levels of the demand data characteristics .......................... 181
6.8 The mean data set for A97-3A ......................................................................................... 182
6.9 The mean and mean squared error for the relative performance of $EF$ and $IFAM^*$ 182
Acknowledgments

First and foremost, I would like to thank my family for all the support and love they have given me during this wonderful and frustrating time. Throughout my life, they have always been a source of inspiration and support, showing me the value of hard work, integrity, a good education and a sense of humor. What I am today, I owe it all to them.

Second, I wish to thank Professor Cynthia Barnhart for not only being a mentor, an advisor, and a colleague, but a friend. Her knowledge of the airline industry and optimization has made preparing this thesis much easier than I could have ever anticipated. Her guidance, encouragement, and editorial comments can never be repaid fully. I also would like to thank Professors Peter Belobaba and George Nemhauser for serving on my committee, and Professors Arnold Barnett, Robert Freund, Thomas Magnanti, Amedeo Odoni, James Orlin and the many others who have made M.I.T. such an exciting and challenging place.

Also, I would like to thank United Airlines and the Charles Stark Draper Laboratory for giving their support throughout my graduate career. Specifically, I would like to acknowledge the direct guidance of Ahmad Jarrah, Raj Sivakumar, John Goodstein, Ram Narasimhan, Nirup Krishnamurthy, and Stephen Kolitz.

An undeniable source of inspiration, motivation, advice, laughter and friendship was Jim Christodouleas. I do not think that I would have been at M.I.T. and made it through had it not been for his constant encouragement. Brian Kantsiper has been a wonderful friend, knowing exactly when I needed to take a break from work and reabsorb myself with the wonders of Central Square fast food, fungo, games, cinema, and ESPN Sportscenter commercials. I also wish to acknowledge Andy Armacost for sharing a common awe and admiration for the greatest game in the world and for constantly putting up with my nagging requests for computer help. Finally, I would like to give my thanks and best wishes to the WatPub crew (Amy, Keely, Rebecca, Leon, Joel, and Beril) for their camaraderie and support, the Princeton Sunday Brunch gang (Vickie, Steve, Jim, and Gregg), and the Galleria lunchtime triumvirate (Arni and Martin).
Chapter 1

Introduction

An airline is faced with a number of daunting problems in scheduling its aircraft to meet customer demands. Once an airline decides when and where to fly (i.e., develops a flight schedule), the next crucial decision is determining the type of aircraft that should be used on each of these flight legs. The formulation of this decision is referred to as the fleet assignment problem. We discuss this problem and specifically address a number of issues associated with the structure of passenger demands and its relationship to the fleet assignment problem.

In this thesis, we do not solve a problem that models the entire airline schedule. Because of the different structure of the flight network and operating requirements between the international and domestic markets at most U.S. carriers, the airline breaks up these operations into two different regimes. Domestic refers to a flight leg that takes off and lands in North America. Many times, an airline divides its fleet so that a group of aircraft in one subset may only be assigned to flight legs in the domestic network and the other subset of aircraft may only be assigned to flight legs in the international network. To balance the flow of aircraft, a purely domestic flight leg might be added to the international flight network. As is customary in the U.S. airline industry, we assume that the same domestic flight schedule is used Monday through Friday, despite vastly different demand patterns. For ease of operations, the assignment of aircraft types to flight legs is typically the same for every day of the business week, despite evidence that, in theory, a differing assignment could result in huge savings. The North American problem we investigate with a repeating schedule is defined as the daily domestic fleet assignment problem.
Before we begin, it will be helpful to define some terms to maintain consistency throughout the rest of the thesis. A flight leg is a non-stop trip of an aircraft from an origin airport to a destination airport (one take-off and one landing). A flight is a series of flight legs that has the same flight number. A market is defined as an ordered origin-destination airport pair, in which passengers wish to fly. Boston Logan International (BOS) - Chicago O'Hare (ORD), for example, is a distinct market from ORD-BOS. We model a round-trip itinerary as two distinct trips in two markets. A Boston-Chicago round-trip is represented as a passenger in the BOS-ORD market and a passenger in the ORD-BOS market. A market may have numerous itineraries which consist of a specific sequence of flight legs, sometimes differentiated only by schedule. The itinerary comprised of flight 789 from BOS to ORD and flight 276 from ORD to Los Angeles International (LAX) is distinct from the itinerary comprised of flight 792 from BOS to ORD and flight 275 from ORD to LAX, of which both are in the BOS-LAX market.

1.1 Overview of the Thesis

In Chapter 2, we discuss the specifics of the fleet assignment problem. Typically, the fleet assignment problem is formulated as some type of linear mixed integer program. Recent research has concentrated on incorporating difficult additional constraints into the problem or improving the solution speed. Despite these improvements, the current methods have shortcomings concerning the objective function. The objective function trades off the large operating costs of large aircraft with the potentially high amount of lost revenue associated with small capacity aircraft. While determining the operating cost of an aircraft is straightforward, the models used to estimate lost revenue due to capacity are, at times, poor approximations. We discuss the inaccuracies of the current models and present a new iterative paradigm for the fleet assignment process.

In Chapter 3, we present the passenger mix model, which is an element of the iterative paradigm. The passenger mix problem is formulated with a linear programming model that approximates the revenue that is attainable given a specific fleeting. Usually difficult, this model is solved quickly by applying a change of variables strategy that has been used to solve general multicommodity flow problems in the transportation and telecommunications industry.
We also show the impact on solution performance as we change the different characteristics of the demand data. In Chapter 4, we present several other applications of the passenger mix problem spanning from the strategic nature of fleet assignment to the tactical revenue management problem to the operational nature of dealing with irregular operations.

In Chapter 5, we develop a framework to compare different fleet assignments by using the passenger mix model in a post processing procedure to accurately estimate the spill associated with a specific fleet assignment. This framework is a valuable tool for comparing numerous methods that approximate the lost revenue from capacity restrictions in the fleet assignment model. Also, we present a new formulation of the fleet assignment model where a decision variable represents many assignments of aircraft types to flight legs not just one aircraft type to one flight leg. Finally, we complete the iterative paradigm by describing the information that we extract from each iteration and is then used to modify the problem for the next iteration.

In Chapter 6, we combine the traditional fleet assignment model with the passenger mix model to form an exact model for optimal fleet assignment, given a specific instance of demand. While difficult to solve for realistic problem sizes, we present ways to improve upon the speed. We use this new model to show how the new methods developed in Chapter 5 compare to the optimal solution. We also describe conditions for which the solution to the traditional fleet assignment model is also the optimal for the combined model. Finally, we conduct experiments to compare the relative importance of capturing network effects versus accurately incorporating demand stochasticity.

In Chapter 7, we summarize the main contributions of this thesis and describe future research directions.

1.2 The Airline Scheduling Process

The fleet assignment model is extensively investigated and reported in the literature, however it is but one piece of the larger puzzle of airline scheduling. Currently, the different scheduling decisions for an airline are determined sequentially, as shown in Figure 1-1. Schedule generation is the process of determining where the airline will fly, how often, and when. Currently, schedules are not created entirely from scratch; they evolve by gradual changes and adjustments to
previous schedules. Relatively little work by the operations research community has been done in this area, since good schedules require complex strategic decisions. The output of this process is a network of flight legs with scheduled arrival and departure times. The next step is fleet assignment in which an aircraft type is assigned to each flight leg in the schedule subject to the flow balance of aircraft. The next step, called aircraft maintenance routing, determines the rotations of specific planes. A rotation refers to the actual sequence of flight legs that a specific aircraft will follow. The primary goal of aircraft maintenance routing is to satisfy the maintenance requirements imposed by the Federal Aviation Administration (FAA). After the aircraft routings are determined, crew work schedules are generated such that costs are minimized.

![Diagram](image)

**Figure 1-1:** The airline scheduling process

Ideally, we desire some large formulation that could combine all of the models of Figure 1-1, therefore giving us a globally optimal schedule and plan. However, the ability to solve this large problem is not possible given current technology. There are attempts to integrate some of the models. For example, in the international problem, fleeting and aircraft routing is done simultaneously. Sometimes, if fleeting and routing are done sequentially, feasible maintenance routings are not possible for a given fleet assignment because the point-to-point nature of international flight networks (Figure 1-2) does not allow certain planes to arrive at maintenance stations on time. Since there is an abundance of maintenance stations in the United States and a large number of routing opportunities within hub-and-spoke domestic flight networks (Figure 1-3), a combined fleeting and routing model is not essential, and often results in insignificant (if any) savings.

Typically, the airlines begin the fleet assignment process 180 days prior to the schedule date. Human schedulers hand off the fleet assignment decisions between 90 and 120 days later.
During this time, schedulers adjust the fleet assignment solutions, since some constraints are
difficult to express mathematically and are not included in the model. Another reason for these
adjustments is the fleet assignment model’s inability to model accurately the revenues associated
with multi-leg passenger itineraries. One of the main goals of this thesis is to investigate ways
that the fleet assignment model can more accurately model multi-leg itinerary revenues to
produce schedules with increased profits.

1.3 Computational Information

Unless otherwise noted, computations are performed on a Hewlett Packard 9000 model D370
(160Mhz clock speed), using the HPUX operating system, version 10.20. The commercial
optimization solver is CPLEX, version 4.0.

Two different data sets are used, both from a major U.S. airline, to test the solution pro-
cedures. Sets F97 and A97 represent flight schedules and passenger demand estimates for
February 1997 and August 1997, respectively. With these data sets, we create a number of
smaller problems for the following reasons:

- It is desirable to see how certain techniques and methods work on both small and large
  problems. Also, by having numerous problems of different sizes, this allows us to determine
  the effect of problem size on solution techniques.

- With the current state of technology, certain solution procedures might work only for
  small problems. We might gain insights using smaller problems that help us determine
  improved solution methods for larger problems.

One piece of datum for each flight leg is the fleet that was eventually used by the airline.
We refer to the airline’s fleet assignment decisions as the current fleetting. To create smaller
problems, we determine a subset of aircraft types to be included, and extract only those flight
legs that were assigned to one member of the subset of fleets. This ensures that a feasible
solution (the current fleetting) exists. Table 1.1 contains the different characteristics of the
problems that we use in this thesis.
We always compare new results to a base case that is solved within the framework of the problem. We do not compare the new fleetings to the current fleeting provided by the airline for the following reasons:

1. For the sake of simplicity, we do not include all of the constraints that are contained in the airline’s operational fleet assignment solver.

2. The fleet assignments provided by the airline are the final result after adjustments by human planners. Therefore, changes are included that incorporate difficult constraints or better handle the network nature of the passenger flows. The magnitude of this human interaction and its effect on the solution is difficult to measure.

1.4 Conclusion

In this introduction, we show where the fleet assignment problem fits into an airline’s scheduling paradigm. We present some motivation for considering the multi-leg nature of passenger itineraries in the fleet assignment problem and describe the computing and data environment that we use throughout this thesis. In the next chapter, we further investigate reasons for rethinking the fleet assignment problem and its solution and we describe our alternative approach.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Number of Fleets</th>
<th>Number of Flight Legs</th>
<th>Number of Markets</th>
<th>Number of Itineraries</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>3</td>
<td>157</td>
<td>6,352</td>
<td>9,845</td>
</tr>
<tr>
<td>F97-4A</td>
<td>4</td>
<td>431</td>
<td>11,173</td>
<td>23,202</td>
</tr>
<tr>
<td>F97-6A</td>
<td>6</td>
<td>823</td>
<td>16,586</td>
<td>42,756</td>
</tr>
<tr>
<td>F97-9</td>
<td>9</td>
<td>2,044</td>
<td>20,928</td>
<td>76,741</td>
</tr>
<tr>
<td>A97-3A</td>
<td>3</td>
<td>173</td>
<td>7,034</td>
<td>11,877</td>
</tr>
<tr>
<td>A97-4A</td>
<td>4</td>
<td>485</td>
<td>12,656</td>
<td>29,299</td>
</tr>
<tr>
<td>A97-6A</td>
<td>6</td>
<td>877</td>
<td>16,736</td>
<td>47,626</td>
</tr>
<tr>
<td>A97-9</td>
<td>9</td>
<td>1,888</td>
<td>21,062</td>
<td>75,484</td>
</tr>
</tbody>
</table>

Table 1.1: The characteristics of the different problems

20
Figure 1-2: An example of a point-to-point airline network: El Al Airlines based in Israel
Figure 1-3: An example of a hub-and-spoke airline network: America West Airlines
Chapter 2

The Fleet Assignment Model

2.1 Overview

In Chapter 1, we discuss the scheduling process at most major carriers in the U.S. passenger airline industry. One component of this process is fleet assignment, a plan to assign aircraft types to each flight leg in a specified flight schedule. In this chapter, we describe the fleet assignment model and present a summary of solution procedures.

The current model uses an objective function to measure revenues that is only approximate. Typically, airlines do not validate this approximation with data. In this chapter, we question some of the underlying assumptions of the current model and present situations in which this model can lead to sub-optimal assignments of aircraft types. We describe a new paradigm for the fleet assignment model in which validating and modifying the model’s revenue approximations is integral to achieving better assignments.

2.2 Model Description

We examine the domestic daily fleet assignment problem. For operational ease, we restrict the flight schedule to be constant throughout the business week, i.e., the flight schedule and the assignment of aircraft types is identical every day. We note that some of the solution methods rely on a hub-and-spoke flight network, such as those prevalent in the U.S. airline industry.
Almost all fleet assignment models can be described as:

\[
\begin{align*}
\text{maximize} & \quad \text{profit from fleeting}, \\
\text{subject to:} & \quad \text{all flights flown by exactly one aircraft type,} \\
& \quad \text{aircraft flow balance,} \\
& \quad \text{only the number of available aircraft are used,} \\
& \quad \text{and other side constraints.}
\end{align*}
\]

A general description of the different components of the fleet assignment model follows. The main output of the model is the assignment of an aircraft type for every flight leg in the network. This output is referred to as either an assignment, a fleet assignment or a fleeting. We use these words interchangeably throughout the thesis.

### 2.2.1 An Illustrative Example

![Diagram](image)

Figure 2-1: The geographic network (a) and the time-space network (b) for the illustrative example

To facilitate the discussion of the fleet assignment model, we include a simple example of a fictitious airline. We assume that our airline has two aircraft, one of type A and one of type
B. Table 2.1 contains some specific characteristics of each aircraft type. The seating capacity is the number of seats that are available on the plane. The minimum turn time refers to the amount of time needed after an aircraft lands to perform all tasks required so the aircraft is ready to fly a subsequent flight leg. These tasks include deplaning passengers and baggage from the previous flight leg, refueling, cleaning the cabin, and boarding passengers and baggage for the next flight leg. In general, larger planes require a longer turn time.

<table>
<thead>
<tr>
<th>Aircraft Type</th>
<th>Seating Capacity</th>
<th>Minimum Turn Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100</td>
<td>30 minutes</td>
</tr>
<tr>
<td>B</td>
<td>200</td>
<td>45 minutes</td>
</tr>
</tbody>
</table>

Table 2.1: Data for aircraft types in the illustrative example

Table 2.2 contains the relevant flight schedule information. The origin and destination are the airports from which the aircraft take off and land, respectively. In Figure 2-1, we show both a geographic and time-space network representation of the flight network. The departure time refers to the scheduled time when the blocks are removed from the wheels of the aircraft prior to rolling back from the gate. Likewise, the arrival time refers to the scheduled time when the blocks are placed under the wheels just after the plane has pulled up to the gate. Typically, the phrase block time refers to the time difference between the removal of blocks for departure and the placement of blocks upon arrival.

<table>
<thead>
<tr>
<th>Flight Number</th>
<th>Origin</th>
<th>Destination</th>
<th>Departure Time</th>
<th>Arrival Time for Aircraft A</th>
<th>Arrival Time for Aircraft B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X</td>
<td>Y</td>
<td>8:00 A.M.</td>
<td>9:30 A.M.</td>
<td>9:20 A.M.</td>
</tr>
<tr>
<td>2</td>
<td>Y</td>
<td>Z</td>
<td>11:00 A.M.</td>
<td>12:00 P.M.</td>
<td>11:55 A.M.</td>
</tr>
<tr>
<td>3</td>
<td>Z</td>
<td>Y</td>
<td>1:00 P.M.</td>
<td>2:00 P.M.</td>
<td>1:55 P.M.</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>X</td>
<td>3:30 P.M.</td>
<td>5:00 P.M.</td>
<td>4:50 P.M.</td>
</tr>
</tbody>
</table>

Table 2.2: Flight schedule information for the illustrative example

2.2.2 The Set of Constraints

Cover Constraints

Traditionally called cover constraints, these set partitioning equations state that we want exactly one aircraft type to be assigned to a flight leg. We require that every flight leg in the schedule
must be flown, and that it is flown only once a day. Therefore, each flight leg has exactly one aircraft type assigned to it.

**Aircraft Flow Balance**

The *aircraft flow balance* constraints ensure the conservation of aircraft throughout the system. For each aircraft type, these constraints are a network flow node-arc incidence matrix based on a time-space network that describes the physical problem. This time-space network is called a *timeline network* in Hane et al. [25]. We present the basic characteristics here for completeness.

![Timeline Network Diagram](image)

**Figure 2-2:** The timeline network associated with the illustrative example

At 8:00 A.M., an aircraft departs from airport X. We define this 8:00 A.M. departure as a *departure event* in the timeline network. At 5:00 P.M., an aircraft arrives (an *arrival event*), yet this aircraft is not available for departure at 5:00 P.M. since the minimum turn time is required for preparing the aircraft for the next flight. To account for this required turn time, flight arcs are incident to the timeline at a node that represents the *ready time*, that is, the arrival time plus the minimum turn time. We define this event as an *availability event*. Only departure events and availability events are marked on the timeline (Figure 2-2). In our example, the network associated with aircraft type A has an arc intersect the timeline at 5:30 P.M. (an arrival at 5:00 P.M. and a 30 minute turn time), while the network associated with aircraft type B has an arc intersect the timeline at 5:35 P.M. (an arrival at 4:50 P.M. and a 45 minute turn time). The flow on these arcs represent an aircraft type assigned to this specific flight leg. Typically,
different aircraft types have different block times for the same flight leg, since some aircraft have faster airspeeds than others. This, along with different turn times, can result in different timeline network structures for each aircraft type.

Between two successive events on the timeline are arcs that represent aircraft staying on the ground. In the literature, these arcs are referred to as ground or sit arcs. Ground arcs are essential to maintain flow balance. While flow on ground arcs is restricted to be integer, we do not state this explicitly. Since all of the flight arcs are restricted to have exactly one aircraft type assigned, the ground arcs will have integer values. A special set of ground arcs, referred to as wrap-around or overnight arcs, are incident from the last event of the day and incident to the first event of the day. These arcs, which allow aircraft to sit on the ground overnight at airports, are important for two reasons, namely:

1. They ensure that a repeatable daily assignment is constructed.

2. They are used to model many side constraints.

**Fleet Count Constraints**

Aircraft purchasing decisions are made years in advance of delivery from the aircraft manufacturers. Within this time interval, traffic patterns might shift such that the current fleet makeup is no longer ideal. Nevertheless, the airline must create and fly a flight schedule that best utilizes their aircraft, while closely matching customer demand. To ensure that a fleet assignment does not use more aircraft of any type than the airline has available, we must count the number of aircraft at every point in time. If aircraft balance is maintained throughout the day, it suffices to do this count once a day. This is done by choosing a time, usually early in the morning (3 A.M. EST for instance), when there are few scheduled flight legs and almost all aircraft are on the ground. We count the number of planes of each type at this specific time and ensure that this number is less than or equal to the number available. In general, these constraints usually have strong effects on the geometry of the problem (Section 2.5).

Typically a major carrier has between one and two dozen different aircraft types. Some of these types can be consolidated into fleet types because of their similar seating capacities, operating costs, and airspeeds. In this paper, we refer to consolidated aircraft types as fleet
types. Computational experience suggests that the solution time increases dramatically as the number of fleet types increases. This consolidation should not be confused with the aircraft type grouping that might occur in other scheduling decisions, such as crew scheduling. In cockpit crew scheduling, the consolidation of aircraft types is based on cockpit design (not necessarily on seating capacity). An aircraft family is defined as a group of aircraft types where a crew is certified to fly all aircraft types in that family. Throughout the rest of the thesis, we use the term aircraft type to describe the physical problem, and fleet type to describe the mathematical abstraction of the physical problem where we have consolidated similar aircraft types for tractability reasons.

Additional Side Constraints

Typically, there are numerous side constraints, in addition to those mentioned above. The importance of these constraints is problem or airline specific. Some examples include:

- Noise restrictions: One example is that only a percentage of arrivals at an airport may be of certain aircraft types because of municipal laws that protect residents from noise pollution.

- Crew considerations: Usually pilots are only certified to fly aircraft in one aircraft family, and we want to ensure that there are enough flying hours for all pilots to meet quotas specified by labor agreements.

- Maintenance constraints: A specific mix of aircraft is required to be on the ground overnight at certain airports so that scheduled maintenance opportunities are built into the schedule.

The consolidation of aircraft types into fleet types is dependent on the nature of the additional side constraints. We do not want to consolidate a 108 seat Boeing 737-500 and a 109 seat Boeing 737-200 if there are numerous side constraints that limit the flow of both aircraft types in different ways.
2.2.3 Spill and Recapture

As mentioned before, a major airline has on the order of one or two dozen different aircraft types (and therefore seating capacities). This, along with the highly stochastic nature of demand, means that it is impossible to match capacity to the demand on every flight leg exactly. There are times when the demand for a flight leg exceeds the seating capacity of an aircraft, and the airline is unable to accommodate all travel requests. This process of turning away a passenger from his/her desired itinerary is called spill. If the passenger travels on the same airline, but on a different itinerary, that passenger is considered recaptured. A passenger is more likely to be recaptured if the airline can offer alternative itineraries that closely match the desired itinerary in departure time, arrival time, length of trip, and the number of stops.

2.2.4 The Objective Function

The objective function of the fleet assignment model is to find a profit-maximizing fleeting, given estimated passenger demand data. From basic economics, profit is revenues minus costs, yet, all costs are not considered in the fleet assignment model. Specifically, the fixed costs associated with operating and maintaining the infrastructure required for a major airline carrier are not considered. Among these costs are annual capital investments, salaried management, rent for gates and hangars. Instead, fleet assignment measures contribution, the net amount of income that the airline receives in providing air travel service. We divide the costs associated with a specific fleeting into two disjoint groups:

1. Operating costs: these are costs that can directly be assigned to a specific flight leg owing to the operation of the flight leg. These costs, such as the minimum amount of fuel, gate rental, and takeoff and landing costs, are independent of the number of passengers on board.

2. Carrying costs: these are the direct costs dependent on the number of passengers flown. These include, but are not limited to, the costs of extra fuel, baggage handling, reservation systems processing, and meals.

The revenue received can be expressed as potential revenue minus uncaptured or spilled revenue because of the fleeting decision. Given passenger demand, potential revenue is a constant.
The uncaptured revenue attributable to the fleeting decision is defined as *spill costs*.

To facilitate a discussion on the objective function, we return to our illustrative example. The deterministic demand data for this example is in Table 2.3 and the operating costs in Table 2.4. For this example, we assume that the carrying costs are negligible.

<table>
<thead>
<tr>
<th>Market (Sequence of Flights)</th>
<th>Itinerary</th>
<th>Number of Passengers</th>
<th>Average Fare</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-Y</td>
<td>1</td>
<td>75</td>
<td>$200</td>
</tr>
<tr>
<td>Y-Z</td>
<td>2</td>
<td>150</td>
<td>$225</td>
</tr>
<tr>
<td>X-Z</td>
<td>1-2</td>
<td>75</td>
<td>$300</td>
</tr>
</tbody>
</table>

**Table 2.3: Demand data**

<table>
<thead>
<tr>
<th>Fleet Type</th>
<th>Flight 1</th>
<th>Flight 2</th>
<th>Flight 3</th>
<th>Flight 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$5,000</td>
<td>$10,000</td>
<td>$10,000</td>
<td>$5,000</td>
</tr>
<tr>
<td>B</td>
<td>$10,000</td>
<td>$19,750</td>
<td>$19,750</td>
<td>$10,000</td>
</tr>
</tbody>
</table>

**Table 2.4: Operating costs**

Since we must have a balance of aircraft throughout the network, we must assign the same fleet type to flights 1 and 4. Likewise, we must assign the same fleet type to flights 2 and 3. Therefore, there are only four feasible fleeting (Table 2.5).

<table>
<thead>
<tr>
<th>Fleeting</th>
<th>Flight 1</th>
<th>Flight 2</th>
<th>Flight 3</th>
<th>Flight 4</th>
<th>Total Operating Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>$30,000</td>
</tr>
<tr>
<td>II</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
<td>$49,500</td>
</tr>
<tr>
<td>III</td>
<td>B</td>
<td>A</td>
<td>A</td>
<td>B</td>
<td>$40,000</td>
</tr>
<tr>
<td>IV</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>$59,500</td>
</tr>
</tbody>
</table>

**Table 2.5: The four possible fleeting**

The maximum potential revenue is $75($200) + 150($225) + 75($300) = $71,250. For the following analysis, we assume that the airline has full discretion in determining which passengers it wishes to accommodate. If we choose fleeting I, then each flight leg has a capacity of 100 seats. The demand for flights 1 and 2 is 150 and 225 passengers, respectively. Therefore, we must spill 50 of the passengers who desire travel on flight 1 and 125 passengers who desire travel on flight 2. Since the fare for the X-Z itinerary is less than the sum of the two local itineraries, we first spill 50 passengers on the X-Z itinerary ($15,000). The remaining demand for flight 1 no longer exceeds capacity. Since the local fare for flight 2 is less than the fare for the X-Z
itinerary, we spill 75 passengers from the Y-Z itinerary (\$16,875). Therefore the spill costs for fleeting I is $15,000 + $16,875 = $31,875. The spill costs for each fleeting are shown in Table 2.6.

<table>
<thead>
<tr>
<th>Fleeting</th>
<th>Operating Costs</th>
<th>Spilled Passengers</th>
<th>Spill Costs</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$30,000</td>
<td>50 X-Z, 75 Y-Z</td>
<td>$31,875</td>
<td>$9,375</td>
</tr>
<tr>
<td>II</td>
<td>$49,500</td>
<td>25 X-Z, 25 X-Y</td>
<td>$12,500</td>
<td>$9,250</td>
</tr>
<tr>
<td>III</td>
<td>$40,000</td>
<td>125 Y-Z</td>
<td>$28,125</td>
<td>$3,125</td>
</tr>
<tr>
<td>IV</td>
<td>$59,500</td>
<td>25 Y-Z</td>
<td>$5,625</td>
<td>$6,125</td>
</tr>
</tbody>
</table>

Table 2.6: The optimal spill costs and contributions for each fleeting

Therefore, fleeting I is the optimal fleeting with a contribution of $9,375. Consider the case when we solve the capacity problems for flights 1 and 2 by greedily spilling the smallest fare from each flight leg first. In this illustrative example, we always spill local passengers in favor of keeping the higher fare connecting passengers. For fleeting I, we spill 50 X-Y passengers at a fare of $200 and 125 Y-Z passengers at a fare of $225. The resulting spill costs and contribution for each fleeting are in Table 2.7.

<table>
<thead>
<tr>
<th>Fleeting</th>
<th>Operating Costs</th>
<th>Spilled Passengers</th>
<th>Spill Costs</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$30,000</td>
<td>50 X-Y, 125 Y-Z</td>
<td>$38,125</td>
<td>$3,125</td>
</tr>
<tr>
<td>II</td>
<td>$49,500</td>
<td>50 X-Y, 25 Y-Z</td>
<td>$15,625</td>
<td>$6,125</td>
</tr>
<tr>
<td>III</td>
<td>$40,000</td>
<td>125 Y-Z</td>
<td>$28,125</td>
<td>$3,125</td>
</tr>
<tr>
<td>IV</td>
<td>$59,500</td>
<td>25 Y-Z</td>
<td>$5,625</td>
<td>$6,125</td>
</tr>
</tbody>
</table>

Table 2.7: The contribution using a myopic greedy algorithm

If we use the greedy model, we are indifferent to either fleeting II or fleeting IV. The reason for the difference between this greedy heuristic and the network approach above is clear. By having a myopic view on both flight legs, we spill passengers that are locally optimal for each flight leg. On the other hand, by keeping the connecting passengers, we use up one unit of capacity on both flight legs. We can improve the solution by spilling the connecting passengers and accommodating the local passengers. For us to realize this, we need a network-wide view of the problem. We define this phenomenon as network effects.
2.2.5 The Attainable Contribution Problem

For Tables 2.6 and 2.7, we need to calculate the contribution (potential revenue minus spill costs minus carrying costs) attained for each specific fleeting. We define this as the attainable contribution problem. The problem description is: Given a flight schedule, the aircraft types that are assigned to each flight leg, and the passenger demand data, determine the contribution that can be attained. While this seems straightforward, there are numerous issues that increase complexity.

Network Effects

Previously, we demonstrate that accommodating or spilling a passenger and not considering how this affects other flight legs in the network can lead to sub-optimal decisions. This is one of the network effects. Another network effect involves recapture. It might be desirable to spill a passenger who is likely to fly on another itinerary on the same airline. Nevertheless, recapturing passengers is dependent on whether there is capacity on the alternate itinerary. Specifically, the two main network effects we wish to consider are the following:

1. If passengers are spilled on a specific flight leg, how does this affect the demand on other flight legs?

2. If spilled passengers can be recaptured, do we have the capacity on the other flight legs to accommodate them?

Stochasticity

In our illustrative example, we determine the optimal spill for one instance of demand. There are two reasons why this might not be sufficient.

1. The attainable contribution problem might have to be solved long before demand is realized. Therefore, we do not know what the actual demand will be.

2. It might be desirable to determine the average attainable contribution for numerous instances of demand, given a specific fleeting. Typically, there is a high level of uncertainty

32
in the demand for air travel service. This uncertainty comes from different day-of-the-week demand distributions, seasonal effects, and special events. Even after accounting for these effects, there is still a large underlying variance.

By considering just the network effects and the stochastic nature of demand, exactly calculating the expected attainable contribution is impossible for even small networks [48]. First, there are numerous random variables to consider. Second, there are numerous interactions in calculating the attainable contribution caused by the network effects. This results in the convolution of many random variables, a mathematically difficult chore.

**Fare Classes and Booking Order**

After the deregulation of the U.S. airline industry in the late 1970's, there is no longer a standard fare for travel from origin to destination. The airline's have segmented the market for travel based on a passenger's willingness to pay for the service. The different levels of segmentation are called *fare classes*. Usually, business travellers are willing to pay a higher price than leisure travellers. Realizing this, the airlines have set up restrictions, called *fences* to facilitate this segmentation. One of these fences is advanced booking, which causes low fare passengers to make their reservations before the high fare passengers. Combining this booking order with the stochasticity of demand leads to a situation where the airline must trade-off low fare passengers willing to pay now and unrealized high fare passengers in the future. The airline does not want to sell too many low fare seats early in the booking process in the event that there are too few available for high fare passengers later. On the other hand, at the time of departure, the airline does not wish to have empty seats on the flight leg caused by saving seats in the hopes of high fare passengers that do not materialize. The method by which the airline decides to accept or turn away the low fare passengers has an impact on the expected attained contribution.

**Other Phenomenon**

There are other phenomenon that affect the attainable contribution. Sometimes, passengers with reservations do not show up, either changing their itinerary or simply canceling their travel plans because of refundable tickets. In the airline industry, these passengers are referred to as *no-shows*. To counter this, many times airlines *overbook* a flight, i.e., sell more seats than are
actually available. Like most other research in this area, we ignore the effects of no-shows and overbooking.

Data Acquisition

As in most real-world situations, data is crucial to the problem. In many instances, problems abound with dirty data, in which human or machine error have resulted in inconsistencies. Even with clean data, some of the data that we desire is not available. For example, the rate at which customers are recaptured is not possible to determine exactly. For the rest of the thesis, we assume that, within reason, the data we desire is available.

Solving the Attainable Contribution Problem

Chapter 3 highlights the different methods presented in previous research and describes a new solution method for the attainable contribution problem. In the next section, we discuss how the attainable contribution problem is typically approximated in traditional fleet assignment models.

2.2.6 Basic Spill Model

![Spill Model Diagram]

Figure 2-3: Spill for a typical demand pattern

In traditional fleet assignment, the attainable contribution problem is approximated by
a simple spill model. This method can model the stochastic nature of demand, however, it ignores booking order and network effects. Usually an average fare is determined for the flight leg; this is easily obtained from historical data. This average fare calculation removes the distinctions between different fare classes and the resulting booking order. After this average fare calculation, the expected spilled costs, $E[c^p_i]$ are calculated for a specific flight leg $i$ by multiplying the average fare, $fare_i$, by the expected number of passengers spilled, $E[l_i]$, i.e.:

$$E[c^p_i] = fare_i \cdot E[l_i].$$  \hspace{1cm} (2.1)

To determine the expected number of spilled passengers, the probability mass function of demand must be known. Denote the demand for flight leg $i$ as a random variable, $Q_i$. We assume that $Q_i$ has a Gaussian distribution with mean $\overline{Q}_i$ and variance $\sigma^2_i$, shown in Figure 2-3. The assigned seating capacity for flight leg $i$ is denoted as $CAP_i$. The shaded region is the spill. The equation for determining the expected number of spilled passengers is [41]:

$$E[l_i] = \sum_{j=CAP_i}^{\infty} (j - CAP_i) P(Q_i = j).$$  \hspace{1cm} (2.2)

**Fare Allocation**

Calculating the average fare is not necessarily straightforward. The network effects of multi-leg passenger itineraries can cause problems when it comes to fare calculations, since a fare is for an entire itinerary, and it is not clear how to divide the fare among the different flight legs in a multi-leg itinerary. Returning to our illustrative example, we examine the itinerary for the passenger market between airport X and airport Z. The fare associated with this itinerary is $300. An intuitive approach is to divide the fare evenly between the two legs. This is equivalent to two independent sets of 75 passengers, one set flying on flight 1 for $150 each, and the other set flying on flight 2 for $150 each. The resulting demand and average fare calculations are 150 passengers paying $175 and 225 passengers paying $200 for flights 1 and 2, respectively. Another approach is to assign the entire fare to all legs of the itinerary. This method results in 150 passengers paying $250 and 225 passengers paying $250 for flights 1 and 2, respectively.
Using these two different allocation schemes, we calculate different estimates of spilled revenue, neither of which is correct (Table 2.8).

<table>
<thead>
<tr>
<th>Fleeting</th>
<th>Equal Allocation (Approximation)</th>
<th>Full Allocation (Approximation)</th>
<th>Actual Spill</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$33,750</td>
<td>$43,750</td>
<td>$31,875</td>
</tr>
<tr>
<td>II</td>
<td>$13,750</td>
<td>$18,750</td>
<td>$12,500</td>
</tr>
<tr>
<td>III</td>
<td>$25,000</td>
<td>$31,250</td>
<td>$28,125</td>
</tr>
<tr>
<td>IV</td>
<td>$5,000</td>
<td>$6,250</td>
<td>$5,625</td>
</tr>
</tbody>
</table>

Table 2.8: The expected spill costs calculated by the different allocation schemes compared to the optimal solution

In Table 2.9, we show the total costs (operating plus spill costs) for each solution. Notice that if we use the equal allocation scheme, we choose the solution that uses fleet type A for flights 1 and 4 and fleet type B for flights 2 and 3. If we use the full allocation scheme, we choose the solution that uses fleet type B on both flights, even though the optimal solution is to use fleet type A on all flights.

<table>
<thead>
<tr>
<th>Fleeting</th>
<th>Equal Allocation (Approximation)</th>
<th>Full Allocation (Approximation)</th>
<th>Optimal Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>$63,750</td>
<td>$73,750</td>
<td>$61,875</td>
</tr>
<tr>
<td>II</td>
<td>$63,250</td>
<td>$68,250</td>
<td>$62,000</td>
</tr>
<tr>
<td>III</td>
<td>$65,000</td>
<td>$71,250</td>
<td>$68,125</td>
</tr>
<tr>
<td>IV</td>
<td>$64,500</td>
<td>$65,750</td>
<td>$65,125</td>
</tr>
</tbody>
</table>

Table 2.9: The total costs for the different allocation schemes

Another issue is determining which passengers are actually spilled. The actual makeup of spilled passengers can have a significant effect on the bottom line in airline revenue because of the enormous price differentials for the various fare classes (e.g., round-trip coach cabin fares from BOS-LAX range from $449 to $3,518 based on restrictions). This is the main idea behind revenue management (or yield management). Revenue management is discussed at length in Section 4.3.3. If the revenue management system works well, it is a safe assumption that the average fare of the set of spilled passengers from a flight is less than the average fare of all passengers wishing a seat on that flight. To model this marginal fare, we multiply the average fare by some multiplier to take into account the effects of revenue management systems. Beloaba and Farkas [10] have done some work on modeling revenue management effects in spill
cost estimations, beyond determining a simple universal multiplier.

2.3 A New Paradigm for Fleet Assignment

The simple allocation rules can do a poor job of measuring spill costs (Table 2.8). While there is no guarantee that poor spill approximations result in poor fleeting decisions, it stands to reason that if we develop strong confidence in our spill calculations, we will have greater confidence in the resulting fleeting decisions. This section describes a new iterative paradigm for fleet assignment that achieves better estimates of spill. We present an overview of the paradigm in general terms, then spend the rest of the thesis detailing the specific components. We also discuss the circumstances under which this paradigm yields good fleeting decisions.

2.3.1 An Iterative Methodology for Fleet Assignment

Figure 2-4 contains a schematic for the new iterative methodology. We begin with initial estimates of spill, then use the fleet assignment solver to determine a fleeting based on these initial estimates. Given this fleeting, we determine the spill and carrying costs with a spill calculation module. We take the output of this model and the operating costs from the fleet assignment model to determine the attainable contribution of this specific fleeting. Based on some pre-specified termination criteria, we either modify the problem to achieve another fleeting or stop. If we modify the problem, we use the fleet assignment solver to determine a new fleeting, and the process starts over again.

As with any iterative methodology, numerous questions are raised. First, is there some improvement guarantee from iteration to iteration? Second, is the process guaranteed to converge to the optimal solution? Third, if the process does converge to the optimal solution, at what rate does it converge?

2.3.2 Spill Calculation

In this module, we wish to determine the spilled revenue given a specific fleeting, while considering as many of the phenomenon as described in Section 2.2.4. We consider two alternative approaches, passenger flow models and simulation.
Figure 2-4: A new fleet assignment paradigm

Passenger Flow Models

With passenger flow models, we determine the expected spill using either large-scale linear programming techniques or heuristics. The advantage of this approach is its ability to determine quickly spill costs using effective solution techniques in the optimization literature. The disadvantage of this approach is it is a relatively inflexible modeling environment. For example, only rough approximations can be used to capture the stochastic nature of demand.
Simulation

Another alternative is to simulate the actual booking process of passengers and airlines. The more accurate and detailed the model, the more accurate the resulting spill calculations. Obviously, the main advantage of using a simulation procedure is the greater flexibility in modeling some of the phenomenon that can occur. The main disadvantage is that the amount of time necessary to determine good solutions often is excessive since many instances of the simulation must be run in order to achieve statistically significant results because of the highly stochastic nature of demand.

2.3.3 Problem Modification

After we have calculated the spill and carrying costs for a given fleeting, we might modify our fleet assignment model in an attempt to improve profits. We highlight two different methods to do this. One is to modify the cost coefficients of the objective function, the other is to change the structure of the model.

Cost Recalculation

Given a fleeting, we can determine which flight legs are causing the spill to occur. We might be able to modify the objective function to model more accurately the spill costs associated with flight legs that are the main sources of spilled passengers.

Altered Problem Structure

We might be able to identify small subsets of flight legs for which it is difficult to capture network effects using linear spill costs. In these situations, we modify the structure of the fleet assignment problem by adding columns that exactly model the spill phenomenon for this subset of flight legs and leaving the rest of the model unchanged.

2.3.4 Fleet Assignment Component

One advantage of our iterative methodology for fleet assignment (Figure 2-4) is that it can use existing fleet assignment models that have been developed, tested, and implemented by the
airlines over the last decade. This allows for a faster and smoother implementation in industry. In the following sections, we present an overview of the fleet assignment model.

2.4 A Mathematical Formulation of the Fleet Assignment Model

We present a mathematical formulation of the fleet assignment model, which is closely based on that found in Hane et al. [25].

2.4.1 Notation

Sets

$A$: the set of airports indexed by $o$.

$L$: the set of flight legs in the flight schedule indexed by $i$.

$K$: the set of different fleet types indexed by $k$.

$T$: the sorted set of all event (departure or availability) times at all airports, indexed by $t_j$. The event at time $t_j$ occurs before the event at time $t_{j+1}$. Also, $|T| = m$; therefore $t_1$ is the time associated with the first event after the count time and $t_m$ is the time associated with the last event before the count time.

$N$: the set of nodes in the timeline network indexed by $\{k, o, t_j\}$

$C(k)$: the set of flight legs that cross the countline when flown by fleet type $k$.

$I(k, o, t)$: the set of inbound flight legs to node $\{k, o, t_j\}$.

$O(k, o, t)$: the set of outbound flight legs from node $\{k, o, t_j\}$.

Decision Variables

$X_{k,i} = \begin{cases} 1 & \text{if flight leg } i \in N \text{ is assigned to fleet type } k \in K; \\ 0 & \text{otherwise.} \end{cases}$

$Y_{k,o,t_j}$: the number of fleet type $k \in K$ aircraft that are on the ground at airport $o \in A$ immediately after time $t_j \in T$. 

40
\( Y_{k,o,t_j} \): the number of fleet type \( k \) aircraft that are on the ground at airport \( o \in A \) immediately before time \( t_j \in T \). If \( t_1 \) and \( t_2 \) are the times associated with adjacent events, then
\[
Y_{k,o,t_1^+} = Y_{k,o,t_2^-}.
\]

\( Z(\mathbf{X}) \): the contribution function associated with the fleeting decision \( \mathbf{X} \).

**Parameters/Data**

\( n_k \): the number of aircraft in fleet type \( k \), \( \forall k \in K \).

### 2.4.2 The Formulation

\[ (FAM) \]

maximize \( Z(\mathbf{X}) \)

subject to:
\[
\sum_k X_{k,i} = 1 \quad \forall i \in L
\]

\[
\sum_{i \in C(k)} X_{k,i} + \sum_{o \in A} Y_{k,o,t_m^+} \leq n_k, \quad \forall k \in K
\] \hspace{1cm} (2.3)

\[
\sum_{i \in I(k,o,t)} X_{k,i} + Y_{k,o,t_j^-} - \sum_{i \in O(k,o,t)} X_{k,i} - Y_{k,o,t_j^+} = 0, \quad \forall \{k,o,t_j\} \in N
\]

\( X_{k,i} \in \{0,1\}, \ \forall i \in L \) and \( k \in K \)

\( Y_{k,o,t_j^+}, Y_{k,o,t_j^-} \geq 0 \) and integer, \( \forall \{k,o,t_j\} \in N \)

### 2.4.3 The Solution Procedure

In Section 2.5, we show that assuming a linear cost function, this mixed integer programming (MIP) formulation of fleet assignment is NP-hard. Nevertheless, by exploiting the structure of the problem, it can be solved efficiently. A problem size that represents a major airline's entire domestic problem, (11 fleets, 2,600+ flights) is consistently solved in under an hour, by Hane, et al. [25].
The full solution procedure is described in Hane, et al. [25], but we mention some of the highlights relevant to this research. The problem is solved as a linear MIP using a commercial optimization solver. The solution process involves a couple of stages of consolidation to decrease the problem size. After the linear programming relaxation is solved, feasible integer solutions are found using a tailored branch and bound algorithm.

**Solving the LP Relaxation**

One of the first steps is to aggregate data. For example, in the Hane formulation, there are sets of flight legs defined as required through flights. Required through flights are a set of flight legs that an airline's marketing department has mandated must be flown by the same aircraft. Obviously, these two flight legs must be assigned the same fleet type. Mathematically, we enforce this restriction for a pair of flight legs $i$ and $j$ by adding the following set of constraints to the problem:

$$X_{k,i} - X_{k,j} = 0, \quad \forall k \in K.$$  \hspace{1cm} (2.4)

Computationally, we combine these two variables into one variable, to decrease the number of variables in the problem.

![Figure 2-5: A timeline segment (a) and node consolidation (b)](image)

Another method of aggregation is node consolidation. We examine the section of timeline displayed in Figure 2-5(a). The links incident to the timeline represent arriving flights and the links incident from the timeline represent departing flights. For the arriving flight arcs, the node where the intersection occurs is the time when the aircraft assigned to the flight leg is
available for its next departure. Therefore, we can make this available time later as long as there is no departure flight leg. In Figure 2-5(b), we show the resulting consolidation; the size of the network decreases by 4 nodes and 4 arcs. In this way, we maintain all possible connections, do not include any new infeasible connections, and conserve aircraft balance while decreasing the size of the problem. Empirically, this can cause dramatic decreases in problem size and run times.

![Diagram](image)

**Figure 2-6:** Arcs that would be removed if the airport is considered an island

The problem size can be decreased further by exploiting the hub and spoke nature of the flight schedule. At spoke airports, there might be only a few flight legs each day, typically following a pattern of alternating arrivals and departures (Figure 2-6). If a departure flight leg is not assigned to the same fleet type as the immediately preceding arrival flight leg, then an extra aircraft is required to remain on the ground overnight at this spoke airport. Typically, this is not optimal because it results in low aircraft utilization. Since optimal solutions usually do not have aircraft overnight at these airports, we can eliminate the possibility of overnighting an aircraft by equating the variables of the arriving and departing flights. This allows us to eliminate a number of ground arcs as well. The X's in Figure 2-6 show the ground arcs that we assume are not used.

Hane et al. [25] tried numerous solution methods offered by a commercial optimization solver to improve run time for the LP relaxation. The interior point method consistently outperformed the primal solution method. The dual simplex method seemed to be on par with the interior point method and many times achieved faster solution times. Steepest edge was
the best pricing policy when using the dual simplex algorithm.

Achieving an Integer Solution

After solving the LP relaxation, an integer solution must be constructed since it is impossible to implement fractional solutions. Empirical work suggests:

1. The bound of the linear programming relaxation relative to the optimal integral solution is typically tight, resulting in integrality gaps much less than 1%.

2. The resulting solution usually has a large number of variables that are integer or close to integer. On the order of 80-90% of the flight legs are assigned solely to one fleet type in the linear programming relaxation of the problem for realistically sized problems.

Using these two properties as rationale, Hane et al. fix variables that are close to one in the solution of the LP relaxation solution to one. This results in a problem that is much smaller, and much quicker to solve when using branch and bound.

Branch and bound is an enumeration method that relies on getting good bounds quickly to prune the search tree early in the process [13]. The effectiveness of branch and bound is a function of problem structure. One effective method used for the fleet assignment problem [25] employs branching not on a single variable, but on groupings of variables that form special ordered sets (SOS). The cover constraints, which are of the form

$$\sum_k X_{k,i} = 1,$$  \hspace{1cm} (2.5)

are referred to as Type 3 SOS constraints. A Type 3 SOS constraint is a set of binary variables, in which exactly one must be set to one and all others to zero. We can develop a more effective branching strategy by dividing the set of variables into two sets, where the sum of the variables in the first set equals one or the sum of variables in the second set equals one. This can lead to a near equal partitioning of the feasible solutions in the branching tree and, in practice, has resulted in some improvements in solution time.

Another element of SOS branching is determining on which set to branch first. In the fleet assignment model, this requires determining on which flight leg to branch first. Hane et al. used
different measures of variance of the objective function cost coefficients of a particular flight leg to prioritize the flight legs. The assignment of a flight leg with a high variance in its objective coefficients might have significant impact on the objective value and allow more pruning parts of the search tree.

2.5 Analysis of the Fleet Assignment Model

Gu et al. [24] study the complexity and behavior of the fleet assignment formulation in Program 2.3. Some of their key results are:

- The problem of determining a feasible fleet assignment with three or more fleets is NP-complete.

- The fleet assignment problem with two fleets but no size constraints is equivalent to a one commodity network flow problem and thus the polyhedron associated with the feasible region is integral. With three fleets, the problem becomes NP-hard.

- An expression for the minimum number of planes for feasibility in the fleet assignment problem with one fleet can be found.

- A lower bound to the minimum number of planes in the K-fleet assignment is equal to \( (\frac{3}{8} \log_2 K) \) times the minimum number of planes in the one fleet problem.

Also, Gu et al. [24] show with a simple example that if there are three fleet types and no fleet size constraints, then the polyhedron is not necessarily integral. Nevertheless, computational experience suggests that even for large problems, the polyhedron has integral extreme points in the area of optimal solutions for typical cost functions. This leads to the conclusion that the set of fleet size constraints is most likely the main source of fractionality in the fleet assignment model. The additional constraints, such as crew considerations and noise restrictions, might cause even more fractionality as they cut the polyhedron.

Divide a feasible solution, \( \mathbf{X} \), of \( \text{FAM} \) into two components: the fleeting variables, and the ground variables, i.e., \( \mathbf{X} = \{\mathbf{x}, \mathbf{y}\} \). We define \( \{\mathbf{x}, \mathbf{y}(\mathbf{x})\} \) as a minimal fleeting solution, if and only if \( \mathbf{y}(\mathbf{x}) \) is the optimal solution to the following program:
\[ (MIN - Y) \]

minimize \[ \sum_{k \in K} \sum_{o \in O} Y_{k,o,t^+} \]

subject to:

\[ Y_{k,o,t^-} - Y_{k,o,t^+} = \sum_{i \in \delta^+(k,o,t)} X_{k,i} - \sum_{i \in \delta^-(k,o,t)} X_{k,i} \quad \forall \{k,o,t\} \in N \]  

\[ Y_{k,o,t^+}, Y_{k,o,t^-} \geq 0 \text{ for all } \{k,o,t\} \in N \tag{2.6} \]

Proposition 1 Given a fleeting \( \bar{\mathbf{x}} \), the \( MIN - Y \) formulation has a unique optimal solution.

Proof: The \( MIN - Y \) formulation is a network flow problem. For each \( \{k,o\} \in (K,A) \) we have a circle of directed arcs (see Figure 2.7). In the figure,

\[ d_j = \sum_{i \in \delta^+(k,o,t_j)} X_{k,i} - \sum_{i \in \delta^-(k,o,t_j)} X_{k,i}. \]  

\[ (2.7) \]

To prove this proposition, we construct the optimal solution, \( \bar{\mathbf{y}} \). We begin by setting \( Y_{k,o,t^+_n} = 0 \). To satisfy the first set of constraints of 2.6, we set

\[ -Y_{k,o,t^+_1} = d_1 - Y_{k,o,t^+_n} \]  

\[ (2.8) \]

Likewise,

\[ -Y_{k,o,t^+_j} = d_j - Y_{k,o,t^+_j-1}, \quad \text{for } j = 2, \ldots, n. \]  

\[ (2.9) \]

Since we have a feasible fleeting, \( \sum_{j=1}^{n} d_j = 0 \). Let \( y = \min_{j=1, \ldots, n} \{Y_{k,o,t^+_j}\} \). To satisfy the second set of constraints of 2.6 for \( j = 1, \ldots, n \)

\[ \bar{Y}_{k,o,t^+_j} = -y + Y_{k,o,t^+_j} \]  

\[ (2.10) \]

This solution \( \bar{\mathbf{Y}} \) is feasible for \( MIN - Y \) and it is the unique optimal solution.\( \square \)

We use Proposition 1 to prove the following theorem.
Theorem 2 A minimal fleeting solution \( \{x, y(x)\} \) where all elements are integer is a vertex of the polyhedron defined by the LP relaxation of FAM.

Proof: We use a contribution function, \( Z(X) \), that is linear with respect to the decision variables. For the fleeting variables that are unity, set the cost coefficient to 0, and for all other fleeting variables set their coefficients to some large number \( M \gg |L| \). All overnight arcs should have their corresponding cost coefficient set to unity, while the cost coefficients for all other ground arcs are set to zero. The optimal objective value of the solution to Program 2.3, is equal to the number of planes on the ground during the count time. Any other solution that uses a different fleeting has a much greater value for this cost vector, and by Proposition 1 of \( \bar{y}(x) \), we have the unique minimal value of the overnight arcs. Therefore we have constructed a cost function, \( c \), for which the minimal fleeting solution is the unique optimal
solution, therefore, \( \{x, y(x)\} \) is a vertex of the LP relaxation polyhedron. □

For the traditional fleet assignment, this theorem suggests that any possible fleeting of interest to us, that is, an integral fleeting using the minimum number of aircraft, is a vertex of the FAM LP-relaxed polyhedron. Therefore, if we have the proper cost function, the optimal solution to the LP relaxation is integral and branch and bound is unnecessary.

2.6 Extensions of the Fleet Assignment Model

In Section 2.4, we present a common formulation of the fleet assignment model. This formulation uses only the direct operating costs and the expected spill costs for determining the best fleeting for a fixed schedule of flights. In this section, we present some other ways to formulate the problem and other models that integrate some of the downstream decisions such as maintenance routing and crew pairing.

2.6.1 Other Formulations

The leg-based formulation of the fleet assignment model is used by a number of airlines, nevertheless, this isn’t the only way to formulate the problem. In fact, for some airlines with additional constraints or particular network structure, it is not the preferred model.

Maintenance Restrictions

Barnhart et al. [4] describe a fleet assignment model that captures aircraft maintenance constraints exactly. The result is a combined fleet assignment and aircraft maintenance routing model. In their formulation, the decision variable is no longer associated with a fleet type flight leg pair, but instead is associated with a fleet type-string pair. A string is a sequence of connected flight legs with the first flight departing and the last flight arriving at maintenance stations. They use this formulation in point-to-point flight networks. Compared to hub-and-spoke networks, point-to-point networks have fewer opportunities to route aircraft, given a particular fleeting. The result is that fleetings generated by Program 2.3 are not guaranteed to allow maintenance feasible aircraft routings. To ensure feasibility, exact maintenance constraints must be added to the model. Hoffman [26] proposes a similar approach for the daily
domestic problem.

**Connection Networks**

The fleet assignment model as formulated in Program 2.3 does not maintain the specific identity of each aircraft, instead all aircraft of a specific fleet type are considered homogeneous. This assumption, however, can be limiting. For example, many airlines have shuttle flight legs which allow quick boarding and deplaning of passengers. If an airline flies successive shuttle flight legs, then the minimum turn time for the aircraft assigned to these flight legs is reduced, say to only 15 minutes instead of 30 minutes. This cannot be modeled by letting an availability event occur 15 minutes after the shuttle flight leg arrival since this would allow an aircraft assigned to a shuttle flight leg to connect with a non-shuttle flight leg in less than 30 minutes. To model conditions such as these, it is necessary to track the connections of aircraft, and the timeline network cannot do this.

Rushmeier and Kontogiorgis [35] propose a fleet assignment model that uses both timeline and connection networks. A connection network differs from a timeline network in that it replaces all ground arcs with connection arcs. A connection arc exists from an availability event of a flight leg to a departure event of another flight leg if the same aircraft can be assigned to both flight legs feasibly. Abara [1] formulated the fleet assignment problem with the decision variables representing the assignment to turns instead of flight legs. The decision variable was $X_{ijk}$ where:

$$X_{ijk} = \begin{cases} 
1 & \text{if flight leg } i \text{ turns to flight leg } j \text{ with fleet type } k, \\
0 & \text{otherwise.} 
\end{cases}$$

(2.11)

**Time Windows**

Rexing et al. [34] consider the situation where the arrival and departure times are allowed to move within a user specified time window. The idea is to solve simultaneously the fleet assignment and departure time scheduling problems. The modeling concept is to make copies of each flight leg spaced at intervals within the time window (Figure 2-8) to represent alternate departure times for each flight leg. The decision is then to select the departure time and fleet
Figure 2-8: Evenly-spaced flight copies within a flight leg's time window

assignment for each flight leg. Therefore, the decision variable is:

\[
X_{k,i,m} = \begin{cases} 
1 & \text{if flight leg } i \text{ is assigned to fleet type } k \text{ for flight departure time } m. \\
0 & \text{otherwise.}
\end{cases}
\]  

(2.12)

Although this formulation can become prohibitively large, Rexing et al. present an iterative process that requires copies of only certain flight legs, thereby reducing memory requirements. While fleet assignment savings of their approach are estimated to be $67,000/day, these savings were achieved by rescheduling less than 6% of the flight legs. Also, this revised flight schedule requires 2 fewer aircraft to fly than the original schedule.

**Maintenance and Crew Considerations**

Clarke et al. [14] adds constraints and variables to the traditional fleet assignment problem in order to capture maintenance and crew issues. Their results show that these additional constraints resulted in dramatically increased solutions times compared to those of the basic fleet assignment model. Solution times increased to 2 - 5 hours from the typical 15-20 minutes required for the basic fleet assignment model.

Lu [32] combines the fleet assignment model with an approximate to the crew pairing model to form a large mixed integer linear program. The goal of this work is to achieve fleet assignments that allowed reduced total fleeting and crew costs.
2.6.2 Warm Start Fleet Assignment

We refer to the fleet assignment problems for which initial feasible assignments do not exist as cold-start or zero-based fleet assignment. Warm-start fleet assignment occurs when an initial feasible fleet assignment is available.

Berge and Hopperstad [12] present a method, called Demand Driven Dispatch or \( D^3 \), to swap assigned aircraft types to those within the same aircraft family on certain flight legs. This dynamic assignment of aircraft generates more cost effective fleet assignments as the date of departure draws closer and the forecast for the final demand on a flight leg becomes more accurate. Beginning with a feasible fleet assignment, two paths of flight legs flown by different aircraft types are identified and if there is improvement, their assignments are switched. For example, in Figure 2-9, the letters represent flight legs and the numbers represent fleet types. There is a swap opportunity by assigning \( i \) and \( j \) to fleet type 2 and \( m \) and \( n \) to fleet type 1. Berge and Hopperstad report improvements in operating profits of 1-5%.

![Diagram of a swap opportunity](image)

Figure 2-9: A swap opportunity (taken from Berge and Hopperstad)

Talluri [42] improved on Berge and Hopperstad's work in the following ways:

- Running time is reduced both empirically and theoretically;

- The algorithm explores more possibilities since it is not restricted to consider only swaps that occur in the same day; and
• The algorithm is optimization-based and is therefore guaranteed to find better same-day swaps, since $D^3$ is a heuristic that is not guaranteed to find all profitable swaps.

2.6.3 Applications

Minor modifications can be made to the basic fleet assignment model to help answer important scheduling questions other than fleeting.

Fleet Planning

The fleet assignment model can be used for fleet planning [1]. Fleet planning is the strategic decision when an airline determines the aircraft it wishes to purchase and either sell or scrap. If we ignore the fleet count constraints, and assign a cost on all arcs that cross the timeline equal to the daily cost of having an aircraft of that specific type in the fleet, one can determine an optimal fleet for a given flight schedule. This can be used to direct an airline’s fleet purchasing decisions in the future.

Schedule Construction

Another application is to make adjustments to the flight schedule. Rexing et al. do this by adjusting departure and arrival times. Another possibility is to change the covering constraints from equality to less-than inequalities:

$$\sum_k X_{k,t} = 1 \quad \Rightarrow \quad \sum_k X_{k,t} \leq 1 \quad (2.13)$$

By including a large number of candidate flights, the optimal solution defines the most profitable schedule. However, this modeling approach has restrictive assumptions. Namely, it is assumed that passenger demands on different flight legs are independent. In reality, however, the demand for a 9:05 A.M. flight leg from Boston to Chicago is a function of whether or not there is a 9:00 A.M. flight leg from Boston to Chicago. In spite of the independence assumption, this model has been used to aid in schedule construction [1]. Daskin and Panayotopoulos [18] assign aircraft types with numerous predefined candidate routes. The output of the model is which routes should be included in the schedule and which aircraft should be assigned to each included
route.

Both US Airways [44] and Klincewicz and Rosenwein [29] investigate the fleeting and scheduling of the weekend flight network. Using the change in Equation 2.13, a schedule contains only those flight legs that are in the solution to the modified fleet assignment model. Constraints are added to ensure balance between the weekend and the regular weekday assignment [44]. For example, an output of the weekday fleet assignment is the number of aircraft of each type on the ground at each airport. These numbers must be the same as the number on the ground both Saturday morning and Sunday night.

2.7 Conclusion

In this chapter we provide a description of the fleet assignment problem. We show the limitations of the model as it is used in practice, specifically in the representation of leg-based spill costs. Even in a small example, we show that simple spill calculation methods give sub-optimal solutions. We present an iterative paradigm that might be employed to improve the solutions that are obtained by conventional methods.
Chapter 3

The Passenger Mix Model with Partial Recapture

3.1 Overview

In Chapter 2, we present an iterative paradigm for the fleet assignment problem. One of the components of this paradigm is a module that accurately calculates the attainable contribution by determining the resulting spilled revenue given a flight schedule and a specific fleetling. There are numerous reasons why an exact calculation of the spilled revenue is difficult; however, there are many ways that we can approximate it.

In this chapter, we present a linear programming formulation of the assignment of passengers to flight legs in an airline network. Current attempts to calculate expected revenue either use revenue management, which has not efficiently been solved for a large network of flights with multiple-leg itineraries, or mathematical programming formulations, which up to this point have not accurately incorporated the concept of recapture. Our model is solved quickly using a column generation approach where many of the constraints can be relaxed and only a nominal number are expressed explicitly.

While our primary goal is to develop a model that approximates spilled revenue for use in a solution procedure for itinerary-based airline fleet assignment, the passenger mix model presented here has applications in many other areas of airline decision making, such as revenue management and irregular operations. These applications are investigated further in Chapter
3.2 Problem Description and Contributions

Given the entire set of unconstrained customer demands for air travel service over a predetermined fixed schedule of flights and a fixed fleet assignment, the airline's objective is to maximize their revenues by accommodating as many high fare passengers as possible. It is not possible to match exactly supply to unconstrained demand because of the requirements of the fleeting decision, such as fleet makeup, the corresponding fixed seating capacity, and the flow balance of aircraft at each airport. For some flights, unconstrained demand exceeds supply and passengers must be spilled to other itineraries of either the same or another airline. In this situation, the best policy is to spill passengers who either have a low fare or a high probability of being recaptured on another itinerary offered by the airline.

Various difficulties arise in determining network effects, booking order, and uncertainty of unconstrained demand. Previous research, as well as the research presented here, addresses only some of these concepts, otherwise intractability results. There appears to be no consensus as to which issue is the most important to model. In fact, the specific application dictates the relative importance of each issue and similarly the selected modeling approach determines which issue is difficult to model. For example, revenue management techniques usually model booking order and stochasticity well, but the numerous interactions between random variables, requiring computationally intractable convolution integrals, make this an impractical approach for solving full network-size systems. Mathematical programming approaches handle these interactions well, but result in approximations to the non-linear nature of random variable probability density functions.

The contributions of Chapters 3 are as follows:

1. We create a linear programming formulation of the attainable contribution problem that can be solved in a reasonable amount of time. Given fleeting decisions, the output of this model is the assignment of passengers to the flight legs in the flight schedule, detailing which passengers are spilled and which alternative itineraries are utilized. We also discuss how this model can be used for other strategic and tactical planning exercises.
2. We demonstrate a novel application of a change of variable strategy that can be used for the general minimum cost multi-commodity flow problem. We also show that this strategy combined with column and constraint generation results in an effective solution procedure for other problems.

3.3 Previous Research

3.3.1 The Stages of the Itinerary Selection Process

We decompose the passenger itinerary selection process into three stages (Figure 3-1). First, for each market, unconstrained passenger demand is modeled as a function of the time of day (Figure 3-2). Typically there is a peak around 9 A.M. in the morning and another in the late afternoon and early evening. This input is used together with information about trip frequency, trip reliability, travel times and perceived safety [36], to solve the airline choice problem, that is to determine the market share of each competing airline. After passengers are allocated to the different airlines, the itinerary choice problem is solved to determine the passengers’ desired itineraries. Typically, these models use an attractiveness factor for each itinerary based on time of day, trip time and number of stops to estimate unconstrained itinerary demand. Given this demand for each itinerary, the airlines accommodate passenger requests. As high demand flight legs fill, passengers change their request if one of the flight legs in their itinerary is full (the itinerary is blocked). We define this process as passenger flow. In reality, the itinerary selection process is not sequential, however, we model it this way to classify the different solution methods.

![Figure 3-1: The airline passenger itinerary selection process](image)

We consider only models that include the passenger flow decision process. Some have
combined both itinerary choice and passenger flow as detailed in the following sections. We are aware of only one model that has included all three components [39]. We begin our research immediately after itinerary choice is complete.

There have been numerous approaches to model the phenomenon of passenger routing requests and airline responses to them. The approaches differ in their solution procedures and in their modeling assumptions. We present a taxonomy of these different approaches in the next section. The first level of classification is whether the model assumes deterministic or stochastic demand. The next level of classification involves the underlying solution approach.

3.3.2 Deterministic Demand: Heuristics

Gagnon [22], along with Air Canada, uses a heuristic method that incrementally assigns passengers to flights. Phillips et al. [33] expand upon this method to include different booking patterns for different itineraries and recapture. Soumis et al. [37] also follow a sequential approach where the model determines the preferred passenger flows, then uses these flows to develop a flight schedule and fleeting. In other words, there are no capacity restrictions for the passenger flow and the goal is to create a flight schedule and fleet assignment that accommodates all passengers along their desired routes.
3.3.3 Deterministic Demand: Optimization

Soumis et al. [38] use a mathematical programming model, called MAPUM, to assign passengers to itineraries (the itinerary choice problem). It considers the passenger flow process by assigning dissatisfaction costs for unattractive itineraries and passenger overload (spill) costs for loading good itineraries. The resulting model is linear in its constraints and non-linear in its objective function. This approach, unlike other work in this area, focuses on user-optimization. That is, the approach models passenger choices based on the assumption that each passenger is trying to optimize their situation given the choices of the airline. In contrast, system-optimization models determine passenger choice decisions that are optimal for the airline.

Glover et al. [23] propose a minimum cost network flow model with side constraints to determine which passenger itineraries to spill. They refer to it as the passenger mix problem. In their time-space network representation, a node represents an airport at a specific point in time. A forward arc represents a flight leg in the flight schedule. The flow capacity on this arc represents the seating capacity based on the aircraft type assigned to the flight leg. A series of backward arcs represents different passenger itineraries and the capacities on these arcs are the unconstrained demands for the different passenger itineraries. The cost on these arcs is equal to the negative of the fare for that passenger itinerary and fare class option. Another network approach for the same problem is proposed by Dror et al. [19].

Phillips et al. [33] propose a packing formulation that is the mathematical formulation of Glover's network model. In his thesis, Farkas [21] proposes a packing formulation similar to Phillips et al. For a reasonably sized network (600 flights, 30,000 itineraries, and five fare classes), this model is solved efficiently using Glover's network approach. These formulations have the limiting assumption that a passenger that does not travel on his/her desired itinerary is lost to the system. We refer to this modeling assumption as no recapture. Another modeling assumption is perfect recapture, in which a passenger always accepts travel on any itinerary in his/her desired market. Somewhere between no recapture and perfect recapture is reality, which we define as partial recapture. With partial recapture, only a percentage of passengers will accept travel on an alternative itinerary, and that percentage varies by desired and alternative itinerary. The spilled passengers that are not recaptured travel on another airline or don't
travel. This recapture model is investigated in section 3.7.5.

3.3.4 Stochastic Demand: Revenue Management

Another approach to the attainable contribution problem is to use seat inventory control models from revenue management. Belobaba [7] gives an overview of previous revenue management research and presents an effective heuristic for setting booking limits given multiple fare-classes on a single isolated flight leg [8]. Curry [16] determines mathematically optimal booking limits in a single flight-leg case. A limitation of these models is their inability to handle network effects.

3.3.5 Stochastic Demand: Simulation and Heuristics

Farkas [21] uses a Monte Carlo revenue management simulation approach. Unconstrained demand is generated for each itinerary. From this demand, a booking request is picked at random and it is determined if the seat inventory control procedure accepts the request. This procedure is repeated until all requests are processed. This overall process is performed a number of times to calculate summary statistics. Researchers at Boeing propose a sequential probabilistic approach dubbed PFLOW [27] to determine which flight leg is the most likely to reach its capacity first. When a flight leg reaches its capacity, it becomes a choke leg. The model then spills passengers from the remaining itineraries that use this leg to other itineraries with no choke legs. After this spill occurs, the next likely flight leg to spill is determined and so on. This process continues until all flight leg demands do not violate capacity restrictions.

3.3.6 Stochastic Demand: Analytical Methods

Some researchers look to the “classical” urban transportation equilibrium models to gain insights about passenger flows. Soumis and Nagurney [39] develop a model that concentrates on user optimization instead of system optimization. This model applies research developed in urban traffic networks to determine passengers’ choice of airlines based on the attractiveness of the path and the capacity available. One limitation of the model is that it assumes that passenger spill is independent of the specific markets. Specifically, the ratio of spilled passengers from market $m$ to all spilled passengers is equal to the ratio of demand from market $m$ to total
demand. Walker [45] presents current efforts at Boeing to expand upon equilibrium work. This model balances both the user optimization objective of Soumis and Nagurney and the system optimization objective of other research.

3.4 The Mathematical Model and Formulation

To solve the attainable contribution problem, we propose a model that is a generalization of Glover's passenger mix model. In our linear programming formulation we consider the partial recapture of spilled passengers. In the case when the unconstrained demand is expressed as the average number of passengers per day, the optimal solution is an allocation of passengers for the instance when the actual demand is exactly equal to the average demand. We detail some interesting properties of this solution in Section 4.3.6. For shorthand, we refer to an itinerary fare class option as IFC. The notation for this will be \((p, f)\) denoting itinerary \(p\) under fare class \(f\).

3.4.1 Notation

Sets

- \(L\): the set of all flight legs, indexed by \(i\).
- \(J\): the set of cabins (e.g., first-class, coach), indexed by \(j\).
- \(P\): the set of all itineraries, indexed by \(p, r,\) or \(q\). Included in this set is the null itinerary. Spilling to the null itinerary represents the airline not offering an alternative itinerary.
- \(M\): the set of all markets, indexed by \(m\).
- \(F\): the set of all fare classes, indexed by \(f, g,\) or \(h\).
- \(F_j\): the subset of all fare classes that use cabin \(j\).

Decision Variable

- \(x_{p,f}^{r,g}\): the number of passengers (possibly fractional) that desire to travel on \((p, f)\), but instead travel on \((r, g)\).
Parameters and Data

\( fare_{p,f} \): the average fare for \((p,f)\).

\( b_{p,f}^{r,g} \): the recovery rate of a passenger desiring \((p,f)\) to \((r,g)\).

\( CAP_{i,j} \): the capacity of cabin \(j\) on flight leg \(i\).

\( D_{p,f} \): the daily unconstrained demand for \((p,f)\).

\( \delta_{i}^{p} \): 
\[
\begin{cases} 
1 & \text{if flight leg } i \text{ is on itinerary } p \\
0 & \text{otherwise}
\end{cases}
\]

3.4.2 The Basic Formulation

Using the above notation, we formulate the problem as follows:

\[
(\text{PMIX})
\]

maximize \[ \sum_{(r,g) \in (P,F)} fare_{r,g} \sum_{(p,f) \in (P,F)} x_{p,f}^{r,g} \]

subject to:

(capacity) \[ \sum_{(r,g) \in (P,F)} \sum_{(p,f) \in (P,F)} \delta_{i}^{p} x_{p,f}^{r,g} \leq CAP_{i,j} \quad \forall (i,j) \in (L,J) \]

(limit) \[ \sum_{(r,g) \in (P,F)} x_{p,f}^{r,g} / b_{p,f}^{r,g} \leq D_{p,f} \quad \forall (p,f) \in (P,F) \]
\[ x_{p,f}^{r,g} \geq 0. \]

In the objective function, we are maximizing the amount of revenue attained. This is calculated by multiplying the average fare for \((r,g)\) and the number of passengers who use this IFC (including those redirected from all other itineraries). The capacity constraints ensure that the number of passengers on each flight leg \(i\) in cabin \(j\) is less than the number of seats available in the cabin. The limit constraints guarantee that the number of people who desire \((p,f)\) traveling on all itineraries is less than the unconstrained demand, \(D_{p,f}\).

To investigate further, assume there are 2.0 passengers traveling on \((r,g)\) who desire \((p,f)\). If the recovery rate \(b_{p,f}^{r,g} = 0.25\), then we must spill 8.0 passengers who desire \((p,f)\) in order to recapture 2.0 passengers on \((r,g)\). We assume that \(b_{p,f}^{r,g} = 1.0\), that is, if a passenger is offered
his/her desired IFC, he/she accepts with probability 1.0. The slack in the limit constraints represents the number of passengers who are spilled to other airlines or who don’t travel. The PMIX formulation is similar to those presented by Farkas [21] and Phillips et al. [33], but relaxes their no recapture assumption.

Using the A97-9 data set from United Airlines with only one fare class for each itinerary, there are 75,484 different itineraries in 21,062 domestic markets in our model. This results in a formulation with 489,517 columns and 77,372 rows. This problem is too large to solve efficiently in a short amount of time, therefore a new formulation is in order. We build our new model to exploit the fact that many of the decision variables are zero in an optimal solution and further, that most of the variables that take on non-zero values in the final solution are of the form $x_{p,f}^{p',f}$, that is, most passengers are assigned to their desired IFC.

3.5 The Keypath Formulation

The passenger mix model is a generalized minimum cost multi-commodity flow (MMCF) problem. When using arc-based formulations for MMCF problems, the number of node-balance constraints is equal to the number of nodes times the number of commodities, which is problematic when either the number of nodes or commodities is large. Typically, path-based formulations of MMCF problems have the advantage of a reasonable number of constraints because node balance is implied by the path nature of the decision variables. Nevertheless, these formulations still have a large number of constraints when the number of commodities is large. To decrease the number of constraints that have to be stated explicitly, we employ a change of variable strategy using keypaths, first proposed by Barnhart et al. [3]. For the passenger mix model, we define a “commodity” as the passengers desiring travel on $(p,f)$ and the keypath as the set of flight legs in the itinerary $p$. If many of the flight legs are not highly capacitated, then most passengers fly on their desired itinerary (the keypath) and we consider explicitly only those passengers that fly on alternative itineraries because of capacity restrictions. This leads to a change of variables in the passenger mix model, specifically:

$t_{p,f}^{r,g}$: the number of passengers who desire travel on $(p,f)$, but whom the airline attempts to redirect onto $(r,g)$.
Note the difference between this decision variable and \( b_{p,f}^{r,g}t_{p,f}^{r,g} \), the actual number of passengers that travel on \((r,g)\) who wanted \((p,f)\).

The unconstrained demand for cabin \(j\) on flight leg \(i\), denoted by \(Q_{i,j}\), can be calculated by:

\[
Q_{i,j} = \sum_{f \in F_j} \delta^p_{i}D_{p,f}, \quad \forall(i,j) \in (L,J).
\] (3.2)

This unconstrained demand might be greater than the capacity, \(CAP_{i,j}\). We wish to find the optimal set of \(t_{p,f}^{r,g}\) to subtract from \(Q_{i,j}\) so that the capacity constraints are satisfied.

The new keypath formulation for the passenger-mix problem is:

\[
(PMIX - KP)
\]

\[
\text{minimize: } \sum_{(r,g) \in (P,F)} \sum_{(p,f) \in (P,F) \setminus (r,g)} (\text{fare}_{p,f} - b_{p,f}^{r,g} t_{p,f}^{r,g})
\]

subject to:

\[
\sum_{(p,f) \in (P,F_j)} \sum_{(r,g) \in (P,F) \setminus (p,f)} \delta^p_{i}t_{p,f}^{r,g}
\]

(capacity)

\[
- \sum_{(p,f) \in (P,F_j)} \sum_{(q,h) \in (P,F) \setminus (p,f)} \delta^p_{i}q,h t_{q,h}^{p,f}
\]

\[
\geq Q_{i,j} - CAP_{i,j} \quad \forall(i,j) \in (L,J)
\] (3.3)

(limit)

\[
\sum_{(r,g) \in (P,F)} t_{p,f}^{r,g} \leq D_{p,f} \quad \forall(p,f) \in (P,F)
\]

\[
t_{p,f}^{r,g} \geq 0.
\]

The \(PMIX - KP\) formulation is obtained by substituting out of the \(PMIX\) formulation the \(x_{p,f}^{r,g}\) variables as follows:

\[
x_{p,f}^{p,f} = D_{p,f} - \sum_{(r,g) \in (P,F) \setminus (p,f)} t_{p,f}^{r,g} \quad \forall(p,f) \in (P,F)
\]

\[
x_{p,f}^{r,g} = b_{p,f}^{r,g} t_{p,f}^{r,g} \quad \forall(r,g) \in (P,F) \setminus (p,f).
\] (3.4)

Both formulations have the same number of constraints and roughly the same number of variables. But by comparing the limit constraints in the two formulations, we see the value of
using keypaths. In the PMIX formulation, we want columns to be as positive as possible since this improves the objective function value. Therefore, the limit constraints in the PMIX formulation are usually binding. In the PMIX - KP formulation, since most of the objective coefficients are positive, we wish to have the decision variables be zero whenever possible, because the more positive a variable, the worse the objective value. As a result, we can relax many of the limit constraints. This leads us to a solution method in which we explicitly express only the constraints when they are violated.

Separating the objective function into two distinct parts yields

\[
\text{maximize} \sum_{(r,g) \in (P,F)} \text{fare}_{r,g} x_{r,g} + \sum_{(r,g) \in (P,F)} \text{fare}_{r,g} \sum_{(r,g) \in (P,F) \setminus (p,f)} x_{r,g}^{r,g}. \tag{3.5}
\]

Substituting Equation 3.4 into Equation 3.5 results in

\[
\text{maximize} \sum_{(r,g) \in (P,F)} \text{fare}_{r,g} \left( D_{r,g} - \sum_{(r,g) \in (P,F) \setminus (r,g)} t_{r,g}^{r,f} \right) + \sum_{(r,g) \in (P,F)} \text{fare}_{r,g} \sum_{(r,g) \in (P,F) \setminus (p,f)} b_{p,f}^{r,g} x_{p,f} \tag{3.6}
\]

Combining like terms, we have

\[
\text{maximize} \sum_{(r,g) \in (P,F)} \text{fare}_{r,g} D_{r,g} - \sum_{(r,g) \in (P,F)} \sum_{(p,f) \in (P,F) \setminus (r,g)} (\text{fare}_{p,f} - b_{p,f}^{r,g} \text{fare}_{r,g}) t_{p,f}^{r,g} \tag{3.7}
\]

The term \( \sum_{(r,g) \in (P,F)} \text{fare}_{r,g} D_{r,g} \) is a constant equal to the amount of revenue that can be attained if there are no capacity restrictions, i.e., the amount of revenue that is attained if every passenger is accommodated on their desired itinerary. We define this value as the unconstrained revenue because this is the maximum amount of expected revenue that the airline can attain if every term \( (\text{fare}_{p,f} - b_{p,f}^{r,g} \text{fare}_{r,g}) \) is positive. Since \( \sum_{(r,g) \in (P,F)} \text{fare}_{r,g} D_{r,g} \) is a constant, we can remove it from the maximization. Multiplying the remaining terms by \(-1\) and minimizing, we have the objective function in PMIX - KP.

If \( (\text{fare}_{p,f} - b_{p,f}^{r,g} \text{fare}_{r,g}) \) is negative, then even with no capacity restrictions, it is better to redirect the passenger from the \((p, f)\) to \((r, g)\) since the fare is higher and the recovery rate...
is such that we expect more revenue by redirecting. Empirically, we observe that this term is usually positive.

3.6 Column and Constraint Generation

3.6.1 An Overview of Column Generation

We still have a large number of variables, specifically, \( \sum_{m \in M} N_m^2 \) variables, where \( N_m \) is the number of itineraries in market \( m \). Most of these variables are zero, thus suggesting a column generation approach. In column generation, we begin with a restricted master problem, in which only some of the variables are actually considered. The restricted master problem solution provides information to the pricing subproblems that are solved to determine if any of the variables omitted from the restricted master problem might improve upon its current solution. If one or more are found, they are added to the restricted master problem and the process is repeated. If none are found, the program stops since an optimal solution has been found. A summary of column generation is in Ahuja et al. [2].

The cost of a path in \( PMIX-KP \) is dependent on the path itself and cannot be allocated to the arcs in the path, as in many multi-commodity flow problems. Therefore, to solve the pricing subproblem, we compute explicitly the reduced cost of all possible columns in order to prove optimality. Empirically, the number of columns is reasonable, so this enumeration does not severely slow down the solution procedure. Section 3.6.3 discusses issues with column pricing.

We start the solution technique by solving an initial restricted master problem. Ideally, we desire the initial restricted master problem to be relatively easy to solve, feasible, a reasonable size, and have associated optimal duals that are close to optimal for the original problem. If these conditions occur, then the pricing subproblems have accurate information about reduced costs. For the passenger mix problem, our initial restricted master problem includes all the
columns associated with spilling to a null itinerary.

\[
\begin{align*}
\text{minimize: } & \quad \sum_{(r,g) \in (P,F)} \sum_{(p,f) \in (P,F)} f_{r,e_p,f} t_{p,f}^- \\
\text{subject to: } & \quad \sum_{(p,f) \in (P,F)} \delta_{i}^{p} t_{p,f}^- \geq Q_{i,j} - \text{CAP}_{i,j} \quad \forall (i,j) \in (L,J) \\
& \quad 0 \leq t_{p,f}^- \leq D_{p,f}. 
\end{align*}
\]

(3.8)

In the above formulation, \( t_{p,f}^- \) represents spilling from \((p,f)\) to the null itinerary. We notice that this is just the special case where we have no recapture. This problem can be solved by using a commercial optimization solver or Glover's network approach. This problem has at least one feasible solution: the trivial solution where all variables are at their upper bound. Since we only have one variable associated with each itinerary, we do not have any of the limit constraints. After this problem has been solved, we must determine which columns to generate.

If \( \lambda_{i,j} \) is the nonnegative dual cost associated with the capacity constraint for cabin \( j \) on flight leg \( i \) and \( \pi_{p,f} \) is the nonpositive dual cost associated with an expressed limit constraint for \((p,f)\), then the reduced cost of some variable \( t_{p,f}^{r,g} \), denoted by \( \tilde{c}_{p,f}^{r,g} \) is:

\[
\tilde{c}_{p,f}^{r,g} = f_{r,e_p,f} - \sum_{i \in P} \lambda_{i,J(f)} - b_{p,f}(f_{r,e_{r,g}} - \sum_{i \in r} \lambda_{i,J(g)}) - \pi_{p,f},
\]

(3.9)

where \( J(f) \) is the mapping from fare class \( f \) to the cabin with which it is associated.

### 3.6.2 The Solution Procedure

To solve the passenger mix problem with partial recapture, we integrate both column generation and constraint generation into the algorithm [28]. This approach is an iterative process with the following components (Figure 3.3):

1. Set \( k = 1 \). Denote an initial subset of columns \( A_1 \) to be used.

2. Solve a problem with the subset of columns \( A_k \).

3. Determine if any constraints are violated and add the violated constraints to the constraint
4. Price some of the remaining columns, and select a group that have reduced cost less than zero. Denote this subset of columns as $A^*$. Add these columns to the current problem to create $A_{k+1} = [A_k | A^*]$.

5. If no columns or rows are added, terminate. Otherwise, $k = k + 1$, go to step 2.

### 3.6.3 Computational Experience with Implementation

One of the implementation problems with column generation is its tendency to exhibit tailing. After several iterations of column generation, a good solution is reached, yet numerous
additional columns are generated since their reduced cost is negative. These added columns do not improve the solution because of degeneracy in the problem. Another problem in column generation is that too many good columns are generated. They all have large negative reduced costs, however, as soon as one is pivoted into the solution, the other columns added might have a less attractive reduced cost.

We experiment with different approaches to overcome these difficulties. In the first approach, we simply add all columns, up to a certain limit, that have negative reduced cost ("All"). In the second approach, for each \((p, f)\), we only add the variable \(t^*_{p,f}g^*\), where \((r^*, g^*)\) is the IFC with the minimum reduced cost ("Best Column"). If none of the variables associated with an IFC have a negative reduced cost, no columns are added for that specific IFC. The run time characteristics as a function of the solution procedure are in Table 3.1, for a variety of problems. We include the time to solve the first iteration in seconds, since the first iteration simply represents the situation with no recapture. For each solution procedure, we also include the total number of columns and rows at the last iteration, the number of iterations and the time in seconds to solve the problem.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Time to Solve 1st Iteration</th>
<th>All</th>
<th>Best Column</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cols</td>
<td>Rows</td>
</tr>
<tr>
<td>F97-3A</td>
<td>0.3</td>
<td>8,647</td>
<td>579</td>
</tr>
<tr>
<td>F97-4A</td>
<td>0.6</td>
<td>20,832</td>
<td>1,010</td>
</tr>
<tr>
<td>F97-6A</td>
<td>1.2</td>
<td>39,372</td>
<td>1,921</td>
</tr>
<tr>
<td>F97-9</td>
<td>2.2</td>
<td>72,357</td>
<td>3,795</td>
</tr>
<tr>
<td>A97-3A</td>
<td>0.2</td>
<td>9,510</td>
<td>384</td>
</tr>
<tr>
<td>A97-4A</td>
<td>0.8</td>
<td>25,079</td>
<td>1,213</td>
</tr>
<tr>
<td>A97-6A</td>
<td>1.6</td>
<td>47,129</td>
<td>2,585</td>
</tr>
<tr>
<td>A97-9</td>
<td>2.8</td>
<td>81,452</td>
<td>5,072</td>
</tr>
</tbody>
</table>

Table 3.1: The effects of the solution procedure on the number of iterations, generated rows and columns, and run times

In Table 3.1, we see that for every data set, using the Best Column approach outperforms the approach of including all negative reduced cost columns at each iteration. It is interesting to note that with the best column approach, fewer columns are added at each iteration, yet this leads to fewer iterations. This counterintuitive result is explained by the implementation of these approaches. Specifically, to limit computer memory requirements, at each iteration a limit
is set on the number of columns that are added. In the early iterations of the All approach, we reach this limit only part way through the pricing process, therefore some markets have many added variables, while others have none. On the other hand, every market in which recapture occurs has some variables added in the Best Column approach. This leads to better dual costs earlier in the solution process.

In general, the column and constraint generation method performs well. Table 3.2 compares the number of recapture variables and limit constraints generated to the actual number in the fully expressed problem for the Best Column approach. For limit constraints, we see that no more than 6% of these constraints need to be included for any problem. Also, we only generate roughly 2% - 3% of the recapture variables in the problem. One unexpected observation is that neither approach experiences the tailing phenomenon. In every instance the problems are solved to optimality in a few iterations, with each iteration improving the objective function value.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Recapture Variables</th>
<th>Limit Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Added</td>
</tr>
<tr>
<td>F97-3A</td>
<td>14,226</td>
<td>749</td>
</tr>
<tr>
<td>F97-4A</td>
<td>71,434</td>
<td>1,242</td>
</tr>
<tr>
<td>F97-6A</td>
<td>175,156</td>
<td>2,388</td>
</tr>
<tr>
<td>F97-9</td>
<td>459,263</td>
<td>3,984</td>
</tr>
<tr>
<td>A97-3A</td>
<td>19,654</td>
<td>455</td>
</tr>
<tr>
<td>A97-4A</td>
<td>87,920</td>
<td>2,558</td>
</tr>
<tr>
<td>A97-6A</td>
<td>212,761</td>
<td>3,778</td>
</tr>
<tr>
<td>A97-9</td>
<td>431,184</td>
<td>7,324</td>
</tr>
</tbody>
</table>

Table 3.2: The number of rows and columns that are added compared with the total number of rows and columns for the Best Column approach

There is one observation that should be noted regarding solution times. Using CPLEX's pre-solve procedures in earlier iterations and ignoring the advanced basis can lead to improved solution times. There does not seem to be consistency, however, as to when a shift to using an advanced basis should be made. Therefore, all procedures are run using pre-solve for the first iteration, and using the advanced basis obtained from the previous iteration for all remaining iterations. Improvements in solution time might be made by determining a proper trade-off between using pre-solve and advanced bases.

69
3.7 Assumptions and Data Issues

3.7.1 Airline Competition

The passenger airline industry is highly competitive, as evidenced by the constant fare wars, and waves of advertising. We consistently make the underlying assumption in all of the data that other airlines do not respond to the actions of a given airline. For some of the operational applications of our passenger mix model, this is reasonable. For the strategic and tactical applications, this might not be the case. For example, the values for fares on an itinerary are constantly changing because of adjustments to the competition. Making an assumption that the current fares are valid three months from now is highly suspect. Likewise, we assume that the competitor's flight network is fixed, thus, the demand data estimates are fixed. If another airline increases service in a competitive market, then this will have an effect on the demand forecasts for all itineraries in that market. Also, the recovery rates are directly related to the services that the other airlines provide. For complete accuracy, we need to consider these issues in our model, however, this is impossible because of the resulting intractability and the unavailability of the data. Therefore, we assume that competitors do not respond to our decisions.

3.7.2 Congestion

In this section, we evaluate the performance of the passenger mix model as a function of the congestion of the system. We define congestion as some measure of the traffic versus the capacity in the system. In other traffic flow problems, the solution time rises as congestion increases. To quantify the magnitude of this phenomenon in our model and solution procedure, we first define a measure of the congestion called the demand factor, which is the ratio of the seats demanded over the seats supplied. The mathematical expression for the system demand factor, \( DF \), is:

\[
DF = \frac{\sum_{(p,f) \in \{(P,F)\}} D_{p,f} \sum_{i \in L} \delta_{i}^{p}}{\sum_{i \in L} \sum_{j \in J} CAP_{i,j}}.
\]  

(3.10)
The term \( \sum_{i \in L} \delta_i \) gives us the number of flight legs in each itinerary. For example, an itinerary that contains three flight legs uses three units of capacity. We then weight this usage by the number of people desiring this itinerary. The denominator is the number of seat-legs that are available. To generate problems with different demand factors, we universally multiply the number of passengers for each IFC by a common factor.

A measure of how well the airline is able to accommodate the passenger requests given a specific fleeting is the system load factor, \( SLF \). The \( SLF \) is determined by the following expression:

\[
SLF = \frac{\sum_{(r,g) \in (P,F)} \sum_{(p,f) \in (P,F)} x^{r,g}_{p,f} \sum_{i \in L} \delta_i}{\sum_{i \in L} \sum_{j \in J} CAP_{i,j}}.
\]  

(3.11)

Note that our definitions of demand factor and system load factor are leg averages. Some airlines prefer to weight each flight leg by the flying miles of that leg, thus giving greater importance to long haul flights.

The system load factor is bounded above by the demand factor, with equality occurring when we accommodate all passenger requests. A large gap between the \( DF \) and the \( SLF \) corresponds to high spill rates. This gap increases as we increase the demand factor. The system load factors, the contribution, and the solution time as a function of demand factor for some of the problems are presented in Table 3.3.

<table>
<thead>
<tr>
<th>DF</th>
<th>SLF</th>
<th>Cont. ($K/day)</th>
<th>Time (sec)</th>
<th>SLF</th>
<th>Cont. ($K/day)</th>
<th>Time (sec)</th>
<th>SLF</th>
<th>Cont. ($K/day)</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>.3976</td>
<td>2.382</td>
<td>0.6</td>
<td>.3886</td>
<td>10.064</td>
<td>9.7</td>
<td>.3760</td>
<td>14.451</td>
<td>19.7</td>
</tr>
<tr>
<td>0.5</td>
<td>.4911</td>
<td>2.962</td>
<td>1.2</td>
<td>.4759</td>
<td>12.340</td>
<td>15.2</td>
<td>.4631</td>
<td>17.650</td>
<td>25.2</td>
</tr>
<tr>
<td>0.6</td>
<td>.5762</td>
<td>3.509</td>
<td>1.8</td>
<td>.5572</td>
<td>14.487</td>
<td>23.3</td>
<td>.5445</td>
<td>20.717</td>
<td>43.9</td>
</tr>
<tr>
<td>0.7</td>
<td>.6535</td>
<td>4.022</td>
<td>2.5</td>
<td>.6299</td>
<td>16.497</td>
<td>39.5</td>
<td>.6183</td>
<td>23.626</td>
<td>67.3</td>
</tr>
<tr>
<td>0.8</td>
<td>.7170</td>
<td>4.482</td>
<td>3.4</td>
<td>.6924</td>
<td>18.344</td>
<td>65.7</td>
<td>.6834</td>
<td>26.382</td>
<td>131.2</td>
</tr>
<tr>
<td>0.9</td>
<td>.7678</td>
<td>4.907</td>
<td>4.7</td>
<td>.7447</td>
<td>20.051</td>
<td>110.5</td>
<td>.7378</td>
<td>28.935</td>
<td>236.0</td>
</tr>
<tr>
<td>1.0</td>
<td>.8112</td>
<td>5.289</td>
<td>9.1</td>
<td>.7875</td>
<td>21.612</td>
<td>168.2</td>
<td>.7826</td>
<td>31.351</td>
<td>514.9</td>
</tr>
<tr>
<td>1.1</td>
<td>.8458</td>
<td>5.630</td>
<td>12.4</td>
<td>.8225</td>
<td>23.009</td>
<td>268.4</td>
<td>.8178</td>
<td>33.601</td>
<td>1,263.1</td>
</tr>
</tbody>
</table>

Table 3.3: The effects of congestion on system load factor, contribution, and run times

71
3.7.3 Passenger Connectivity

We define \( P_L \) as the set of local itineraries, i.e., itineraries with only one flight leg and \( P_C \) as the set of connecting itineraries. Obviously, \( P_L \) and \( P_C \) are disjoint sets and \( P_L \cup P_C = P \). We define the passenger connectivity ratio (PCR) as a measure for the percentage of connecting passengers in the system. The PCR is defined as:

\[
PCR = \frac{\sum_{(p,f) \in (P_C,F)} D_{p,f}}{\sum_{(p,f) \in (P,F)} D_{p,f}}.
\] (3.12)

If we change the PCR by multiplying the connecting demand by some factor, then the demand factor is not the same. Therefore, to modify the PCR and/or \( DF \), we multiply the local traffic by one factor, \( \theta_L \), and the connecting traffic by another, \( \theta_C \). If \( \overline{DF} \) is the desired demand factor and \( \overline{PCR} \) is the desired passenger connectivity ratio, we have:

\[
\overline{DF} = \frac{\theta_L \left( \sum_{(p,f) \in (P_L,F)} D_{p,f} \right) + \theta_C \left( \sum_{(p,f) \in (P_C,F)} D_{p,f} \sum_{i \in L} \delta^p_i \right)}{\sum_{i \in L} \sum_{j \in J} CAP_{i,j}} \text{, and}
\] (3.13)

\[
\overline{PCR} = \frac{\theta_C \left( \sum_{(p,f) \in (P_C,F)} D_{p,f} \right)}{\theta_L \left( \sum_{(p,f) \in (P_L,F)} D_{p,f} \right) + \theta_C \left( \sum_{(p,f) \in (P_C,F)} D_{p,f} \right)}.
\] (3.14)

Solving for the two unknowns, \( \theta_L \) and \( \theta_C \), results in

\[
\theta_L = \frac{\left( 1 - \frac{\overline{PCR}}{PCR} \right) \left( \frac{\overline{DF} \sum_{i \in L} \sum_{j \in J} CAP_{i,j} \sum_{i \in L} \delta^p_i}{\sum_{(p,f) \in (P_C,F)} D_{p,f} \sum_{i \in L} \delta^p_i} \right) \left( \sum_{(p,f) \in (P_C,F)} D_{p,f} \right)}{1 + \left( 1 - \frac{\overline{PCR}}{PCR} \right) \left( \sum_{(p,f) \in (P_C,F)} D_{p,f} \sum_{i \in L} \delta^p_i \right)} \text{, and}
\] (3.15)

\[
\theta_C = \frac{\overline{DF} \sum_{i \in L} \sum_{j \in J} CAP_{i,j} - \theta_L \left( \sum_{(p,f) \in (P_L,F)} D_{p,f} \right)}{\sum_{(p,f) \in (P_C,F)} D_{p,f} \sum_{i \in L} \delta^p_i}.
\] (3.16)

Table 3.4 shows the results when we change the passenger connectivity ratio, but keep the
demand factor at the level shown at the top of each column. One interesting observation is that the solution time decreases as the passenger connectivity ratio increases. Also, the contribution decreases as we increase the PCR. This occurs mainly for two reasons.

1. The system load factor has a tendency to decrease as we increase the PCR. This holds for the first two cases but in the second case, it begins to actually rise, but then decreases.

2. The unconstrained revenue decreases as we increase the PCR and keep the DF fixed. This reflects that in general the fare of a passenger flying on a multi-leg itinerary is less than the sum of the local fares on each leg.

3.7.4 Demand Data

Throughout this thesis, we refer to the values associated with the number of passengers on a given itinerary as demand data. This is not entirely accurate, since we do not have knowledge of the actual demand until it is realized at the day of departure. What we refer to as demand data is, in reality, demand estimates or demand forecasts. The basis for these demand estimates is historical demand data. In this section, we describe some inaccuracies in the historical demand data that need to be taken into account to determine accurate demand estimates.

Unconstrained Data

In this chapter, we qualify the phrase demand data by the word unconstrained. Unconstrained demand is defined as the number of passengers that fly on a specific itinerary if there is infinite

<table>
<thead>
<tr>
<th>PCR</th>
<th>SLF</th>
<th>Cont. ($K/day)</th>
<th>Time (sec)</th>
<th>SLF</th>
<th>Cont. ($K/day)</th>
<th>Time (sec)</th>
<th>SLF</th>
<th>Cont. ($K/day)</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>.6012</td>
<td>3,907</td>
<td>2.2</td>
<td>.5983</td>
<td>17,864</td>
<td>39.3</td>
<td>.6050</td>
<td>30,947</td>
<td>166.7</td>
</tr>
<tr>
<td>0.10</td>
<td>.5914</td>
<td>3,388</td>
<td>1.9</td>
<td>.5958</td>
<td>15,745</td>
<td>30.6</td>
<td>.6140</td>
<td>27,535</td>
<td>118.3</td>
</tr>
<tr>
<td>0.20</td>
<td>.5727</td>
<td>2,920</td>
<td>1.4</td>
<td>.5875</td>
<td>13,788</td>
<td>24.7</td>
<td>.6177</td>
<td>24,395</td>
<td>80.8</td>
</tr>
<tr>
<td>0.30</td>
<td>.5481</td>
<td>2,505</td>
<td>1.1</td>
<td>.5793</td>
<td>12,027</td>
<td>23.2</td>
<td>.6180</td>
<td>21,494</td>
<td>62.2</td>
</tr>
<tr>
<td>0.40</td>
<td>.5221</td>
<td>2,143</td>
<td>1.2</td>
<td>.5709</td>
<td>10,455</td>
<td>22.4</td>
<td>.6145</td>
<td>18,864</td>
<td>59.2</td>
</tr>
<tr>
<td>0.50</td>
<td>.4969</td>
<td>1,828</td>
<td>1.4</td>
<td>.5600</td>
<td>9,038</td>
<td>22.5</td>
<td>.6079</td>
<td>16,461</td>
<td>59.2</td>
</tr>
</tbody>
</table>

Table 3.4: The effects of passenger connectivity on system load factors, contribution and run times

73
capacity on each flight leg. A naive implementation relying on historical data might inaccurately predict actual unconstrained demand for the flight leg, since the capacity of flight legs has historically constrained booking requests. The problem of determining unconstrained demand forecasts from historical demand data is a particularly difficult one.

For example, Figure 3-4 shows the unconstrained demand function for an arbitrary flight leg. If the capacity for this flight leg is 30, then in isolation of other effects, historical demand has the shape shown in Figure 3-5. In the specific instances when the realized demand is greater than 30, a historical demand of 30 is shown, since we turn away the excess passengers. The demand distribution, however, typically resembles Figure 3-6 because of no-shows and overbooking.

Determining the demand for a specific IFC is much more difficult. There is not a straightforward barrier that truncates the historical demand. The complexity of revenue management systems might cause the following to occur: one day an airline accepts only three of five requests for a specific itinerary; on another day, the airline might accept seven of seven requests. Typically, efforts in this area of unconstraining historical itinerary demand constitute a competitive advantage for an airline. As a result, little is written about it. So, for this research project, we assume that unconstrained demand data are available.
Figure 3-5: The idealized historical constrained demand function

An important assumption is that the passengers who desire travel on the same IFC are homogeneous. Specifically, this has importance in modeling recapture. We assume that every passenger desiring a particular IFC has the same probability of accepting some other alternative itinerary if offered.

**Historical Aggregation**

Belobaba and Farkas [10] discuss the inaccuracies of spill models when using aggregated data. Specifically, they define two types of aggregation. *Vertical aggregation* refers to the practice of combining passengers from different fare classes, described in Section 2.2.4. *Horizontal aggregation* refers to aggregating data from different days of the week, which typically have dissimilar traffic patterns. Their work shows significant spill underestimation using simple spill models and aggregated data.

The demand data in this research are both vertically and horizontally aggregated. In this section, we investigate how a disaggregated data set compares to an aggregated data set. The first step is to develop a disaggregation model, beginning by considering vertical aggregation. We examine each itinerary and disaggregate it as follows. We pre-determine the number of fare
classes to use in our model, denoted $FC_{\text{max}}$. We do not want to have many fare classes if the demand for a specific itinerary is small, since this will lead to really small numbers for these $IFC$'s. Therefore, the number of fare classes used for itinerary $p$, denoted $f_{cp}$, is estimated as follows:

$$f_{cp} = \min\{FC_{\text{max}}, [2D_pFC_{\text{max}}]\}.$$  \hspace{1cm} (3.17)

We use a simple formula for breaking out the number of passengers in each fare class. If the fare classes are ordered from 1 to $f_{cp}$, with fare class 1 being the highest yield class, then

$$D_{p,f} = \frac{f}{\left(\sum_{g=1}^{f_{cp}} g\right)}D_p, \hspace{1cm} f = 1, \ldots, f_{cp},$$  \hspace{1cm} (3.18)

where $D_p$ is the average demand for itinerary $p$. To determine the fare value for each fare class, we use another parameter, the fare value spread, $FC_{sp}$. Let

$$fare_{p,1} = (FC_{sp})^2 fare_{p,f_{cp}}.$$  \hspace{1cm} (3.19)
To maintain the same average fare, $\bar{\text{fare}}_p$, for the aggregated itinerary, the fare value for each fare class is calculated by the following relationship:

$$
\text{fare}_{p,f} = \frac{\left( \sum_{g=1}^{f_{cp}} g \right) (FC_{sp})^{\frac{2(f-1)}{f_{cp}-1}}}{\sum_{g=1}^{f_{cp}} g(FC_{sp})^{\frac{2(g-1)}{f_{cp}-1}}} \text{fare}_p.
$$

(3.20)

To model horizontal disaggregation, we assume that an instance of demand is universally multiplied by some factor $\gamma$. That is, for instance $j$, the demand for itinerary $p$ is $D_p^j = \gamma_j D_p$. To determine the set of multipliers we use two parameters. The first parameter, $SP_{HA}$, is the horizontal aggregation spread, measuring the spread between the highest and lowest multiplier of the demand data. The second parameter, $N_{HA} > 1$, is the number of instances we wish to examine. If $P_j$ is the probability of instance $j$ occurring, then

$$
\gamma_j = 1 - SP_{HA} \frac{1 - j + \sum_{i=1}^{N_{HA}} P_i (i - 1)}{N_{HA} - 1}, \quad j = 1, 2, ..., N_{HA},
$$

(3.21)

where we have assumed that $\gamma_j - \gamma_{j-1} = \gamma_{j+1} - \gamma_j$ for every $j$.

### 3.7.5 Recovery Rates

Another difficult problem is determining the recovery rates for alternative itineraries. No matter the sophistication of the model used for determining recovery rates, the values are ballpark figures at best. With that said, the basis for calculating the recovery rates in these experiments is the Quantitative Share Index ($QSI$). This industry standard measures the “attractiveness” of an itinerary relative to the entire set of other itineraries (including competing airlines) in that market. In Section 3.7.1, although we discuss how airline competition affects demand estimates, we assume that this index is fixed.

The model that we use for passenger recapture is:

1. Passengers wishing to fly on a specific $IFC$ come to the airline with their collective booking requests.
2. Based on the passenger mix model, the airline partitions these passengers into one or more groups. The airline offers one IFC option to each group. It might be the desired IFC, an alternative IFC, or the null itinerary. All other IFC's of the airline are closed to the passenger except the alternative that is offered.

3. Based on the recovery rates, the passengers decide whether to fly on the offered IFC or try another airline. If the passengers try another airline or are offered the null itinerary, their revenue is assumed lost to the airline.

For a specific passenger, the passenger tells the airline which IFC they desire, and the airline returns with the IFC it is willing to offer. There is a probability that the passenger accepts this alternative. Therefore, the value that we desire is to find the probability of a person accepting the alternative IFC given that this is the only option that they are offered by this airline. We assume that if the alternative IFC is the same as what is desired, the passenger accepts it with probability 1.0.

We now use the QSI to measure the probability of accepting the alternative IFC. The sum of QSI corresponding to all itineraries (including competitors) in a market is equal to one. The sum of the QSI for one airline is an approximate measure of its market share for that specific market. Therefore, if the airline offers a passenger only one of their itineraries, it is effectively removing all of their other itineraries from this market. Let \( q_p \) denote the QSI value of itinerary \( p \). Let \( Q_m \) represent the sum of all QSI values in market \( m \) for the airline, i.e., \( Q_m = \sum_{p \in m} q_p \). Then, the base recovery rate, \( \bar{b}_{p,f}^r \), is:

\[
\bar{b}_{p,f}^r = \begin{cases} 
1.0 & \text{if } p = r \text{ and } f = g \\
\frac{q_r}{1-Q_m + q_r} & \text{otherwise.}
\end{cases} \tag{3.22}
\]

The base recovery rate is a measure of the probability of accepting the alternative itinerary as long as fare and difference in departure time isn’t a factor. The attractiveness of the alternative is based on the time of day of departure, length of trip, and number of connections.

While this base recovery rate gives us a starting point, we modify it for different fare classes and similarities in departure (or arrival) times. Obviously, if the airline offers someone a higher fare class than what they want, they might choose another airline if it has space available at the
lower fare class. If a passenger is offered an itinerary that has a similar departure and arrival time as the desired itinerary, he/she is more likely to accept the alternative. Therefore, the recovery rate should be higher if there are similar departure and arrival times, and lower if there are drastic differences in departure and arrival times. Also, the fare class needs to be considered in modifying the recovery rate. Typically high fare class passengers are time sensitive business travellers, while low fare class passengers are usually price-sensitive leisure travellers.

The results of this process are rough estimates of the recovery rates. Accurately determining recovery rates is difficult if not impossible. In Section 3.7.5, we measure how sensitive the model is to changes in the recovery rates.

The Magnitude of Recapture

If recapture accounts for only a small percentage of contribution, it isn’t necessary to model it exactly. This has important ramifications since solving models that consider recapture is much more difficult than solving models that ignore it. Using the recapture model described above, we compare the contributions of no recapture and partial recapture solutions. We also show the effects of modeling partial recapture on the dual costs associated with the flight leg capacity constraints. Note that the base recovery rate is used for the calculations in Tables 3.5 and 3.6.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>No Recapture Contribution ($/day)</th>
<th>Recapture Contribution ($/day)</th>
<th>Absolute Change ($/day)</th>
<th>Percentage Increase (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>5,266,476</td>
<td>5,438,283</td>
<td>171,807</td>
<td>3.26</td>
</tr>
<tr>
<td>F97-4A</td>
<td>9,185,166</td>
<td>9,381,016</td>
<td>195,850</td>
<td>2.13</td>
</tr>
<tr>
<td>F97-6A</td>
<td>14,823,592</td>
<td>15,253,906</td>
<td>430,314</td>
<td>2.90</td>
</tr>
<tr>
<td>F97-9</td>
<td>22,122,871</td>
<td>22,423,069</td>
<td>300,198</td>
<td>1.36</td>
</tr>
<tr>
<td>A97-3A</td>
<td>3,657,230</td>
<td>3,664,274</td>
<td>7,044</td>
<td>0.19</td>
</tr>
<tr>
<td>A97-4A</td>
<td>8,547,672</td>
<td>8,574,384</td>
<td>26,712</td>
<td>0.31</td>
</tr>
<tr>
<td>A97-6A</td>
<td>15,313,194</td>
<td>15,512,526</td>
<td>199,332</td>
<td>1.30</td>
</tr>
<tr>
<td>A97-9</td>
<td>23,247,134</td>
<td>23,624,770</td>
<td>377,636</td>
<td>1.62</td>
</tr>
</tbody>
</table>

Table 3.5: The effect of recapture on contribution

From Table 3.5, we see that considering recapture has a large impact on contribution. For the two full size problems, the percentage increase in contribution with recapture is roughly 1.5% compared to the contribution without recapture. This improvement translates to a daily
contribution over $300K/day ($110M/year).

The dual prices associated with the flight leg capacity constraints in the passenger mix model represents the additional contribution of adding one more seat to the flight leg capacity, or equivalently, the contribution of the last passenger spilled on that flight leg. In Table 3.6, we notice that most non-zero duals change value when incorporating recapture. For the full sized problem, this average change is on the order of $25 - $30 per flight leg. However, we notice that there is at least one dual for which the difference in the dual price is $400. As we will show in Chapter 4, dual prices are one mechanism for estimating seat values for use in advanced revenue management schemes to approximate network effects and control seat inventory. Vastly different dual costs resulting from including or excluding recapture has a significant impact on these network revenue management schemes.

**Recovery Rate Sensitivity**

Finally, we examine the sensitivity of the optimal solution to perturbations in the recovery rates. We perturb the recovery rate in the following way:

1. Select a specific recovery rate $\tilde{b}_{p,f}$. Determine whether this is to be increased or decreased. This can be done either randomly or systematically.

2. Randomly determine a percentage change, $\beta$, which is bounded by the maximum magnitude of a change, $\beta_{\text{max}}$.

3. Calculate the new recovery rate, $\tilde{b}'_{p,f}$, based on a convex combination of the recovery rates.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Flights</th>
<th>Num. of Non-Zero Duals</th>
<th>Change in Duals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No Recapure Recapture</td>
<td>Num. Total</td>
</tr>
<tr>
<td>F97-3A</td>
<td>157</td>
<td>24</td>
<td>27</td>
</tr>
<tr>
<td>F97-4A</td>
<td>431</td>
<td>47</td>
<td>50</td>
</tr>
<tr>
<td>F97-6A</td>
<td>823</td>
<td>73</td>
<td>78</td>
</tr>
<tr>
<td>F97-9</td>
<td>2,044</td>
<td>125</td>
<td>125</td>
</tr>
<tr>
<td>A97-3A</td>
<td>173</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>A97-4A</td>
<td>485</td>
<td>53</td>
<td>57</td>
</tr>
<tr>
<td>A97-6A</td>
<td>877</td>
<td>88</td>
<td>89</td>
</tr>
<tr>
<td>A97-9</td>
<td>1,888</td>
<td>182</td>
<td>181</td>
</tr>
</tbody>
</table>

Table 3.6: The effect of recapture on the dual variables
percentage and the bound in the direction we are moving. Specifically, if we are increasing
the rate, then \( \tilde{b}^r_{p,f} = (1-\beta)b^r_{p,f} + \beta \) and if we are decreasing the rate, then \( \tilde{b}^r_{p,f} = (1-\beta)b^r_{p,f} \).

4. To disallow huge changes for recovery rates that are close to 0 or 1, we also have the
following limits. In the case when we are increasing the recovery rate, \( \tilde{b}^r_{p,f} \leq \omega b^r_{p,f} \); in the
case when we are decreasing the recovery rate, \( \tilde{b}^r_{p,f} \geq 1 - \omega (1 - b^r_{p,f}) \). We set \( \omega = 5 \).

<table>
<thead>
<tr>
<th>( \beta_{\text{max}} )</th>
<th>Contribution ($/day)</th>
<th>Non-Zero Duals</th>
<th>Change from Base Case</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Total</td>
<td>Average</td>
</tr>
<tr>
<td>Base</td>
<td>23,624,770</td>
<td>181</td>
<td>NA</td>
</tr>
<tr>
<td>.10</td>
<td>23,657,700</td>
<td>183</td>
<td>157 $1,071.65</td>
</tr>
<tr>
<td>.20</td>
<td>23,690,105</td>
<td>179</td>
<td>165 $1,802.92</td>
</tr>
<tr>
<td>.30</td>
<td>23,715,441</td>
<td>176</td>
<td>162 $2,524.41</td>
</tr>
<tr>
<td>.40</td>
<td>23,736,613</td>
<td>178</td>
<td>178 $3,136.83</td>
</tr>
<tr>
<td>.50</td>
<td>23,753,251</td>
<td>178</td>
<td>180 $3,624.16</td>
</tr>
</tbody>
</table>

Table 3.7: The sensitivity of the passenger mix model with random perturbations in the recovery rates

<table>
<thead>
<tr>
<th>( \beta_{\text{max}} )</th>
<th>Contribution ($/day)</th>
<th>Non-Zero Duals</th>
<th>Change from Base Case</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Total</td>
<td>Average</td>
</tr>
<tr>
<td>Base</td>
<td>23,624,770</td>
<td>181</td>
<td>NA</td>
</tr>
<tr>
<td>.10</td>
<td>23,699,873</td>
<td>182</td>
<td>164 $1,537.60</td>
</tr>
<tr>
<td>.20</td>
<td>23,767,778</td>
<td>179</td>
<td>164 $2,599.48</td>
</tr>
<tr>
<td>.30</td>
<td>23,827,205</td>
<td>178</td>
<td>167 $3,438.61</td>
</tr>
<tr>
<td>.40</td>
<td>23,881,181</td>
<td>179</td>
<td>170 $4,086.87</td>
</tr>
<tr>
<td>.50</td>
<td>23,932,438</td>
<td>180</td>
<td>171 $4,791.58</td>
</tr>
</tbody>
</table>

Table 3.8: The sensitivity of the passenger mix model with increased perturbations in the recovery rates

Tables 3.7, 3.8, and 3.9 detail the results of this perturbation experiment for data set A97-9. We see how the changes affect the contribution, the spill costs, and some measures of
the magnitude of change in the duals associated with the flight legs. The first group of data
are the results from random perturbation in both directions. The second set of data results
from random perturbations when we only increase the recovery rates. Lastly, the third set of
data results from random perturbations when we only decrease the recovery rates. Even when
perturbations are limited to small variations (\( \beta_{\text{max}} = 0.10 \)). This can have an impact of almost
\$30K/day (\$10M/year) in the case of random perturbations. The dual costs vary by almost
$11 and the maximum dual cost change is over $200. This leads us to think that the model is fairly sensitive to the recovery rates used and therefore much thought must go into their use.

### 3.7.6 Fare Data

In the competitive environment of airlines, fare wars occur with regular frequency. Even in more peaceful times, airlines continually adjust fares to match their competitors' within 24 or 48 hours of changes. Predicting how fare values change over the planning horizon is difficult. Usually, most models do not consider uncertainty in fare values and assume deterministic values. We use this assumption throughout the rest of the thesis.

### 3.7.7 Parallel Markets

A parallel market is two or more origin-destination airport pairs that customers can reasonably substitute for one another [36]. For example, a passenger travelling from the New York City area to the metro Washington DC area might be relatively indifferent to the origin airport (Laguardia, JFK, or Newark) and/or the destination airport (Washington National, Dulles, or Baltimore-Washington International). In this case, there are alternative itineraries in nine different markets that should be considered. In our computational experiments, we do not consider such shifts to parallel markets, however, the model is capable of handling such shifts.

<table>
<thead>
<tr>
<th>$\beta_{\text{max}}$</th>
<th>Contribution ($/\text{day}$)</th>
<th>Non-Zero Duals</th>
<th>Number</th>
<th>Total</th>
<th>Average</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>23,624,770</td>
<td>181</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>.10</td>
<td>23,602,706</td>
<td>183</td>
<td>120</td>
<td>$304.56</td>
<td>2.54</td>
<td>$36.82</td>
</tr>
<tr>
<td>.20</td>
<td>23,581,667</td>
<td>182</td>
<td>130</td>
<td>$497.28</td>
<td>3.60</td>
<td>$36.78</td>
</tr>
<tr>
<td>.30</td>
<td>23,560,749</td>
<td>182</td>
<td>146</td>
<td>$669.92</td>
<td>4.59</td>
<td>$49.09</td>
</tr>
<tr>
<td>.40</td>
<td>23,541,581</td>
<td>182</td>
<td>149</td>
<td>$878.81</td>
<td>5.90</td>
<td>$62.61</td>
</tr>
<tr>
<td>.50</td>
<td>23,523,327</td>
<td>182</td>
<td>151</td>
<td>$1,058.81</td>
<td>7.01</td>
<td>$62.61</td>
</tr>
</tbody>
</table>

Table 3.9: The sensitivity of the passenger mix model with decreased perturbations in recovery rates

82
3.8 The Passenger Mix Model with Stochasticity

One of the main limitations of the model presented is that it ignores the stochasticity of the demand data. For now, we limit our consideration to uncertainty in the number of passengers and not to uncertainty in the fares. We assume that we can accurately forecast the fare for each IFC and the first two moments of the passenger demand distribution for each itinerary.

3.8.1 Demand Distribution

Airline passenger demand is highly uncertain. This variability comes from numerous sources: day-of-the-week and seasonal effects, fare wars, increased competition in key markets, growth during economic booms, and contraction during recessions. Even if we account for all of these conditions, there is still significant randomness. Ideally, we wish to consider this randomness in all of the airlines' decision processes.

There have been numerous studies on the shape of flight leg demand distributions [6], [30], [41]. All of the following distributions have been considered to model demand: Gaussian, log-normal, Poisson, and gamma. Although each distribution model has advantages and disadvantages, the Gaussian distribution is typically used as a good approximation for the demand distribution for a flight leg. When the standard deviation is large with respect to the mean, however, it is possible to have positive probability associated with negative demand. Also, the Gaussian distribution is symmetrical which isn't necessarily an accurate representation of demand with small means. Because of the possibilities of negative demand and the symmetry around the mean, some researchers prefer either a Poisson or gamma distribution, especially when trying to model itinerary data which has small demand numbers and skewness.

Typically, the variance of demand on a flight leg is dependent on the mean of the flight leg [6]. There are two simple variance models which are usually employed. One is to set the standard deviation equal to a constant multiplied by the expected demand [41]. This constant, defined as a $K$-factor or the coefficient of variation, is typically between 0.2 and 0.5. Another is to set the standard deviation equal to a constant multiplied by the square root of the mean [10]. This constant is referred to as a $Z$-factor, and typically has values between 1.0 and 2.5. If the data are available, the variance is sometimes calculated by using historical demand distributions.
In the passenger mix model, we wish to model demand for a specific itinerary, not for a specific flight leg. Some highly traveled itineraries (direct flights from JFK-SFO) can be accurately approximated by a Gaussian distribution, other itineraries, such as Albuquerque to Allentown, have extremely small expected demand, on the order of a couple passengers per month. If a Gaussian distribution is used for these itineraries, instances of negative demand will have significant probability.

3.8.2 A Formulation with Stochastic Demand

![Marginal Fare Value($) vs Seats Reserved](image)

Figure 3-7: An expected marginal seat revenue curve

We present a formulation of the passenger mix model that approximates the uncertainty in demand [21], in an attempt to calculate a more accurate estimate of the attainable contribution. For each \((p, f)\), instead of a demand of \(D_{p,f}\) passengers, we have a number of segments of demand which when summed together will be much larger than \(D_{p,f}\) passengers. The segment represents a group of passengers that have similar expected marginal seat revenue. For example, Figure 3-4 contains a discretized Gaussian distribution for an itinerary with an expected demand of 20.0
and a standard deviation of 6.0. Assume that the fare for this IFC is $100. The bar graph in Figure 3-7 shows the expected marginal seat revenue of this distribution. The dotted line shows the deterministic model's representation of the expected marginal seat revenue. Specifically, the deterministic model assumes that 20.0 passengers have $100 expected marginal seat revenue and additional passengers have $0. The heavy line represents another approximation. The fare associated with each segment, in this case, is discounted to model the fact that this segment of passengers will not be realized 100% of the time.

That is, we show the expected revenue received by allocating one more seat to this itinerary. The heavy line shows how we are capturing stochasticity with deterministic demand. For our mathematical formulation, we define a new set and a new traffic variable:

$$S(p, f):$$ the ordered set of different segments for $$(p, f)$$.  

$$x_{p,f,s}^{r,g}$$: the number of passengers in segment $$s$$ who desire travel on $$(p, f)$$, but instead travel on itinerary $$r$$ under fare class $$g$$.

Our new formulation is:

$$(SPMIX)$$

maximize: \[ \sum_{(r,g) \in (P,F)} fare_{r,g} \sum_{(p,f,s) \in (P,F,S(p,f))} \beta_{p,f}^{s} x_{p,f,s}^{r,g} \]

subject to:

\[ \sum_{(r,g) \in (P,F)} \sum_{(p,f,s) \in (P,F,S(p,f))} \delta_{i,j} x_{p,f,s}^{r,g} \leq CAP_{i,j} \quad \forall (i,j) \in (L,J) \tag{3.23} \]

\[ \sum_{(r,g) \in (P,F)} x_{p,f,s}^{r,g} / b_{p,f,s} \leq D_{p,f,s} \quad \forall (p,f) \in (P,F,S(p,f)) \]

\[ x_{p,f,s}^{r,g} \geq 0. \]

where \(\beta_{p,f}^{s}\) is a factor to approximate the expected marginal revenue for segment $$s$$. One possibility is to define \(\beta_{p,f}^{s}\) as:

$$\beta_{p,f}^{s} = P \left( \frac{d_{p,f}}{\sum_{i=1}^{s-1} D_{p,f,i}} > \delta_{s} \right),$$

(3.24)
i.e., $\beta_{p,f}^s$ is the probability that the realized demand is greater than or equal to $\sum_{i=1}^{s-1} D_{p,f,s}$.

Farkas [21] shows a serious problem with this model. In the partitioning of the cabin and assignment of a partition of seats to each itinerary, we might overestimate spill. For example, say that for some solution we have $\sum x_{p,f,s}^{p,f} = 20$, i.e., we reserve twenty seats on every flight leg in itinerary $p$ for passengers flying on $(p,f)$. If only fifteen passengers from this itinerary-fare class actually fly, then this model says that there are five seats on each one of these flight legs that are empty. In reality, if these flight legs are in high demand, the airline fills these seats with other itineraries.

3.9 Conclusion

In this chapter we present a linear programming formulation that models the flow of airline passengers over a flight schedule. Unlike previous research, this model includes the partial recapture of passengers when the desired itinerary is not available. Also, we present a change of variables that has, in practice, allowed us to reduce drastically the number of explicit constraints in the problem. This results in a tractable problem, and an effective and quick solution procedure using column and row generation methods. We show that run times are affected significantly by congestion in the system and the amount of network effects.

Most importantly, we show that recapture represents a significant amount of contribution in determining the attainable contribution. Also, the empirical evidence suggests that this model is sensitive to the values of the recovery rates. This suggests that future research is needed to investigate the concept of recapture. One of the more notable issues involved in different recapture models is how drastically dual costs are affected in the passenger mix model. We show in Chapter 4 that the dual costs from the passenger mix model can be used in network revenue management. This is one of numerous uses of the passenger mix model in the airline industry.
Chapter 4

Applications of the Passenger Mix Model

4.1 Overview

In Chapter 3, we describe the passenger mix model which can be used as a spill calculator for the iterative fleet assignment paradigm in Chapter 2. The passenger mix model has numerous applications beyond a simple calculator. It is interesting to note that this model has applications in all scopes of decision-making faced by the airline, from operational to strategic. In this chapter, we present an application in each of the following scopes: operational, tactical, and strategic.

4.2 Irregular Operations

In the scheduling process (Chapter 1), an airline develops a schedule and its plan for execution with little concern for the potential problems that might cause a break down of this plan. These problems include but are not limited to severe weather, mechanical failure of equipment, and delayed crews. When one or more problems occur, the schedule and plan must be altered, often quickly, to satisfy numerous constraints. The result can be flight cancellations or significant delays. The aftermath of unexpected events that have a significant impact on the carrier's schedule is defined as an irregular operation [15]. These cancellations and delays, in turn, might
lead to re-fleeting and rerouting of aircraft, maintenance reallocation, crew rescheduling, and ultimately, the rerouting of passengers. Today, this replanning process is carried out manually. Clarke [15] calls this the airline recovery problem and proposes an optimization-based heuristic solution approach that can be incorporated into various decision support systems.

One of the many issues that must be addressed by the professionals at the airline operations control center is how their decisions affect the airline’s ability to transport all passengers to their ultimate destination in a timely fashion. Clarke cites the model presented in Chapter 3 as a viable solution technique for this subproblem of determining passenger flows in irregular operations because of its flexibility and speed.

4.2.1 Implementation

To implement the passenger mix model in an irregular operations framework, we must rethink the nature of how passengers are modeled. In the passenger mix model presented in Chapter 3, the passenger was undecided on which airline he/she would actually fly. In this case, the passenger has already bought a ticket, and it is now the responsibility of the airline to provide transportation from their origin (or current location) to their final destination.

Since the demand is realized, in the absence of overbooking, the quantity $Q_i - CAP_i$ is always less than zero and therefore all of the capacity constraints are satisfied. If there are no significant delays or cancellations, then this problem is trivial, since no passengers need to be rerouted. However, irregular operations cause delays, re-fleeting, and cancellations. In this section, we present some methods to modify the passenger mix model to capture these situations.

Unaffected Passengers

From the standpoint of customer service, it is bad business to re-route passengers whose itineraries are not affected by the irregular operations. First, let $L_U$ be the set of flight legs that are unaffected by the irregular operations. We define a flight leg as unaffected if the flight leg is not cancelled, the fleeting does not change and the scheduled departure or arrival time does not shift by more than some tolerance. If all of the flight legs in a passenger’s itinerary, $p$, are in the set $L_U$, then that passenger is considered unaffected and should not be rerouted.
To enforce this, we do not include spill variables associated with this itinerary in the passenger mix model. We do, however, include the demand, \( D_p \), in the calculation of \( Q_i \).

**Flight Leg Re-Fleetings and Cancellations**

Because of irregular operations, the airline might either re-fleet a flight leg with a new fleet type or cancel the flight leg. We model these situations similarly in the passenger mix model. We formulate the passenger mix model as usual, however, on the right hand side of the capacity constraints of Program 3.3, we change \( Q_i - \overline{CAP}_i \) to \( Q_i - \overline{CAP}_i \), where \( \overline{CAP}_i \) is the number of seats on the new fleet type. For flight leg cancellations, \( \overline{CAP}_i = 0 \).

**Significant Delays**

When a significant delay occurs, this might cause some passenger itineraries to become invalid. Specifically, if there is a significant delay on the first flight leg of a passenger's itinerary, he/she might not be able to make the connection for the second leg. In these cases, we change the limit constraints from

\[
\sum_{(r,g)\in(P,F)} t^r,g_{p,f} \leq D_{p,f}
\]  \hspace{1cm} (4.1)

to

\[
\sum_{(r,g)\in(P,F)} t^r,g_{p,f} = D_{p,f},
\]  \hspace{1cm} (4.2)

e.i., we force the model to spill all passengers on invalid itineraries. Also, we eliminate all variables that represent spill to invalid itineraries.

**Recovery Rates**

Since the airline's preference is to reassign most passengers to alternate routes on the same airline, we assume that \( b_p^r \) is equal to 1.0. Recovery rates at or close to 1 can also be justified by considering that many irregular operations are caused by severe weather, this affects all airlines, leading to minimal space on other airlines to which the passenger can be spilled.
4.2.2 Different Objective Functions

Clarke suggests a number of different objective functions for the passenger flow recovery problem. One of these is minimizing overall passenger delay time. The passenger mix model can be converted quite easily for this purpose. For simplicity and ease of notation, we assume only one fare class and use the keypath formulation. Instead of

\[ \text{fare}_p - b_p^r \text{fare}_r \]  \hspace{1cm} (4.3)

as the cost coefficient, we change it to

\[ (1 - b_p^r) \psi + b_p^r d(p, \hat{r}), \]  \hspace{1cm} (4.4)

where \( \hat{r} \) is the newly rescheduled itinerary \( r \), \( p \) is the originally scheduled itinerary, \( \psi \) is a penalty in units of time for passengers that the airline decides to transfer to another airline, and \( d(p, \hat{r}) \) is the difference in arrival time at the destination airport of itinerary \( p \) and itinerary \( \hat{r} \).

4.3 Revenue Management

Another application of the passenger mix model involves some of the tactical and operational issues involved in revenue management. The passenger mix model can be used to incorporate network effects into the revenue management decision process. Also, it gives an upper bound on the expected contribution that can be achieved by an arbitrary revenue management process. It is not in the scope of this section to give a fully detailed literature review of the entire body of work done in revenue management. For this, the reader is directed to Weatherford and Bodily [46], Belobaba [7], and Williamson [47]. The literature review in this section closely follows chapter 3 in Farkas [21].

4.3.1 A Revenue Management Primer

One of the hot topics in the airline operations research community has been revenue management. The goal of revenue management is to maximize profits by selling seats to low fare
passengers for which there is no high fare passenger demand. Originally, this process was called 
*yield management*, since passengers were partitioned into classes based on their yield (fare per 
mile). Now, many prefer the more accurate title of *revenue management or perishable asset 
revenue management* (*PARM*) [46]. It can be shown with simple examples that achieving the 
best yield per passenger does not necessarily mean maximizing revenues and profits. Revenue 
management has been used by airlines, cruise lines, car rental agencies, hotels, and numerous 
other industries.

Wong et al. [48] outline five components of revenue management:

1. **Segmentation:** The identification of market segments willing to pay different prices for 
   use of the same resource.

2. **Pricing:** The development of products that carry different prices and service character-
   istics

3. **Restricting:** The development of rules or “fences” to prevent customers willing to pay 
   higher prices from lower prices.

4. **Forecasting:** Predicting the demand in the different segments.

5. **Control:** An approach to the allocation of the shared resources to maximize some pre-
   determined quantity, usually revenue.

These components are highly dependent on one another. For example, the effectiveness 
of allocating resources is only as good as the forecasts of demand. Similarly, the ability to 
forecast demand accurately is determined by how finely the market is segmented. Typically, 
segmentation, pricing and restricting are issues that are considered by marketing and sales 
departments. Forecasting has fallen under the domain of demand modelers. Control has been 
the main focus of operations researchers, and is usually referred to as *seat allocation, seat 
assignment, or seat inventory control* in the airline industry. Because of complexities that are 
discussed in Section 4.3.3, this is still an ongoing area of research. We focus on control and 
assume that segmentation, pricing, and forecasting results are inputs to the control process.

For the sake of simplicity, we present the problem of determining a seat inventory control 
scheme for a single flight leg in isolation. The goal is to maximize the expected revenues for that
specific flight leg. It is assumed that the costs associated with implementing the seat inventory control scheme (computers, software, training) are negligible compared to the potential benefits that might be derived from it.

4.3.2 Leg-Independent Seat Inventory Control

We begin with a single flight leg that is fleetted with an aircraft having a capacity of $S$ seats. There is a set, $F$, of fare classes whose elements $\{f_1, f_2, \ldots, f_{|F|}\}$ represent different market segments and are sorted in ascending order by contribution (the received fare - the carrying and booking costs), i.e., the contribution from a passenger in fare class $f_i$ is less than the contribution from a passenger in fare class $f_j$ if $i < j$. The research assumptions typically used are [16]:

1. Lower-valued fare classes book before higher-valued fare classes.

2. There are no cancellations.

3. Demand among fare classes is independent.

4. A denied request is lost revenue to the airline, that is, a passenger will not buy a higher valued ticket (passenger sell-up) or book on another flight on the same airline (recapture).

For illustrative reasons, let’s assume just two fare classes. A static model allocates $S_1$ seats to fare class $f_1$ and $S - S_1$ seats to fare class $f_2$. If $c_i$ is the contribution from one passenger booked on fare class $f_i$, then the optimal allocation $S_1^*$ is such that

$$\frac{P_1(S_1^*)}{P_2(S - S_1^*)} = \frac{c_2}{c_1}, \quad (4.5)$$

where $P_i(s)$ is the probability that the demand for fare class $f_i$ is at least $s$ passengers. This simple formula is based on the idea that we set the seat allocation such that we cannot increase contribution by allocating one more seat to $f_1$ by decreasing the allocation $f_2$ and vice versa. Because of the discrete nature of demand, an integer value of $S_1^*$ that satisfies Equation 4.5 might not exist.

The discussion assumes that fare classes are not nested, meaning that an availability in the lower fare class is not given to the upper fare class. Littlewood [31] determines an optimal
control procedure in a nested environment with two fare classes by relying on the dynamic nature of the booking process. Since the lower fare classes book before the higher fare classes, he uses the expected marginal approach where up to $S_1$ lower fare class passengers are booked as long as

$$c_1 \leq c_2 P_2 (S - S_1).$$

(4.6)

Today, most airlines offer numerous fare classes based on different restrictions, thus necessitating an approach that is applicable for numerous fare classes. Belobaba [8] proposes an approach (later named EMSRa) in which the pairwise expected marginal revenues of the fare classes are equal. His approach, however, does not account for the nesting of multiple fare classes. Curry [16] solves for the optimal booking limits (OBL) with multiple nested fare classes on a single flight leg. The numerous convolution integrals in his approach, however, results in difficult computations. Belobaba [9] proposes the heuristic EMSRb that uses a weighted average contribution for all higher fare classes. Computational experience shows that the differences between EMSRb and OBL is negligible, with EMSRb requiring much easier calculations. In fact, as some of the assumptions stated above are relaxed, EMSRb can actually outperform OBL.

The implementation of these seat inventory controls can be either static or dynamic. In a static environment, the booking limits are determined prior to any bookings and are based solely on initial forecasts. These booking limits are maintained throughout the entire booking process. In a dynamic environment, the booking limits are altered every few days. Once some of the bookings are realized, the demand is forecasted again and new booking limits are determined. Utilizing a dynamic inventory control scheme is standard practice at airlines today.

4.3.3 Network Effects in Revenue Management

From the above discussion and results from section 2.2.4, we note that leg-independent approaches might not be adequate when we consider multi-leg passengers. Farkas [21] overviews three different approaches in use at airlines around the world. We present these approaches in order of increasing levels of sophistication.
Fare Class Control

Based on amenities and restrictions, airlines segment the demand in a market into various fare classes designated by letter codes (Table 4.1). The computer reservation systems (CRS) are standardized such that each airline enters into the system the number of seats that are available in each fare class. The system assumes nesting, therefore, if there is a seat available in fare class H, then that seat can be filled by a passenger who buys a ticket in that fare class or higher (H, M, B, etc.). This seat is off limits to passengers who wish to buy a ticket in the Q or V fare classes. Because of these industry standards, it is required that a revenue management scheme segments the market into the standard CRS fare class designations.

<table>
<thead>
<tr>
<th>Fare Class</th>
<th>One-way Fare</th>
<th>Restrictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>$1,212.00</td>
<td>First class cabin, no stopovers, open return</td>
</tr>
<tr>
<td>Y</td>
<td>837.00</td>
<td>Coach cabin, no stopovers, open return</td>
</tr>
<tr>
<td>B</td>
<td>571.00</td>
<td>Flight legs must be confirmed in advanced, otherwise same as Y class</td>
</tr>
<tr>
<td>M</td>
<td>231.00</td>
<td>Non-refundable, fares only available certain times of day, plus all B class restrictions</td>
</tr>
<tr>
<td>H</td>
<td>159.00</td>
<td>Must be booked 7 days before departure, plus all M restrictions</td>
</tr>
<tr>
<td>Q</td>
<td>129.00</td>
<td>Must be booked 14 days before departure, plus all M class restrictions</td>
</tr>
<tr>
<td>V</td>
<td>107.00</td>
<td>Same as Q class, but more restrictive on times of day.</td>
</tr>
</tbody>
</table>

Table 4.1: The different restrictions for each fare class. The fare is for a one-way ticket from Boston (BOS) to Chicago (ORD) for March 16, 1998

To determine the demand (and thus the booking limits) on each flight leg in each fare class, the demand for all M fare class designations from all itineraries are aggregated together regardless of the actual itinerary fare. To calculate the contribution associated with each fare class, a weighted average of fare of the different itineraries is used.

In the fare class (FC) control system, a passenger is accepted if there is space in his specific fare class on every flight leg in his itinerary. This might lead to poor decisions in the following circumstances. We might turn away a multi-leg itinerary (DSM-ORD-LHR) passenger in the B class with a high total fare ($1,174) because there is no seat in B class on the short haul Des Moines - Chicago O'Hare leg. Instead, a local Y class passenger on DSM-ORD ($608) is accepted. In simulations, FC control can perform poorly because of this aggressive acceptance of higher fare class local passengers over multi-leg passengers in lower fare classes.
Value Class Control

In the value class (VC) control system, the standard reservation system fare class codes are ignored. An itinerary is put into a structure not based on the fare class but based on the system revenue or value. For example, itineraries with fares between $200 and $249 are put into one value class, fares between $250 and $299 into another and so forth.

To control the booking process on each flight leg, a weighted average fare is calculated for each value class. Then, the nested leg-independent booking limits are determined for each leg using an expected marginal revenue approach where the different value classes act as fare classes. When a booking request is received, the control system determines to which value class this request belongs and checks to see if there is available seating for this value class on each flight leg. If there is availability on all flight legs, the request is accepted. To work with standard CRS’s, these value classes are broken into standard fare class codes. It is possible that a Y class passenger might actually be assigned to a Y class on one leg of his itinerary and M class on another leg.

The value class control evaluates an itinerary based mostly on its associated system revenue, which can be an improvement over just considering fare class. VC control does not consider, however, the multiple units of capacity that multi-leg itineraries require. As the revenue ranges become smaller, the system tends toward a greedy approach, always selecting high fare connecting passengers. As flight leg demands become more balanced this system might perform worse than fare class control.

Bid Price Heuristics

A compromise between the fare class and the value class control systems, is a heuristic approach called bid pricing. This method follows the same procedure as the VC control system. When a request is received, it is mapped into a value class system. To account for the multi-leg nature of passenger itineraries, the request is mapped into a value class for each flight leg based on the fare minus the network value of the other flight legs in the itinerary. These network values for each flight leg are the displacement costs. Roughly, they represent a measure of the contribution that is lost by taking up that seat. Obviously, flight legs with a high demand compared to capacity will have high displacement costs.
4.3.4 Itinerary Blocking

Because of the small demand numbers in some markets (on the order of 1 or 2 passengers per month), we do not suggest the use of the passenger mix model to determine a partition of the flight leg cabin. Nevertheless, the results of the model might provide useful insights, including:

1. The itineraries that could be aggregated on specific flight legs.

2. The itineraries that should be blocked on specific flight legs. For example, upon review of the model results, we might notice that it is a good idea to divert LGA-LAX traffic away from the Chicago O’Hare hub towards the Denver hub.

4.3.5 Dual Variables as Bid Prices

The optimal dual variable for a flight leg represents the value of an extra seat on it. This can be used as a displacement cost in a virtual nesting revenue management system, or a basis for the bid price in a bid-price control revenue management system. In fact, both Williamson [47] and Talluri and van Ryzin [43] show that using a deterministic linear programming model for passenger flow (such as the passenger mix model) is one of the best methods to determine bid prices in network revenue management schemes.

4.3.6 Bounds for Revenue Management

As researchers develop more advanced heuristics for network revenue management, an important question is how much improvement can be attained by better approaches? In this section, we present some methods that can be used to determine an upper bound to the expected contribution achieved by an arbitrary seat inventory control scheme.
The Attainable Contribution Upper Bound

Consider the passenger mix problem without recapture and one cabin per flight.

\[
(PMIX - NR) \quad \begin{align*}
\text{maximize:} & \quad \sum_{(p,f) \in (P,F)} \text{fare}_{p,f} \cdot x_{p,f} \\
\text{subject to:} & \quad \sum_{(p,f) \in (P,F)} \delta_{i}^{p} x_{p,f} \leq CAP_{i} \quad \forall i \in L \\
& \quad 0 \leq x_{p,f} \leq D_{p,f} \quad \forall (p,f) \in (P,F)
\end{align*}
\]

(4.7)

where \( x_{p,f} \) is the number of passengers that travel on their desired IFC. We wish to see how the objective function of \( PMIX - NR \) changes as we vary \( D_{p,f} \). We denote the vector of \( D_{p,f} \) as \( \mathbf{d} \). Let \( Z_{PMIX}(\mathbf{d}) \) be the objective value of \( PMIX - NR \) with demand vector \( \mathbf{d} \). Let’s also define \( Z_{IC}(\mathbf{d}) \) as the contribution achieved using some arbitrary seat inventory control scheme, \( IC \). First, note that \( Z_{PMIX}(\mathbf{d}) \geq Z_{IC}(\mathbf{d}) \). This follows because the amount of potential contribution with perfect information from some specific demand instance, \( \mathbf{d} \), is greater than the realized contribution using a seat inventory control scheme, \( IC \), which has imperfect information.

Next, we propose the following lemma, which is a straightforward corollary of Theorem 5.1 in Bertsimas and Tsitsiklis [13]:

**Lemma 3** The function \( Z_{PMIX}(\mathbf{d}) \) is a concave function of \( \mathbf{d} \).

**Proof:** Let \( S \) be the set of demands such that the feasible region of \( (PMIX - NR) \) is non-empty and \( \mathbf{c} \) be the vector of fares in the objective function. Assume that the demand is independent and \( S \) is just the positive orthant, which is a convex set. Let \( \mathbf{d}_{1} \) and \( \mathbf{d}_{2} \) be two elements of this set. Let \( \mathbf{x}_{1} \) and \( \mathbf{x}_{2} \) be the corresponding optimal solution to \( (PMIX - NR) \) using demands \( \mathbf{d}_{1} \) and \( \mathbf{d}_{2} \), respectively. Thus \( Z_{PMIX}(\mathbf{d}_{1}) = \mathbf{c}^{\prime} \mathbf{x}_{1} \) and \( Z_{PMIX}(\mathbf{d}_{2}) = \mathbf{c}^{\prime} \mathbf{x}_{2} \). Let \( \lambda \) be a scalar in the closed interval \([0,1]\). Because of the convexity of \( S \), \( \tilde{\mathbf{d}} = \lambda \mathbf{d}_{1} + (1 - \lambda) \mathbf{d}_{2} \) is in the feasible region. The solution \( \tilde{\mathbf{x}} = \lambda \mathbf{x}_{1} + (1 - \lambda) \mathbf{x}_{2} \) is a feasible solution to \( (PMIX - NR) \) for the instance of demand \( \tilde{\mathbf{d}} \). Since the solution \( \tilde{\mathbf{x}} \) is feasible and we are max-
imizing, \( c'\tilde{x} \leq Z_{PMIX}(\tilde{d}) \). Yet, \( c'\tilde{x} = \lambda c'x_1 + (1 - \lambda) c'x_2 \), thus \( \lambda Z_{PMIX}(d_1) + (1 - \lambda) Z_{PMIX}(d_2) \leq Z_{PMIX}(\lambda d_1 + (1 - \lambda) d_2) \). Therefore, \( Z_{PMIX}(d) \) is a concave function of \( d \).

Actually, demand could be considered a random variable \( D \). First, let’s define the the following:

\[
E[Z_{PMIX}(D)] = \sum_d P(D = d)Z_{PMIX}(d).
\]

(4.8)

where \( P(D = d) \) is the probability that the random variable demand \( D \) is equal to the instance \( d \). We now have the following theorem:

**Theorem 4** The objective value of the problem \((PMIX - NR)\) using average demand is an upper bound to the expected contribution for any seat inventory control scheme. Mathematically,

\[
Z_{PMIX}(E[D]) \geq E[Z_{IC}(D)]
\]

(4.9)

**Proof:** Let’s take an arbitrary instance of demand, \( d \). It’s obvious that the passenger mix model gives the optimal seat allocation for this demand because of perfect information. Therefore, \( Z_{PMIX}(d) \geq Z_{IC}(d) \). Since this will be true for every instance of demand, \( E[Z_{PMIX}(D)] \geq E[Z_{IC}(D)] \). From Lemma 3, \( Z_{PMIX}(d) \) is a concave function of \( d \). Using Jensen’s inequality, \( Z_{PMIX}(E[D]) \geq E[Z_{PMIX}(D)] \). Combining the two sets of inequalities, we get \( Z_{PMIX}(E[D]) \geq E[Z_{PMIX}(D)] \geq E[Z_{IC}(D)] \).

From Theorem 4, we can quickly calculate an upper bound to the potential contribution by solving just one linear program. Notice that this is the case for any control scheme. This includes both static and dynamic controls.

**A tighter bound**

From the proof of the theorem above, we can obtain a tighter bound by solving a series of linear programs. We see that \( E[Z_{PMIX}(D)] \) is a tighter bound than \( Z_{PMIX}(E[D]) \). To solve for
\[ E[Z_{PMIX}(D)] \] exactly, however, we need to solve a huge number of linear programs. Using simulation, we can empirically determine an estimate to this bound.

**Bound tightness**

We wish to determine the tightness of the bounds that we have presented. To calculate \( E[Z_{IC}(D)] \), the expected contribution using an arbitrary seat inventory control, we use a Monte Carlo simulation procedure described in Farkas [21]. For seat inventory control we use the static fare-class based control, described in Section 4.3.3, that accepts a passenger if all legs of the itinerary have seats available for that fare class. For this experiment we use three(3) fare classes and assume that *all* passengers in a lower fare class book before *every* passenger in a higher fare class. The booking pattern is randomized within a fare class, i.e., each passenger is randomly chosen from all local and connecting itineraries. After the seat inventory control contribution is calculated, we calculate the optimal booking using the passenger mix problem. This process is repeated a prespecified number of times (namely 100). The mean of the runs for both schemes are used as estimates for \( E[Z_{PMIX}(D)] \) and \( E[Z_{IC}(D)] \). We calculate the mean squared error to determine a confidence interval for these estimates. For itinerary demand we use an integral gamma distribution that is described in Appendix B.

We wish to determine the tightness of the upper bounds as a function of four factors. For each of the four factors, we set three different values, described as (low, medium, high). These four factors are:

1. **Congestion**: To model congestion, we set the demand factor at (0.5, 0.75, 1.0).

2. **Network Effects**: We vary the passenger connectivity ratio (0.1, 0.25, 0.4).

3. **Fare Value Spread**: We set the fare value spread to these values (1.25, 1.5, 2.0) (See Section 3.7.4).

4. **Demand Uncertainty**: We set the z-factor of the demand to (1.0, 1.5, 2.0).

Tables 4.2 and 4.3 report results measuring the tightness of the bounds as a function of the above factors for problem data set F97-3A. We pair two factors and simultaneously vary them. In the first table, we vary the demand factor and passenger connectivity ratio, while keeping
the fare value spread and the Z-factor at their medium level. In the second table, we vary the fare value spread and the z-factor, while keeping the demand factor and passenger connectivity ratio at their medium levels.

<table>
<thead>
<tr>
<th>Demand Factor</th>
<th>PCR Low(.10)</th>
<th>PCR Middle(.25)</th>
<th>PCR High(.40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (.50)</td>
<td>(.995,.002)</td>
<td>(.987,.002)</td>
<td>(.985,.002)</td>
</tr>
<tr>
<td></td>
<td>(.100,.002)</td>
<td>(.996,.002)</td>
<td>(.999,.002)</td>
</tr>
<tr>
<td>Middle (.75)</td>
<td>(.977,.002)</td>
<td>(.978,.002)</td>
<td>(.970,.002)</td>
</tr>
<tr>
<td></td>
<td>(.997,.002)</td>
<td>(.997,.002)</td>
<td>(.996,.002)</td>
</tr>
<tr>
<td>High (.99)</td>
<td>(.954,.001)</td>
<td>(.961,.002)</td>
<td>(.960,.002)</td>
</tr>
<tr>
<td></td>
<td>(.997,.001)</td>
<td>(.995,.002)</td>
<td>(.998,.002)</td>
</tr>
</tbody>
</table>

Table 4.2: Tightness of bound as a function of passenger connectivity ratio (network effects) and the demand factor (congestion)

<table>
<thead>
<tr>
<th>Z-Factor</th>
<th>Fare Class Spread Low(1.25)</th>
<th>Fare Class Spread Middle(1.50)</th>
<th>Fare Class Spread High(2.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low (1.0)</td>
<td>(.979,.001)</td>
<td>(.983,.001)</td>
<td>(.989,.001)</td>
</tr>
<tr>
<td></td>
<td>(.998,.001)</td>
<td>(.999,.001)</td>
<td>(.999,.001)</td>
</tr>
<tr>
<td>Middle (1.5)</td>
<td>(.978,.002)</td>
<td>(.978,.002)</td>
<td>(.986,.002)</td>
</tr>
<tr>
<td></td>
<td>(.100,.002)</td>
<td>(.997,.002)</td>
<td>(.1002,.002)</td>
</tr>
<tr>
<td>High (2.0)</td>
<td>(.972,.002)</td>
<td>(.974,.002)</td>
<td>(.975,.002)</td>
</tr>
<tr>
<td></td>
<td>(.998,.002)</td>
<td>(.998,.002)</td>
<td>(.997,.002)</td>
</tr>
</tbody>
</table>

Table 4.3: Tightness of bound as a function of the fare class spread (segmentation) and the Z-factor (demand variance)

For each cell in the table, the contribution associated with the average demand passenger mix model $Z_{PMIX} (E[D])$ is normalized to 1.0. The first row contains the estimate of $E[Z_{FC}(D)]$, (the expected contribution using the static FC control), and the mean squared error of the estimate. The second is an estimate of the tighter bound of $E[Z_{PMIX}(D)]$ and the mean square error of this estimate. To determine a 95% confidence interval, we multiply the mean square error by approximately 2.0 (the student t value from statistics). This serves as the half-width of the confidence interval. For example, the 95% confidence interval for the true mean with the low demand factor and low passengers connectivity ratio is between .991 and .999.

There are some observations we can make from the tables:

- The difference between the bound by solving the average demand passenger mix model
and the series of passenger mix models with different demand is less than 0.5%. This doesn't seem to change much as we vary the different parameters.

- The level of the demand factor significantly affects the gap between $E[Z_{FC}(D)]$ and $Z_{PMIX}(E[D])$. As we increase the congestion in the system, the size of the gap increases.

- There are some trends with the other factors, though these do not seem to be as significant as congestion.

Actually, there are some issues that lead us to believe that the bounds might even be tighter than what is described above. First, we simulated a fare class seat inventory control scheme. As was highlighted in Section 4.3.3, this has a tendency to be outperformed consistently by a bid price heuristic which is in place at many major airlines. Second, we use a gamma distribution for each $IFC$ that has a standard deviation equal to the Z-factor times the square root of the mean. Since the gamma distribution is a continuous distribution, the value is rounded to an integer (see Appendix B). This discretized gamma distribution has a larger variance than the original continuous distribution. For distributions with large means, this change in the standard deviation is almost negligible; for distributions with small means, the increased standard deviation could be on the order of 10 to 15%. The fare class buckets were optimized with respect to the standard deviation of the continuous distribution (which was known) and not the discretized distribution. Finally, we use a static inventory control scheme. By using a dynamic control scheme, the seat inventory control scheme might perform better and thus the bound would be tighter.

### 4.4 Fleet Assignment

#### 4.4.1 Attainable Contribution Problem

The main reason for our development of the passenger mix model is to solve the attainable contribution problem for the iterative fleet assignment process. We can use the passenger mix model to determine better estimates of spill costs in lieu of the approximations of the spill models currently used. Using the passenger mix model, we have a basis for comparison of different fleetings. Also, by using information from the passenger mix model (such as dual
costs), we can gain insights about the relationship between passenger itineraries and the fleet assignment model and thus determine ways that we can improve the traditional fleet assignment model. We further explore these areas in Chapter 5.

### 4.4.2 A Combined Fleet Assignment Model

The passenger mix formulation can be combined with a traditional fleet assignment model to create an itinerary based fleeting model [21]. In traditional fleet assignment models, the spill calculation for each fleet-flight leg decision variable is calculated using the forecasted demand for travel service on that specific flight leg independent of other flight legs in the network. When spill does occur, no attempt is made to determine which passengers will be spilled and how this might affect the demand on other flight legs in the network. For example, consider a flight leg from Honolulu (HNL) to LAX. For many passengers, the ultimate destination is not Los Angeles, but many destinations throughout the continental United States. If spill occurs on the HNL-LAX flight leg, numerous other flight legs (LAX-BOS, LAX-ORD, etc.) are affected and the resulting spill calculations for those flight legs will be inaccurate.

These two models can be combined using the capacity constraint in the passenger mix model. This constraint is modified by replacing the capacity value on the right hand side of the equation with a sum in the constraint matrix as follows:

\[
\sum_k SEATS_{k,j} X_{k,i} + \sum_{(p,f) \in (P,F)} \sum_{(r,g) \in (P,F)} \delta_{i,p,f}^{r,g} - \sum_{(p,f) \in (P,F)} \sum_{(r,g) \in (P,F)} \delta_{i,p,f}^{r,g} t_{r,g}^{i,p,f} \ge Q_{i,j}
\]  

(4.10)

where \(SEATS_{k,j}\) is the number of seats in cabin \(j\) for fleet type \(k\) and \(X_{k,i}\) is a binary decision variable equal to one when fleet type \(k\) is assigned to flight leg \(i\) and 0 otherwise.

### 4.5 Conclusions

In this chapter, we present a number of applications of the passenger mix model. It can be used in an operational setting when passengers must be rerouted because of significant delays, flight leg re-fleetings, and cancellations caused by severe weather and equipment failure.
The passenger mix model can also be used tactically in seat inventory control for revenue management systems. It also provides an upper bound to the expected contribution that can be attained. We show that the upper bound calculated by using average demand in the passenger mix model is within 0.5% of the upper bound calculated by averaging the passenger mix model using many different instances of demand. Also, we show that the quality of these upper bounds decrease as we increase the congestion in the system.
Chapter 5

Improving the Traditional Fleet Assignment Model

5.1 Overview

The goal of this thesis is to provide an analysis of various solution techniques of the airline fleet assignment problem with respect to the network effects of multi-leg passenger itineraries. Before we create a new large model that optimizes the fleet assignment and passenger mix decisions simultaneously, we consider the current state of affairs in the airline industry. Airlines have spent large amounts of money developing their fleet assignment solvers over the last decade. Therefore, a procedure that determines better near-optimal fleeting decisions quickly using the current fleet assignment solvers as an integral part of the process has advantages (initially anyway) over an entirely new formulation that requires major re-engineering. In Chapter 2, we present a new paradigm for the fleet assignment model that includes current solvers as an ingredient of the solution procedure. One of the crucial components of this new paradigm is a model that calculates the attainable contribution (the potential revenue minus the spill and carrying costs). Exactly calculating the attainable contribution is extremely difficult. In the current fleet assignment model, the attainable contribution model is a rough approximation that does not consider the network effects of multi-leg passenger itineraries. In Chapter 3, we present a linear programming model that approximates the attainable contribution with network effects. We prove in Chapter 4 that using this formulation with average demand
provides an upper bound on the expected contribution.

In this chapter, we describe, in detail, the current state of the practice in approximating the attainable contribution for the fleet assignment problem. We also describe some simple modeling rules that better estimate the attainable contribution and result in better fleetings than those achieved by the current state of the practice. We conclude this chapter with some advanced iterative modeling methods for the fleet assignment model that will help drive the fleet assignment model toward better fleetings.

Unless otherwise noted, we use the following assumptions in both Chapters 5 and 6:

1. We do not consider recapture explicitly in the fleet assignment model. This allows us to isolate the network effects solely associated with the effects of spill on other dependent flight legs. However, most of the models presented can model the effects of recapture.

2. To further isolate the network effects, we solve the fleet assignment problem for the specific case when passenger demand is deterministic. We discuss how to generalize the following analyses for the case when demand is uncertain.

3. We assume that there is only one fare class. In the linear programming passenger mix model (Chapter 3) that we use to measure spill, the inclusion of fare classes is modeled as identical itineraries with different fares.

This research is one of the first to understand how multi-leg itineraries affect the fleet assignment problem. With the above assumptions in place, we are determining what is the optimal fleet assignment given deterministic demand and no recapture. In this chapter, we gain insights that will help us develop procedures that are general enough to be applied to the real problem of fleet assignment given multiple fare classes, uncertain demand, and recapture.

### 5.2 The Current State of the Practice

In Chapter 2, we outline the mechanics of solving the fleet assignment problem, as well as present some of its analytical properties. Using an illustrative example, we touch on the ambiguity involved in modeling the attainable contribution. In this section, we describe the most common attainable contribution modeling methods in fleet assignment currently employed at the major
airlines. We then compare and contrast the relative advantages and disadvantages of each method. However, to do this, we first develop a framework that allows us to compare different methods consistently.

5.2.1 A Sequential Framework

The goal of the fleet assignment model is to assign fleet types to each flight leg in accordance with the physical and operational constraints, while maximizing the contribution to the airline. In the illustrative example in Chapter 2, we show that a naive implementation of basic spill models might not result in an optimal assignment, since it is impossible to model exactly the network effects of passenger requests using costs that are linear with respect to the traditional fleet decision variables. Therefore, objective function values of the traditional fleet assignment models are only approximations. It is likely that two different approximation approaches of the contribution for the fleet assignment model might result in the same fleeting, but have different objective values. Once a fleeting is determined, we employ a comparison procedure that determines the attainable contribution as a function solely of the fleeting decision, regardless of the method that was used to obtain that fleeting.

Throughout this chapter, we implement a framework that follows this criterion (Figure 5-1). First, we use some method to determine a feasible fleeting, which has an associated operating cost (the right branch in Figure 5-1) regardless of the number of passengers in the system. Second, we solve for the attainable contribution given the fleeting determined by the solver (the left branch in Figure 5-1). By taking the optimal objective value of the attainable contribution problem and subtracting the operation costs determined in the first stage, we have the fleeting contribution for that specific fleeting. Let us denote the contribution by the variable Z. We use this value to compare the different approaches for the fleet assignment problem.

In Chapters 5 and 6, the values that are reported have units of dollars per day. For emphasis, we sometimes express this value in terms of an estimate of dollars per annum, simply multiplying the dollars per day by 365. We use the assumption that each day is identical.
Figure 5-1: The sequential framework to compare different fleet assignment methods

5.2.2 The Current Methodology

The Representative Fare Spill Model

We use the same notation that was defined in Chapter 3. Let us denote $t_p$ as the number of spilled passengers from itinerary $p$. In the deterministic case, the contribution can be found by the following formula:

$$Z = \sum_{p \in P} \text{fare}_p \cdot D_p - \sum_{p \in P} \text{fare}_p \cdot t_p - \sum_{k \in K} \sum_{i \in L} c_{k,i} X_{k,i}, \quad (5.1)$$

where the first sum represents the unconstrained revenue and the last sum represents the operating costs of a specific fleetting $\mathbf{X}$. Since the first term is a constant, finding the minimum of the spilled revenue plus operating cost is sufficient. Traditional fleet assignment solvers ap-
proximate the spilled revenue assuming it is linear with respect to the fleet assignment decision variables, i.e.,

\[ \sum_{p \in P} fare_p \cdot t_p \approx \sum_{i \in L} \sum_{k \in K} c_{k,i}^{sp} X_{k,i}, \quad (5.2) \]

where \( c_{k,i}^{sp} \) is an approximation of the spill costs given that fleet type \( k \) is assigned to flight leg \( i \). To approximate these spill costs, the standard fleet assignment solvers implemented in industry use a basic spill model in which a representative spill fare, \( fare_i^{sp} \), for each leg is determined. This spill fare is multiplied by the number of spilled passengers, on flight leg \( i \) to determine the spilled revenue on each flight leg,

\[ c_{k,i}^{sp} = fare_i^{sp} \cdot \max \left\{ 0, \sum_{p \in P} t_i^p D_p - SEATS_k \right\}, \quad (5.3) \]

where \( SEATS_k \) is the number of seats on an aircraft of fleet type \( k \). We define the model used in Equation 5.3 as the representative fare spill model (RFSM). Typically, the spill fare for flight leg \( i \) is closely related to the average fare for flight leg \( i \), \( fare_i \). This relationship is usually linear with the average fare being multiplied by some constant, \( \kappa \):

\[ fare_i^{sp} = \kappa \cdot fare_i. \quad (5.4) \]

Note that the constant, \( \kappa \), is universal; it is not dependent on either the flight leg or the fleet type assigned to the flight leg. One might assume \( \kappa = 1 \), however, since the airlines partially control passenger bookings with revenue management systems, the airline might deny the lower fare passengers on flights that are filling up during the booking process in order to increase contribution. As a result, the average fare of spilled passengers on a specific flight leg is usually less than the average fare of all passengers desiring travel on that flight leg, and \( \kappa < 1 \).

**Two Standard Fare Allocation Schemes**

The question of determining the average fare for a specific flight leg isn't straightforward. In Chapter 2, two different methods for calculating the average fare are briefly described. We
denote the first as the full fare (FF) allocation scheme. The full fare allocation scheme assigns all of the fare to each leg of the itinerary. The average fare on flight leg $i$ is:

$$\bar{\text{fare}}_{FF} = \frac{\sum_{p \in P} \text{fare}_p \delta^p_i D_p}{\sum_{p \in P} \delta^p_i D_p}. \quad (5.5)$$

The second method we will denote as the mileage-based pro-rated fare (MF) allocation scheme. In the MF scheme, the fare is allocated among all legs of the itinerary based on the costs of carrying the passenger on each flight leg. As an approximation for the relative costs of transporting the passenger, airlines use the flying miles for each flight leg. The average fare on leg $i$ using the MF allocation scheme is therefore:

$$\bar{\text{fare}}_{MF} = \frac{\sum_{p \in P} \frac{\text{mile}_i \cdot \text{fare}_p \delta^p_i}{m_p} D_p}{\sum_{p \in P} \delta^p_i D_p}, \quad (5.6)$$

where $\text{mile}_i$ is the flying miles of flight leg $i$, and $m_p = \sum_{i \in L} \delta^p_i \text{mile}_i$.

Whether the full fare or the mileage-based pro-rated fare performs better is subject to the nature of the passenger flows and the structure of the flight network. In Section 5.2.4, we analyze the potential advantages and disadvantages of both methods and show empirically which performs better for a number of test problems extracted from two demand scenarios with different flight networks. Before we do that, let's focus on the one free parameter in these schemes, the spill fare to average fare ratio, $\kappa$.

The Lack of Quasiconcavity of the Contribution Function

Let $Z_S(\kappa)$ be the contribution when we use the value $\kappa$ in Equation 5.4 with fare allocation scheme $S$ for the sequential fleet assignment, passenger mix framework. Our best case for allocation scheme $S$, $Z^*_S$, is defined as

$$Z^*_S = \max_{\kappa \geq 0} Z_S(\kappa). \quad (5.7)$$

The ease of exactly finding $Z^*_S$ depends on whether or not $Z_S(\kappa)$ is a quasiconcave function.
in \( \kappa \). Let \( f : E \to \mathbb{R} \), where \( E \) is a non-empty convex set. The function \( f \) is said to be quasiconcave if for each \( x_1 \) and \( x_2 \in E \), the following inequality is true [5]:

\[
f [\mu x_1 + (1 - \mu) x_2] \geq \text{minimum} \{ f(x_1), f(x_2) \} \quad \forall \mu \in (0, 1).
\] (5.8)

Another way to view quasiconcavity is with level sets. The level sets of a function \( f \) are defined as follows:

\[
E_\alpha = \{ x \in E : f(x) \geq \alpha \}.
\] (5.9)

If \( f \) is a quasiconcave function, then \( E_\alpha \) is a convex set (Bazaraa et al. [5]). Ideally, we would like to have \( Z_S(\kappa) \) be a strictly quasiconcave function, which is defined the same as a quasiconcave function except that a strict inequality is used in Equation 5.8 unless both \( x_1 \) and \( x_2 \) are both optimal solution points. This eliminates any functions that have "flat spots" except at the maximum. If a function is strictly quasiconcave, then we can find the global maximum of the function by performing a simple line search. We can rule out the possibility that \( Z_S(\kappa) \) will be strictly quasiconcave. If two different values \( \kappa_1 \) and \( \kappa_2 \) in the fare calculation give the same fleeting \( \mathbf{X} \), then this will result in the same spill values in the attainable contribution module. Therefore, these values will result in the same contribution. However, if these contributions are not the global maximum, then the function is not strictly quasiconcave.

Figure 5-2 shows the behavior of \( Z_{FF}(\kappa) \) and \( Z_{MF}(\kappa) \) as we vary \( \kappa \) for problem data set A97-4A. The first thing we notice is that these functions are not quasiconcave. For example, the highlighted area around \( \kappa = 0.68 \) results in a lower contribution as compared to the areas on either side. This means that finding \( Z_S^* \) exactly will be extremely difficult if not impossible. When we develop our ideas of an iterative framework, we describe methods to calculate \( \tilde{Z}_S^* \), an estimate of \( Z_S^* \).

### 5.2.3 Spill Estimation

The shape of the graph offers us another insight. As \( \kappa \to 0 \), we notice that the quality of the solution begins to deteriorate rapidly. However, as we begin to increase \( \kappa \) beyond the best value, the contribution function seems to flatten out. Therefore, if we err in our determination
Figure 5-2: The contribution as a function of the spill fare to average fare ratio

of the spill fare to average fare ratio, it is better to overestimate, than underestimate.

Figures 5-3 and 5-4 present a comparison of the estimated spill using a basic spill model and
the more sophisticated passenger mix model. The reasoning behind why low values of $\kappa$ are so
detrimental is straightforward. On the one hand, low $\kappa$ causes us to underestimate the cost of
spill and hence total system cost. On the other hand, low $\kappa$ leads us to use smaller aircraft,
increasing actual spill and the ultimate objective value.

Using complementary arguments, we understand why overestimating spill fares is not as
damaging. If we overestimate spill fares, the fleet assignment model is more likely to utilize the
more expensive, higher capacity aircraft. With more capacity in the system, this results in fewer
spilled passengers. With fewer spilled passengers, this results in less spill that is inaccurately
estimated, therefore our errors are less.

If we combine the two graphs above, we notice something interesting. Figure 5-5 is a graph
where the quality of the solution (as measured by the difference between a specific solution
and the best solution found for all fare allocation schemes in question) is plotted against the
Figure 5-3: A comparison of the estimated spill and the actual spill when using the full fare allocation scheme

estimated spill inaccuracy, which we denote by $SI$. The quality of the solution is found by:

$$Z_S(k) = \min \limits_S \left[ \tilde{Z}^*_S \right]. \quad (5.10)$$

The spill inaccuracy for a fleeting $\mathbf{X}$ is defined as the estimated spill in the fleet assignment model minus the actual spill calculated by the attainable contribution solver:

$$SI(\mathbf{X}) = \left( \sum \limits_{i \in L} \sum \limits_{k \in K} c_{k,i}^{\mathbf{X}} \mathbf{X}_{k,i} \right) - \left( \sum \limits_{p \in P} fare_p \cdot t_p \right). \quad (5.11)$$

Our intuition tells us that a better estimate of the spill results in a better fleeting decision. We expect to see a V-shaped plot with the vertex at the origin, i.e., if we exactly estimate spill, we expect to obtain the optimal solution. However, in Figure 5-5 we see that this is not the case; the vertex is off-center. There are a couple of explanations for this.
Figure 5-4: A comparison of the estimated spill and the actual spill when using the mileage-based fare allocation scheme

- For one flight leg, we may overestimate spill for all fleet type assignments, while for another flight leg, we underestimate spill. These inaccuracies cause a bad fleeting decision: however, the total spill value is equal to the actual spill value.

- While we have an accurate estimate of spill for a given fleeting, we might be inaccurately estimating spill on alternative fleeting.

5.2.4 Mileage-Based Pro-Rated Fare Versus Full Fare Allocation

Of the two allocation schemes presented, let's compare them to see which performs better empirically. To solve for $\hat{Z}_S^*$, an estimate to $Z_S^*$, we use the following procedure:

$$\hat{Z}_S^* = \max \{Z_S(0.00), Z_S(0.02), Z_S(0.04), \ldots, Z_S(2.00)\},$$  \hspace{1cm} (5.12)
Figure 5-5: The quality of the contribution as a function of spill inaccuracy

i.e., we calculate the contribution for a number of different values of \( \kappa \) and choose the maximum, which we denote \( \kappa^*_S \). Table 5.1 contains the results. Included in the table for both allocation schemes is the maximum contribution found, the optimal spill fare to average fare ratio (or a range), and the spill inaccuracy, \( SI^* \), when using \( \kappa^*_S \). If there is a range of \( \kappa^*_S \) that yields the same \( Z^*_F \), then the minimum inaccuracy of these solutions is presented in the table. Bold values are the maximum contribution for each problem.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>( Z^*_F ) ($/day)</th>
<th>( \kappa^*_F )</th>
<th>( SI^* )</th>
<th>( Z^*_M ) ($/day)</th>
<th>( \kappa^*_M )</th>
<th>( SI^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>3,325,628</td>
<td>1.18-1.38</td>
<td>304,244</td>
<td>3,325,628</td>
<td>1.18-1.38</td>
<td>231,445</td>
</tr>
<tr>
<td>F97-4A</td>
<td>5,141,331</td>
<td>1.16-1.20</td>
<td>329,018</td>
<td>5,136,779</td>
<td>1.36</td>
<td>678,739</td>
</tr>
<tr>
<td>F97-6A</td>
<td>7,201,066</td>
<td>1.08</td>
<td>126,442</td>
<td>7,176,229</td>
<td>.88</td>
<td>-678,805</td>
</tr>
<tr>
<td>F97-9</td>
<td>8,201,192</td>
<td>.98</td>
<td>-426,465</td>
<td>8,199,869</td>
<td>1.12</td>
<td>-340,669</td>
</tr>
<tr>
<td>A97-3A</td>
<td>1,448,092</td>
<td>.54-70</td>
<td>-13,355</td>
<td>1,448,092</td>
<td>.54-70</td>
<td>-12,500</td>
</tr>
<tr>
<td>A97-4A</td>
<td>4,901,319</td>
<td>.68-72</td>
<td>-11,652</td>
<td>4,901,197</td>
<td>.64-66</td>
<td>-108,672</td>
</tr>
<tr>
<td>A97-6A</td>
<td>8,012,664</td>
<td>.92</td>
<td>117,021</td>
<td>8,014,395</td>
<td>.92</td>
<td>-54,478</td>
</tr>
<tr>
<td>A97-9</td>
<td>11,208,105</td>
<td>.84</td>
<td>-823,272</td>
<td>11,155,890</td>
<td>1.38</td>
<td>837,843</td>
</tr>
</tbody>
</table>

Table 5.1: A comparison of the full fare and partial fare allocation schemes

We can draw a couple of observations from the results. For this set of problems, the \( FF \) allocation scheme seems to perform better than the \( MF \) allocation scheme. In the two full
size problems (F97-9 and A97-9), the FF allocation scheme outperformed the MF allocation scheme by $1,323/day ($483K/year) and $52,215/day ($19M/year), respectively.

However, Table 5.1 raises some concerns when it comes to a solution procedure. For the two larger problems (6A and 9) of each scenario, we see that there is a significant range of optimal k’s. For the FF allocation scheme, the range is from .84 to 1.08, while the range of the MF allocation scheme is .88 to 1.38. This range leads us to the inconsistency in the spill inaccuracy values.

To get the best solution for the A97-9 data set using the FF allocation scheme, we underestimated the amount of spill by over $800K/day. However, for the same problem, to get the best solution using the MF allocation approach, we needed to overestimate the spill by $800K/day. On the other hand, for data set F97-6A, we overestimate spill by $125K/day and underestimate spill by $678K/day for the FF and MF schemes, respectively. There isn’t any consistency even for each scheme. We highlight two points:

1. Accurately determining system-wide spill for the optimal solution does not guarantee good fleeting decisions; it must be done for all fleeting decisions.

2. While the best fleeting decisions for each allocation scheme may be adequate, the current methods are erratic in determining how to find the best fleeting decision. There cannot be much faith in a fleeting decision when the anticipated spill is so vastly different than the spill that will actually occur.

The quality of solution depends strongly on the fare allocation methodology, but it is not clear how to calculate it. We therefore consider it imperative to generate a higher fidelity model of spill costs which we explore in the next section.

5.3 Simple Improvements to the State of the Practice

In the previous section, we highlight some of the problems of the current model, specifically in the area of spill estimation. Some of these problems might be corrected by viewing the problem differently. This section offers simple methods to improve the current model, requiring little effort in changing existing fleet assignment codes. We conclude this section with some results that give an idea of their effectiveness.
In Section 5.2.2, we present two different schemes for allocating fare to flight legs: full fare and mileage-based pro-rated fare. If most passenger itineraries have only one flight leg in which passengers will be spilled, then we expect the full fare scheme to work better. However, if there are a significant number of passengers whose itineraries have multiple flight legs in which passengers can be spilled, then a full-fare scheme may lead to double-counting of lost revenue. In this section, we quantify this notion.

The use of the mileage-based pro-rated scheme is an outgrowth of airline accounting procedures. To determine the profitability of a flight leg, an airline will divide the fare of a multi-leg itinerary based on the costs of transporting a passenger on each leg. There are a couple of ways to approximate the relative costs. The method described above uses the flight mileage to compare the different legs. Another possibility is to use the flying time of each flight leg to divide the fare. While this allocation scheme seems logical for accounting purposes, it is illogical for a fleet assignment model. The main determination on whether or not revenue is realized is based on capacity, not costs.

### 5.3.1 Equal Fare Allocation

Since available space on every flight leg of a passenger's itinerary is essential in order to realize that demand, a natural instinct is to divide the fare equally among all flight legs in the itinerary. We will denote this allocation scheme as the *equal fare (EF) allocation scheme*. In this sense we are modeling the process as if we have one passenger on each flight leg, paying a partial fare, so it is similar to the mileage-based pro-rated fare scheme. However, we treat each flight leg equally instead of giving weight to the longer flight legs. The calculation for the average fare on each flight leg becomes

$$
\text{fare}^{EF}_{i} = \frac{\sum_{p \in P} \text{fare}_{\text{legs}_p} \delta^p_i D_p}{\sum_{p \in P} \delta^p_i D_p},
$$

where \( \text{legs}_p = \sum_{i \in L} \delta^p_i \).

Figure 5-6 shows how the equal fare allocation performed as we varied the spill fare to
Figure 5-6: A comparison of the equal fare and the full fare allocation scheme

average fare ratio for data set A97-4A. Table 5.2 shows how the equal fare allocation scheme performs compared to the full fare and the cost-based pro-rated schemes.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>$Z^*_S$ ($/day$)</th>
<th>Best Method</th>
<th>$\kappa^*_S$</th>
<th>$Z^*_{EF}$ ($/day$)</th>
<th>$\kappa^*_{EF}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>3,325,628</td>
<td>MF,FF</td>
<td>1.18-1.38</td>
<td>3,325,628</td>
<td>1.60</td>
</tr>
<tr>
<td>F97-4A</td>
<td>5,141,331</td>
<td>FF</td>
<td>1.16</td>
<td>5,117,606</td>
<td>.90-1.14</td>
</tr>
<tr>
<td>F97-6A</td>
<td>7,201,066</td>
<td>FF</td>
<td>1.08</td>
<td>7,208,352</td>
<td>1.72</td>
</tr>
<tr>
<td>F97-9</td>
<td>8,201,192</td>
<td>FF</td>
<td>.98</td>
<td>8,204,675</td>
<td>1.02</td>
</tr>
<tr>
<td>A97-3A</td>
<td>1,448,092</td>
<td>MF,FF</td>
<td>.54-.70</td>
<td>1,448,159</td>
<td>1.08</td>
</tr>
<tr>
<td>A97-4A</td>
<td>4,901,319</td>
<td>FF</td>
<td>.72</td>
<td>4,904,975</td>
<td>1.14-1.18</td>
</tr>
<tr>
<td>A97-6A</td>
<td>8,014,395</td>
<td>MF</td>
<td>.92</td>
<td>8,037,422</td>
<td>1.24</td>
</tr>
<tr>
<td>A97-9</td>
<td>11,208,105</td>
<td>FF</td>
<td>.84</td>
<td>11,227,126</td>
<td>1.20</td>
</tr>
</tbody>
</table>

Table 5.2: A comparison of the equal fare allocation scheme compared to the methods of full fare and cost-based pro-rated fare allocation

In Figure 5-2, the graph of the performance of the $MF$ and the $FF$ allocation schemes were almost identical. However, in Figure 5-6, the graph of the $EF$ scheme is different than that of the $FF$ scheme. Actually, the shape of the $EF$ scheme is similar to a stretched out version of the $FF$ scheme. An important outgrowth of this similarity is the width of the plateau of good solutions. For this problem, when using the $FF$ allocation scheme, using a spill to average fare
ratio between 0.6 and 0.8 results in a series of solutions that are all close to the best solution in value. However, when using the EF scheme, the range of good solutions is from 0.8 to 1.4. Since it is not possible to exactly get the best \( \kappa \) value in every instance, having a larger near-optimal solution range is beneficial.

For almost the entire set of problems, the EF scheme outperforms both the MF and the FF scheme (Table 5.2). The EF scheme increases contributions over the FF scheme by $3,483/day ($1.3M/year) and $19,021/day ($6.9M/year) for the F97-9 and the A97-9 data sets, respectively. Compared to the commonly used MF scheme, the EF scheme provides savings of $4,806/day ($1.8M/year) and $71,236/day ($26.0M/year).

5.3.2 A Categorization of Passenger Itineraries

Before we investigate other techniques further, it is helpful to analyze systematically the nature of passenger itineraries. First, let us present some necessary definitions:

\( K(i) \): the set of fleet types that may be assigned to flight leg \( i \). This will be a subset of all fleet types \( K \). The two sets might not be equal due to noise restrictions, flying range restrictions, etc.

\( CAP_i \): the minimum feasible capacity of flight leg \( i \).

\( \overline{CAP}_i \): the maximum feasible capacity of flight leg \( i \).

The minimum feasible capacity, \( CAP_i \), can be computed by

\[
CAP_i = \min_{k \in K(i)} \{ SEATS_k \},
\]

while the maximum feasible capacity is found similarly,

\[
\overline{CAP}_i = \max_{k \in K(i)} \{ SEATS_k \}.
\]

The unconstrained demand on a flight leg is

\[
Q_i = \sum_{p \in P} \delta^p_i D_p.
\]
We define a flight leg as *uncapacitated* if $Q_i \leq \text{CAP}_i$. We define a flight leg as *capacitated* if $Q_i > \text{CAP}_i$. Sometimes, we may want to further classify capacitated flight legs. If $\text{CAP}_i < Q_i \leq \overline{\text{CAP}}_i$, then we define flight leg $i$ as *potentially capacitated*, since the likelihood of spilling passengers is dependent on the fleet assignment. If $Q_i > \overline{\text{CAP}}_i$, then flight leg $i$ is *overcapacitated*, since passengers will be spilled, regardless of the fleet assignment. Table 5.3 shows the number of flight legs in each classification for the data sets that we are using in this thesis.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Total # of Flight Legs</th>
<th>Uncapacitated Num.</th>
<th>Uncapacitated %</th>
<th>Pot. Capacitated Num.</th>
<th>Pot. Capacitated %</th>
<th>Overcapacitated Num.</th>
<th>Overcapacitated %</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>157</td>
<td>120</td>
<td>76.4</td>
<td>20</td>
<td>12.7</td>
<td>17</td>
<td>10.8</td>
</tr>
<tr>
<td>F97-4A</td>
<td>431</td>
<td>339</td>
<td>78.7</td>
<td>74</td>
<td>17.2</td>
<td>18</td>
<td>4.2</td>
</tr>
<tr>
<td>F97-6A</td>
<td>823</td>
<td>655</td>
<td>79.6</td>
<td>145</td>
<td>17.6</td>
<td>23</td>
<td>2.8</td>
</tr>
<tr>
<td>F97-9</td>
<td>2,044</td>
<td>1,730</td>
<td>84.6</td>
<td>257</td>
<td>12.5</td>
<td>57</td>
<td>2.8</td>
</tr>
<tr>
<td>A97-3A</td>
<td>173</td>
<td>98</td>
<td>56.6</td>
<td>72</td>
<td>41.6</td>
<td>3</td>
<td>1.8</td>
</tr>
<tr>
<td>A97-4A</td>
<td>485</td>
<td>349</td>
<td>72.0</td>
<td>128</td>
<td>26.4</td>
<td>8</td>
<td>1.6</td>
</tr>
<tr>
<td>A97-6A</td>
<td>877</td>
<td>603</td>
<td>68.7</td>
<td>254</td>
<td>29.0</td>
<td>20</td>
<td>2.3</td>
</tr>
<tr>
<td>A97-9</td>
<td>1,888</td>
<td>1,366</td>
<td>72.4</td>
<td>453</td>
<td>24.0</td>
<td>69</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Table 5.3: A classification of flight legs based on the passenger demand and potential capacity

Let us focus on the full data set for August 1997 (A97-9). We see that almost 75% of the flight legs are uncapacitated, meaning that the airlines have little problem accommodating passengers on these flight legs. Roughly a quarter of the flight legs might have spill on them. Less than 5% of flight legs are overcapacitated and will always spill passengers.

While the above classifications of flight legs is interesting, for the sake of determining spill, we are more interested in how the passengers "flow through" the flight network. Given these
definitions, we are now able to classify passenger demand. We will use the following taxonomy:

<table>
<thead>
<tr>
<th>Type 0:</th>
<th>All flight legs in the itinerary are uncapacitated.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1:</td>
<td>Exactly one flight leg in the itinerary is capacitated.</td>
</tr>
<tr>
<td>Type 2:</td>
<td>Itinerary has more than one potentially capacitated flight leg, but no overcapacitated legs.</td>
</tr>
<tr>
<td>Type 3:</td>
<td>In the itinerary, exactly one leg is overcapacitated and at least one other flight leg is potentially capacitated.</td>
</tr>
<tr>
<td>Type 4:</td>
<td>Two or more flight legs in the itinerary are overcapacitated.</td>
</tr>
</tbody>
</table>

The reasons for this mutually exclusive and collectively exhaustive classification of passengers will be apparent after the discussion in Section 5.3.4.

Table 5.4 shows the classification of passenger demand for the full-sized data sets of the two different scenarios. For each data set, we categorize each passenger into one of the passenger demand types, then calculate the percentages in each category based on the number of passengers and the revenue.

<table>
<thead>
<tr>
<th></th>
<th>F97-9</th>
<th>A97-9</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Passengers</td>
<td>Revenue</td>
</tr>
<tr>
<td></td>
<td>Number</td>
<td>%</td>
</tr>
<tr>
<td>Type 0</td>
<td>93,204</td>
<td>65.1</td>
</tr>
<tr>
<td>Type 1</td>
<td>47,277</td>
<td>33.0</td>
</tr>
<tr>
<td>Type 2</td>
<td>1,355</td>
<td>1.0</td>
</tr>
<tr>
<td>Type 3</td>
<td>757</td>
<td>0.5</td>
</tr>
<tr>
<td>Type 4</td>
<td>520</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Table 5.4: A classification of passenger demand

In Table 5.4, we see that most passengers are either Type 0 or Type 1. For the fleet assignment problem without recapture, we do not need to consider Type 0 passengers, since all of their flight legs are uncapacitated and are not in jeopardy of being spilled, i.e., these passengers are traveling on uncapacitated flight legs therefore the spill costs for these flight legs are always zero no matter which fleet type is assigned. If a passenger is Type 1, then we are only concerned about one flight leg on that passenger's itinerary, the one which is capacitated. If we spill this passenger, it will not affect the spill costs of any other flight leg. Therefore,
in the case of no recapture, Type 0 and Type 1 passengers will not have any network effects associated with them. This may not be the case when we consider recapture. As passengers are spilled from other itineraries, uncapacitated flight legs may be the flight legs that are used to recapture passengers. Therefore, the "residual" capacity of uncapacitated flight legs determine the amount of passengers that can be recaptured.

The network effects associated with the fleet assignment model are caused by Type 2, 3, and 4 passengers, which represent less than 5% in terms of passengers and potential revenue. Therefore, the improvements in contribution of our fleet assignment model will come by only focusing on this 5%.

5.3.3 Another Illustrative Example

In order to handle these passengers, we introduce an illustrative scheduling scenario.

![Diagram showing flight network for the illustrative example]

Figure 5-7: Flight network for the illustrative example

The flight network for the illustrative example is shown in Figure 5-7 and the passenger demand data is in Table 5.5. The relevant characteristics of the three fleet types are found in Table 5.6. For this example, we will not concern ourselves with flow balance of aircraft. We are more concerned with fare allocation schemes, their problems, and how we might correct these problems.
5.3.4 Capacity-Based Fare Allocation

Ideally, we want to develop an allocation scheme that will do a better job of not overestimating or underestimating spilled revenue. Due to the network effects of passengers with multiple capacitated leg itineraries, it is impossible to find something that works exactly with the traditional fleet assignment model. Using the classification of passengers described in the previous section, we propose two allocation schemes. We define the first as the Capacity-Based Fare (CF) allocation scheme. The fare is distributed among the flight legs in the itinerary based on passenger classification as follows:

| Type 0: Ignore                                                                 |
| Type 1: Assign all of the fare to the capacitated leg.                        |
| Type 2: Divide the fare equally among all potentially capacitated legs.       |
| Type 3: Assign all of the fare to the overcapacitated leg.                    |
| Type 4: Divide the fare equally among all overcapacitated flight legs.        |

The second scheme does not make any distinctions between overcapacitated and potentially capacitated flight legs, and is a mix of the CF and the EF scheme. We define this second scheme as the Capacitated Equal Fare (CEF) allocation scheme. Specifically, it assigns the fare equally among all capacitated flight legs, i.e.,

<table>
<thead>
<tr>
<th>Market</th>
<th>Number of Passengers</th>
<th>Fare</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFO-ORD</td>
<td>30</td>
<td>$200</td>
</tr>
<tr>
<td>SFO-BOS</td>
<td>50</td>
<td>$375</td>
</tr>
<tr>
<td>SFO-ATL</td>
<td>80</td>
<td>$425</td>
</tr>
<tr>
<td>LAX-ORD</td>
<td>40</td>
<td>$150</td>
</tr>
<tr>
<td>LAX-BOS</td>
<td>50</td>
<td>$325</td>
</tr>
<tr>
<td>LAX-ATL</td>
<td>90</td>
<td>$325</td>
</tr>
<tr>
<td>ORD-BOS</td>
<td>51</td>
<td>$200</td>
</tr>
<tr>
<td>ORD-ATL</td>
<td>70</td>
<td>$200</td>
</tr>
</tbody>
</table>

Table 5.5: Demand data
Type 0: Ignore
Type 1-4: Divide the fare equally among all capacitated legs.

Capacity-Based Fare Motivation

Type 0 passengers do not affect the spill calculations since they travel on legs for which there will be no spill, no matter the fleet assignment. For Type 1 passengers, there is only one flight leg that will determine if a passenger is spilled. If this passenger is spilled, it is due to this capacitated flight leg, so we assign all of the fare to that specific flight leg. We prove in Chapter 6 that if all passengers are either Type 0 or Type 1, then using capacity-based fare allocation will result in an optimal fleeting.

A Type 2 passenger could potentially be spilled because of capacity problems on any of the capacitated flight legs. Therefore, we will divide the fare equally for all potentially capacitated flight legs. While this has the potential of under-estimating the spilled revenue on a potentially capacitated flight leg, it will do so less than cost-based or the equal fare allocation scheme since we have eliminated from consideration the uncapacitated flight legs that will not cause any spill.

If a passenger is Type 3, then there will be spill on the one overcapacitated flight leg. By assigning all of the fare to the overcapacitated flight leg, the average fare on the potentially capacitated flight legs will be less which may result in a smaller aircraft assigned to the potentially capacitated flight leg. By doing this, we might decrease operating costs. This point is more apparent by re-examining the illustrative example in this chapter. The demand on flight 4 is 230 passengers, therefore, at least 30 passengers must be spilled. If we have a leg independent view of the problem, then spilling the ORD-ATL passengers seems to be the optimal policy.

<table>
<thead>
<tr>
<th>Fleet Type</th>
<th>Seating Capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>100</td>
</tr>
<tr>
<td>B</td>
<td>150</td>
</tr>
<tr>
<td>C</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 5.6: Fleet type data

123
However, if we spill 30 LAX-ATL passengers, then the new demand on flight 2 becomes 150, and we could assign fleet type B without any more spill. Therefore, if the operating cost difference between fleet type B and fleet type C for the LAX-ORD is greater than $3,750 (30 passengers times the difference between the fare for LAX-ATL and ORD-ATL), then it is better to assign fleet type B rather than fleet type C to flight 2.

To determine a proper allocation scheme for Type 4 passengers, we combine the logic for Type 2 and Type 3 passengers. We assume that the overcapacitated flight legs will be spilling passengers anyway, and as in the Type 3 passengers, we wish to encourage smaller aircraft on the potentially capacitated flight legs. However, since we wish not to discriminate between the overcapacitated flight legs, we divide the fare equally among these legs.

**Computational Performance of the Capacity-Based Allocation Schemes**

We compare the results of the capacity-based fare allocation scheme with the best method of the previous implemented allocation schemes in Table 5.7. Except for the four fleet problem for the February data set, the $EF$ allocation scheme outperforms the $CF$ scheme consistently. However, the $CEF$ scheme seems to be on par with the $FF$ allocation scheme. On the other hand, the $CEF$ scheme outperforms the $EF$ scheme on both full size problems.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Z_S^* ($/day)</th>
<th>Best Scheme</th>
<th>κ_S</th>
<th>Z_CF^* ($/day)</th>
<th>κ_CF</th>
<th>Z_CEF^* ($/day)</th>
<th>κ_CEF</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>3,325,628</td>
<td>All</td>
<td>-</td>
<td>3,325,628</td>
<td>1.18-1.36</td>
<td>3,325,628</td>
<td>.76-1.42</td>
</tr>
<tr>
<td>F97-4A</td>
<td>5,141,331</td>
<td>FF</td>
<td>1.16</td>
<td>5,141,876</td>
<td>1.14</td>
<td>5,140,389</td>
<td>.88</td>
</tr>
<tr>
<td>F97-6A</td>
<td>7,208,352</td>
<td>EF</td>
<td>1.72</td>
<td>7,200,010</td>
<td>1.20</td>
<td>7,211,007</td>
<td>.92</td>
</tr>
<tr>
<td>F97-9</td>
<td>8,204,675</td>
<td>EF</td>
<td>1.12</td>
<td>8,207,754</td>
<td>.64</td>
<td>8,217,011</td>
<td>.32</td>
</tr>
<tr>
<td>A97-3A</td>
<td>1,448,159</td>
<td>EF</td>
<td>1.08</td>
<td>1,448,092</td>
<td>.54-.70</td>
<td>1,448,092</td>
<td>.54-.70</td>
</tr>
<tr>
<td>A97-4A</td>
<td>4,904,975</td>
<td>EF</td>
<td>1.14-1.18</td>
<td>4,900,589</td>
<td>.62</td>
<td>4,901,320</td>
<td>.70</td>
</tr>
<tr>
<td>A97-6A</td>
<td>8,037,422</td>
<td>EF</td>
<td>1.24</td>
<td>8,009,642</td>
<td>.92</td>
<td>8,030,979</td>
<td>.46</td>
</tr>
<tr>
<td>A97-9</td>
<td>11,227,126</td>
<td>EF</td>
<td>1.20</td>
<td>11,204,094</td>
<td>.82</td>
<td>11,232,458</td>
<td>.48</td>
</tr>
</tbody>
</table>

Table 5.7: A comparison of the capacity-based allocation schemes and the other fare allocation schemes
5.3.5 Probabilistic Allocation Schemes

In the $CF$ allocation scheme, we treat all potentially capacitated legs equally for Type 2 passengers. Likewise for Type 4 passengers, we considered all overcapacitated flight legs the same. In the $CEF$ allocation scheme, we treat all capacitated legs equally. We allocate the fare to the different capacitated flight legs equally for a Type 2 passenger, assuming that each leg was equally as important for the capturing of the passenger fares. This may not always be the case.

Returning to the illustrative example, assume that fleet type A cannot be assigned to either flight 2 or flight 3. Passengers in the LAX-BOS market are Type 2 passengers, since both flight 2 and flight 3 are potentially capacitated. Flight 2 has a demand of 180 passengers and flight 3 has a demand of 151 passengers. Let's assume that either feasible aircraft type is equally likely to be assigned to the flight legs. The probability of there being spill on both flight legs is .25 (the probability of fleet type B being assigned to both flight legs). However, let's look at a specific passenger. If we assume that any passenger on a capacitated flight leg is equally likely to be spilled, then the probability of a specific passenger on flight 3 being spilled because of this leg is roughly .003 (a 50% chance of the 150 seat aircraft being assigned, and roughly a .6% chance of being the spilled passenger). However, the probability of a specific passenger being spilled on flight 2 is .08 (a 50% chance of the 150 seat aircraft being assigned and a 16% chance of being one of the spilled passengers). Therefore, the probability of being spilled on flight 2 is over 20 times more likely than flight 3. We might not want to consider these flight legs equally important from the standpoint of the passenger in the LAX-BOS market, since the probability of being spilled by each flight leg is not equal.

We consider pro-rating the fare among the different capacitated flight legs based on the probability of a specific passenger being spilled. We will define this as the probabilistic pro-rated fare ($PF$) allocation scheme, which implies an average fare of:

$$
\overline{fare_i}^{PF} = \frac{\sum_{p \in P} \frac{p_i.fare_p}{\text{prob}_p} b_i^p D_p}{\sum_{p \in P} b_i^p D_p},
$$

(5.17)
where the probability of being spilled on flight leg \( i \) is calculated by:

\[
pr_i = \frac{\sum_{k \in K(i)} n_k \frac{\max(0, Q_i - SEATS_k)}{Q_i}}{\sum_{k \in K(i)} n_k},
\]

and

\[
prob_p = \begin{cases} 
1 & \text{if } \sum_{i \in L} \delta^p pr_i = 0 \\
\sum_{i \in L} \delta^p pr_i & \text{otherwise}
\end{cases}.
\]

Another method assigns all of the fare to the flight leg that has the greatest probability of spilling. We define this method as the greatest probability fare (GF) allocation scheme. The performance of the probabilistic allocation schemes are shown in Table 5.8.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>( Z_S^* ) ($/day)</th>
<th>Best Method</th>
<th>( Z_{PF}^* ) ($/day)</th>
<th>( \kappa_{PF}^* )</th>
<th>( Z_{GF}^* ) ($/day)</th>
<th>( \kappa_{GF}^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>3,325,628</td>
<td>All</td>
<td>3,325,628</td>
<td>1.16-1.38</td>
<td>3,325,628</td>
<td>1.14-1.38</td>
</tr>
<tr>
<td>F97-5A</td>
<td>5,141,876</td>
<td>CF</td>
<td>5,141,876</td>
<td>1.12</td>
<td>5,141,876</td>
<td>1.14</td>
</tr>
<tr>
<td>F97-6A</td>
<td>7,211,007</td>
<td>CEF</td>
<td>7,200,677</td>
<td>1.04</td>
<td>7,200,010</td>
<td>1.18</td>
</tr>
<tr>
<td>F97-9</td>
<td>8,217,011</td>
<td>CEF</td>
<td>8,200,287</td>
<td>.48</td>
<td>8,199,315</td>
<td>.60</td>
</tr>
<tr>
<td>A97-3A</td>
<td>1,448,159</td>
<td>EF</td>
<td>1,448,092</td>
<td>.54-.70</td>
<td>1,448,092</td>
<td>.54-.70</td>
</tr>
<tr>
<td>A97-5A</td>
<td>4,904,974</td>
<td>EF</td>
<td>4,899,905</td>
<td>.84</td>
<td>4,900,810</td>
<td>.82</td>
</tr>
<tr>
<td>A97-6A</td>
<td>8,037,422</td>
<td>EF</td>
<td>8,013,345</td>
<td>.96</td>
<td>8,010,814</td>
<td>1.00</td>
</tr>
<tr>
<td>A97-9</td>
<td>11,232,458</td>
<td>CEF</td>
<td>11,189,321</td>
<td>.68</td>
<td>11,177,998</td>
<td>.40</td>
</tr>
</tbody>
</table>

Table 5.8: A comparison of the probabilistic allocation schemes

Using the representative spill fare model, the two fare allocation schemes that performed the best seemed to be the equal fare and the capacity-based equal fare model. From the results presented in the previous sections, the mileage-based pro-rated fare scheme that is most commonly used in industry performed poorly compared to the other allocation schemes. We further investigate the possible reasons for this in Section 5.4.

### 5.3.6 Spill Integration

In the previous sections, we concentrate on methods to allocate fare to each flight leg to obtain a better estimate of the costs associated with spilling passengers. Given these more accurate costs,
the representative spill fare model aggregates this data into an average fare that is multiplied by some universal constant to take into account the partial control of passenger bookings afforded by revenue management systems. In the representative spill fare model, we approximate the spill costs on flight leg $i$ as:

$$c_{k,i}^{sp} \approx \sum_p \delta^p_i \text{fare}_p t^k_p,$$

where $t^k_p$ is the number of passengers spilled from itinerary $p$ when fleet type $k$ is assigned to flight leg $i$. The representative spill fare model sets this value equal to the following:

$$c_{k,i}^{sp} = \kappa \cdot \text{fare}_i \cdot \left( \max \left\{ 0, \sum_{p \in P} \delta^p_i D_p - SEATS_k \right\} \right).$$

We see one potential problem with this model. We assume that the average fare of spilled passengers is independent of the type of aircraft that is assigned to the flight leg. Moreover, this assumption means that the average fare of spilled passengers is independent of the number of passengers spilled. Reviewing the illustrative example presented in this chapter, assume that we cannot assign fleet type B to flight 4 because of operational considerations. If the airline has perfect control of passenger bookings, the optimal solution is to spill 30 passengers in the ORD-ATL market when fleet type C is assigned. The average fare of spilled passengers in this example is $200. However, if we assign fleet type A, then the optimal solution is to spill 70 passengers in the ORD-ATL market and 60 passengers in the LAX-ATL market. In this case, the average fare of spilled passengers is roughly $258, which is much different than the $200 average fare when fleet type C is assigned.

The above example highlights a potential problem. In a situation with uncertain demand data, Belobaba and Farkas [10] showed that the average fare of spilled passengers is dependent on the number of passengers that are spilled. Even with deterministic demand data, the optimal spill fare will be based on the number of passengers spoiled.

To rectify this, we introduce a method, denoted as the integrated spill model (ISM), to determine the spill costs. In this method, the spill costs of each flight leg and fleet type combination, $c_{k,i}^{sp}$, is estimated as the optimal objective function of the following program:
\[
\text{minimize } \sum_{p \in P} \text{fare}_p^i t_p \\
\text{subject to: } \sum_{p \in P} \delta^i_{p} t_p \geq Q_i - \text{SEATS}_k \\
0 \leq t_p \leq D_p \quad \forall p \in P,
\]

where \( \text{fare}_p^i \) is the fare of itinerary \( p \) assigned to flight leg \( i \) based on one of the allocation schemes described above. The optimal solution to this program can be found with a simple greedy algorithm. For each flight leg \( i \), we sort, in ascending order, passenger itineraries by \( \text{fare}_p^i \). To determine the spill for fleet type \( k \), we remove the first \( Q_i - \text{SEATS}_k \) passengers from the sorted list.

The ISM assumes that each flight leg is independent. However, we have assumed that the majority of the network effects have been captured by the fare allocation scheme being used. In Chapter 6, we prove that the ISM coupled with the capacity-based fare allocation scheme will give the optimal fleet assignment when all passengers are either Type 0 or Type 1. Table 5.9 shows the performance using the ISM instead of the representative fare spill model, RFSM. We denote the different fare allocation schemes with the ISM by replacing the \( F \) with an \( S \), e.g., CES denotes using the capacitated equal fare allocation scheme with integrated spill.

When comparing the RFSM to the ISM, the RFSM has a distinct advantage in terms of solution quality. We use an iterative approach by which we determine numerous fleetings, then chose the best solution. This is done so the potential benefits of using the RFSM are not shrouded by a bad choice of \( \kappa \). On the other hand, the ISM solves the problem in one iteration. However, in almost every data set, the ISM outperforms the RFSM. Examining Table 5.9, we see some trends:

- The MS solution technique (mileage-based pro-rated fare allocation with integrated spill) performed the worst for most runs. For the two full-sized problems, the MS scheme was $288K/day ($105M/year) and $293K/day ($107M/year) worse than the best solution found.
- Using the integrated spill model resulted in the best solution in all but one of the data sets.
<table>
<thead>
<tr>
<th>Data Set</th>
<th>$Z_{MF}^*$ ($/\text{day})$</th>
<th>$Z_{SF}^*$ ($/\text{day})$</th>
<th>Best Method</th>
<th>$Z_{FS}^*$ ($/\text{day})$</th>
<th>$Z_{MS}^*$ ($/\text{day})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>3,325,628</td>
<td>3,325,628</td>
<td>All</td>
<td>3,325,628</td>
<td>3,325,628</td>
</tr>
<tr>
<td>F97-4A</td>
<td>5,136,779</td>
<td>5,141,876</td>
<td>SF,PF,GF</td>
<td>5,140,685</td>
<td>5,131,643</td>
</tr>
<tr>
<td>F97-6A</td>
<td>7,176,229</td>
<td>7,208,352</td>
<td>EF</td>
<td>7,216,325</td>
<td>7,112,211</td>
</tr>
<tr>
<td>F97-9</td>
<td>8,199,869</td>
<td>8,217,011</td>
<td>CEF</td>
<td>8,224,292</td>
<td>7,935,877</td>
</tr>
<tr>
<td>A97-3A</td>
<td>1,448,092</td>
<td>1,448,159</td>
<td>EF</td>
<td>1,448,159</td>
<td>1,447,819</td>
</tr>
<tr>
<td>A97-4A</td>
<td>4,901,197</td>
<td>4,904,975</td>
<td>EF</td>
<td>4,913,419</td>
<td>4,906,387</td>
</tr>
<tr>
<td>A97-6A</td>
<td>8,014,395</td>
<td>8,037,422</td>
<td>EF</td>
<td>8,043,037</td>
<td>7,898,639</td>
</tr>
<tr>
<td>A97-9</td>
<td>11,155,890</td>
<td>11,232,458</td>
<td>CEF</td>
<td>11,246,977</td>
<td>10,953,626</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data Set</th>
<th>$Z_{ES}^*$ ($/\text{day})$</th>
<th>$Z_{CS}^*$ ($/\text{day})$</th>
<th>$Z_{CES}^*$ ($/\text{day})$</th>
<th>$Z_{PS}^*$ ($/\text{day})$</th>
<th>$Z_{GS}^*$ ($/\text{day})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-4A</td>
<td>5,117,715</td>
<td>5,067,344</td>
<td>5,141,310</td>
<td>5,140,795</td>
<td>5,071,446</td>
</tr>
<tr>
<td>F97-6A</td>
<td>7,200,854</td>
<td>7,145,390</td>
<td>7,215,180</td>
<td>7,215,784</td>
<td>7,149,382</td>
</tr>
<tr>
<td>F97-9</td>
<td>8,215,423</td>
<td>8,159,432</td>
<td>8,223,717</td>
<td>8,224,092</td>
<td>8,214,604</td>
</tr>
<tr>
<td>A97-3A</td>
<td>1,447,752</td>
<td>1,447,264</td>
<td>1,447,116</td>
<td>1,446,643</td>
<td>1,448,465</td>
</tr>
<tr>
<td>A97-4A</td>
<td>4,911,506</td>
<td>4,913,197</td>
<td>4,913,196</td>
<td>4,913,692</td>
<td>4,912,024</td>
</tr>
<tr>
<td>A97-6A</td>
<td>8,024,555</td>
<td>8,039,557</td>
<td>8,043,632</td>
<td>8,043,519</td>
<td>8,036,807</td>
</tr>
<tr>
<td>A97-9</td>
<td>11,234,985</td>
<td>11,240,090</td>
<td>11,247,114</td>
<td>11,246,828</td>
<td>11,227,382</td>
</tr>
</tbody>
</table>

Table 5.9: The performance of spill integration compared to the representative approach
Specifically, the CES, FS and PS methods were always within $2,000/day ($730K/year) of the best solution found for every data set.

- For the two full sized problems, the best solution resulted in savings of $24,423/day ($8.9M/year) and $91,224/day ($33.3M/year) over the commonly used MF scheme.

5.4 Analysis of Solution Techniques

5.4.1 Performance Trends

From the above results in Table 5.9, there is no question that the integrated spill model performs better than the representative fare spill model. As evidenced by the poor spill inaccuracy results, the RFSM does not do a good job of accurately estimating spill. The key to the success of ISM comes from the relaxation that the average spill fare is a constant regardless of the assignment.

One of the first things we notice about the results from above is the consistently good performance of the CES, FS, and PS solution techniques. There are two attributes of these schemes that makes them different from the others:

1. A value at least greater than or equal to the entire fare is distributed to the capacitated flight legs in each passenger itinerary.

2. Each capacitated flight leg in a passenger itinerary has at least some fare assigned to it.

By assigning a large percentage of the fare to the capacitated flight legs, this guarantees that the assignment of larger aircraft to capacitated flights is a priority. When using the ISM model, the spill costs with each flight leg will always be greater than with the ES scheme. It also appears that the PS solution technique does a proper job of dividing the fare between the different capacitated flight legs.

While the CS and GS fare schemes satisfy the first attribute, they might not assign any cost of spill for a specific itinerary to potentially capacitated flight legs. This might lead to assigning a smaller aircraft than necessary to potentially capacitated flight legs. This is indeed the case. Too many Type 3 and 4 passengers in the mix lead to too many spills on potentially capacitated flight legs. The common thread with the CES, FS, and PS solution techniques is
that they all assign some fare to each capacitated flight leg for each passenger. The CS and GS schemes underestimated the spill of capacitated flight legs that were the least likely to spill passengers.

However, the CS scheme still performs well in some circumstances (specifically the August data set), and not so well in others (the February data set). The main reasons for this performance variability can be explained by Table 5.3. The CS scheme does well on the data sets with a low ratio of overcapacitated to potentially capacitated flight legs, which drives the likelihood that we will underprice legs. The reason above explains this distinction. In the February data set, too much of the spill costs are assigned to the overcapacitated flight legs, thus resulting in not enough large planes assigned to the potentially capacitated flight legs.

5.4.2 Performance as Demand Characteristics Change

A valid question is determining whether or not the performances have something to do with the characteristics of the demand data. Two of the main characteristics that we examine are the congestion in the system and the level of network effects. The congestion of the system is measured by the leg average demand factor (DF), which is calculated by dividing the number of seats requested by the supply of seats. The level of network effects is measured by the passenger connectivity ratio (PCR), which is the number of connecting passengers divided by the number of total passengers. Table 5.10 presents the DF and PCR for the different data sets that are used. As mentioned before, the F97 series represents a data set for February 1997 which is typically a low demand month (a .608 demand factor), and the A97 series represents a data set for August 1997, typically a high demand month for the airline industry (a .699 demand factor). One interesting note is that the passenger connectivity ratio is virtually the same for both scenarios. From just the limited number of data points, it is difficult to make any judgments based on demand factor and passenger connectivity ratios.

To generate more data, we have selected one of the data sets (A97-4A) and performed the following two experiments:

1. We vary the relative congestion in the system, i.e., we change the demand factor, DF, as described in Section 3.7.2 (Figure 5-8).
2. Keeping the demand factor at a constant medium level ($DF = 0.65$), we vary the relative number of connecting passengers in the system, i.e., we change the passenger connectivity ratio, $PCR$, as described in Section 3.7.3 (Figure 5-9).

![Diagram showing normalized contribution vs demand factor for different data sets.]

Figure 5-8: The performance of various solution techniques as the demand factor is varied

In Figures 5-8 and 5-9:

- As the demand factor increases, the integrated spill models perform better than the representative fare models, especially the $CS$ model. At the medium demand factors, the representative fares perform the best.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>DF</th>
<th>PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>.642</td>
<td>.029</td>
</tr>
<tr>
<td>F97-4A</td>
<td>.653</td>
<td>.053</td>
</tr>
<tr>
<td>F97-6A</td>
<td>.583</td>
<td>.098</td>
</tr>
<tr>
<td>F97-9</td>
<td>.608</td>
<td>.227</td>
</tr>
<tr>
<td>A97-3A</td>
<td>.629</td>
<td>.046</td>
</tr>
<tr>
<td>A97-4A</td>
<td>.643</td>
<td>.078</td>
</tr>
<tr>
<td>A97-6A</td>
<td>.650</td>
<td>.112</td>
</tr>
<tr>
<td>A97-9</td>
<td>.699</td>
<td>.226</td>
</tr>
</tbody>
</table>

Table 5.10: The demand characteristics of the different data sets
Figure 5-9: The performance of various solution techniques while the passenger connectivity ratio is varied (keeping the demand factor constant)

- The integrated spill models perform well at low passenger connectivity ratios, while the spill fare models perform better at higher passenger connectivity ratios, especially the CF allocation scheme.

5.5 The Iterative Methodology

In Chapter 2, we present an iterative methodology (Figure 5-10) for the fleet assignment model and discuss traditional fleet assignment solvers. In the previous sections of this chapter, we present methods to improve upon the estimates of spill before solving the initial fleet assignment. In Chapter 3, we present a new model for spill calculation and compare it to other methods of spill calculation. Now that we have thoroughly examined these other components, we are now ready to describe how we can combine them into an iterative procedure that could potentially yield better fleet assignment solutions. Like many other iterative solution procedures for general problems, we must now specify two important components: the procedure for modifying the problem from iteration to iteration, and the criteria to be used for determining when to terminate the algorithm.
Figure 5-10: An iterative methodology for the fleet assignment model

An iterative methodology will only be effective if useful information is passed from one iteration to the other. The important consideration in this process is which information should be passed. We will discuss a couple of different methods that can be used to achieve possibly better fleet assignment solutions. The first is a simple approach that will update the objective cost coefficients in the traditional fleet assignment model. A second approach will be a model that slightly alters the traditional fleet assignment model. It will include columns that represent multiple assignments of different fleet types to subsets of flight legs.
5.6 Cost Coefficient Modification Approach

All of the different allocation schemes that are presented earlier in this chapter only affect the cost coefficients of the traditional fleet assignment model. These changes do not affect the constraints of the problem. We show that by just slightly altering the cost coefficients, substantial differences to the quality of the fleet assignment solution are realized. In Chapter 2, we prove that all feasible minimal fleetings are extreme points of the fleet assignment polyhedron. Therefore, if the objective cost coefficients are properly set, then the true optimal fleet assignment will be an extreme point of the fleet assignment polyhedron. This is the basis for considering an iterative approach that just updates the costs of the fleet assignment model.

In Figure 5-10, if the termination criteria is not satisfied, then we alter the problem to get a better solution potentially. In this section we present some modification techniques. Based on the spill models presented in this chapter, we propose:

1. For the representative fare spill model, we alter our value of the spill fare to average fare ratio, $\kappa$.

2. For the capacity-based allocation schemes, we alter the classification of flight legs and passenger itineraries based on the previous iteration.

3. For the probabilistic allocation scheme, we alter the probability of a passenger being spilled based on the previous iteration.

5.6.1 Spill Fare to Average Fare Ratio Modification

Accurate Spill Measurement

Since the contribution function, with regards to the spill fare to average fare ratio, $\kappa$, is not quasiconcave, we cannot perform a simple line search to find an optimal $\kappa$. One potential
algorithm is as follows:

**Step 0**: Set $\kappa^1 = 1.0$ and $j = 1$.

**Step 1**: Let $X^j$ be the fleeting decision using $\kappa^j$.

Let $C^{sp}(\kappa^j, X^j) = \sum_{k \in K} \sum_{i \in L} c_{k,i}^{sp}(\kappa^j) \cdot X_{k,i}^j$.

**Step 2**: Solve the spill calculation module using $X^j$, where $Z(X^j)$ is the optimal objective value

**Step 3**: Set $\kappa^{j+1} = \kappa^j \frac{Z^{PMIX}(X^j)}{C^{sp}(\kappa^j, X^j)}$ and $j = j + 1$.

**Step 4**: If the termination criteria is met, STOP!
Otherwise, go to Step 1.

One termination criterion is to stop when at iteration $j$, $X^j = X^{j-1}$. Another criterion is to stop if $\kappa^j = \kappa^{j-1}$, or for practical purposes $|\kappa^j - \kappa^{j-1}| < \epsilon$, where $\epsilon$ is some sufficiently small tolerance. For our purposes, we use the criteria that both of the above conditions must be met to terminate. Since the termination of this algorithm is not guaranteed, a limit on the number iterations should be included. This process has no guarantee that a solution at iteration $j$ is better than iteration $j - 1$. Therefore, we record the best solution obtained. In the practical sense, this process may be able to give a set of good solutions that can be evaluated by human planners.

The above routine assumes that accurately measuring the spill will in general lead to good solutions. However, we show in Section 5.2.4 that this is not always the case. The performance of this method has lead to poor solutions.

**Step Approach**

If time is not an issue, the iterative method presented in Section 5.2.4, can work well with RFSM. The graph of the contribution versus $\kappa$ typically has the same shape. As $\kappa$ increases form 0.0, a steep incline in contribution exists, then a plateau of good solutions, and finally a tailing area as the ratio is overestimated. By doing a number of iterations spread out across various $\kappa$ values, one can get a sense of the graph and pinpoint areas for more extensive searching.
5.6.2 Capacity-Based Fare Allocation Modification

When using the $CF$ and $CEF$ allocation schemes, we can use information from previous iterations to change our classification of flight legs, thus changing the classification of the passengers. For example, let’s assume that there is some flight leg where the unconstrained demand is 101 passengers. Let’s assume that there are three fleet types with 100, 150, and 200 seats. This flight leg will be classified as potentially capacitated. However, after an iteration, let’s say that the flight leg was assigned the 150 seat aircraft. Changing the flight leg’s classification to uncapacitated, may affect the fare distribution of a number of passengers who use this flight leg, assigning more fare to other capacitated flight legs. Therefore, one iterative procedure is as follows:

**Step 0:** Set $j = 1$. Set $\overline{CAP}_{i}$ and $\overline{CAP}_{i}$ as usual.

**Step 1:** Let $X^j$ be the fleeting decision at iteration $j$.

Set $\overline{CAP}_{i} = \overline{CAP}_{i} = \sum_{k \in K} SEATS_{k}X^j_{k,i}$

and reclassify flight legs and passengers.

**Step 2:** Set $j = j + 1$.

**Step 3:** If the termination criteria is met, STOP!

Otherwise, go to Step 1.

The straightforward termination criteria is to end when $X^j = X^{j-1}$. In fact, once this condition occurs, then future iterations will always yield the same solution. After the first iteration, we notice that we will either have uncapacitated aircraft or overcapacitated aircraft, therefore, the classifications of Type 2 or Type 3 passengers are eliminated. The disadvantage of this approach is that it might too quickly fix the fleetings of aircraft thus converging to a poor quality solution.

We would like an approach that will not fix an aircraft type into a specific fleeting as quickly and allow us to still classify flight legs as potentially capacitated. We propose the follow alteration to the above algorithm.
Step 0: Set $j = 1$. Set $\overline{CAP}_i$ and $\overline{CAP}_i$ as usual

Step 1: Let $X^{j,LP}_i$ be the fleeting decision at iteration $j$ of the LP relaxation. Let $K^j(i)$ be the set of fleet types $k$ such that $X^{j,LP}_{k,i} > 0$

Step 2: Let $\overline{CAP}_i = \min_{k \in K^j(i)} \{SEATS_k\}$,
$\overline{CAP}_i = \max_{k \in K^j(i)} \{SEATS_k\}$, and reclassify flight legs and passengers.

Step 3: Set $j = j + 1$.

Step 4: If the termination criteria is met, STOP! Otherwise, go to Step 1.

The termination criterion is $|X^{j,LP}_i - X^{j-1,LP}_i| < \varepsilon$.

5.6.3 Probabilistic Allocation Scheme Modification

For the set of probabilistic allocation schemes, we might use the LP relaxation as an estimate for the probability of a specific fleet type being assigned to a flight leg. In the initial phase of these allocation schemes, an estimate of the probability of fleet type $k$ being assigned to flight leg $i$ was $n_k / (\sum_{k \in K(i)} n_k)$. The problem modification routine for probabilistic allocation schemes is:
Step 0: Set $j = 1$.

Step 1: Let $X^{j,LP}$ be the fleeting decision at iteration $j$ of the LP relaxation. Let $K^j(i)$ be the set of fleet types $k$ such that $X^{j,LP}_{k,i} > 0$.

Step 2: Let $pr_i = X^{j,LP}_{k,i} \max(0,Q_i-SEATS_i)/Q_i$ and recalculate the spill costs on each flight leg.

Step 3: Set $j = j + 1$.

Step 4: If the termination criteria is met, STOP! Otherwise, go to Step 1.

The termination criterion is $|X^{j,LP} - X^{j-1,LP}| < \varepsilon$.

5.7 Multiple Flight Column Approach

While changing the cost coefficients is one approach to capture the network effects of spill, we can not guarantee the optimal fleeting will be found. The main problem is that we model the passenger itinerary behavior linearly with respect to our flight specific decision variable. However, the assignment of fleet types to adjacent flight legs might have a dramatic impact on spill. Let's consider a formulation where the spill is linear with respect to the decision variable. Primarily done for capturing maintenance restrictions in Barnhart et al. [4] and Hoffman [26], the decision variable was the assignment of one fleet type to a string of flight legs. We build upon this idea by having a decision variable represent the assignment of many fleet types to many (not necessarily sequential) flight legs.

5.7.1 The Illustrative Example Revisited

Again considering the illustrative example, we can find the optimal solution by comparing the contribution (using the passenger mix model) for each fleeting and selecting the best one. This exhaustive method, however, requires us to express each fleeting explicitly. Let's assume, however, that fleet type A cannot be assigned to flight 1. Therefore, flight 1 is uncapacitated.
and the fleeting decision for this flight leg has no bearing on the resulting spill. In this case, we can determine the fleeting decision for flight leg 1 separately from the fleeting decision of the other flight legs. In a real fleet assignment decision process, we concern ourselves about aircraft flow balance. We generalize and mathematically formulate this approach.

5.7.2 New Decision Variables and Assignment Combinations

Let's divide the entire set of flight legs into \( N \) disjoint subsets, \( \{ L_n \} \), where \( q_n = |L_n| \). Further, let us denote the members of subset \( L_n \) as

\[
L_n = \{ i_1, i_2, ..., i_{q_n} \}.
\]  

(5.23)

We define a set of variables for each subset as

\[
X_{i_1, i_2, ..., i_q}^{k_1, k_2, ..., k_{q_n}} = \begin{cases} 
1 & \text{if flight leg } i_1 \text{ is assigned to fleet type } k_1 \\
& \text{AND flight leg } i_2 \text{ is assigned to fleet type } k_2 \text{ AND, etc.} \\
0 & \text{otherwise.} 
\end{cases}
\]

There is one variable for every assignment combination of fleet types to flight legs, thus the number of variables associated with subset \( L_n \) is equal to

\[
|K(i_1)| |K(i_2)| \cdots |K(i_{q_n})| = o(|K|^{q_n}),
\]  

(5.24)

i.e., it is exponential in the size of the subset. Let \( U(L_n) \) be the set of assignment combinations for subset \( L_n \), indexed by \( U \). An assignment combination, \( U \), is defined by a set of flight legs and an associated fleet type for each flight leg, i.e.,

\[
U = \{ (i_1, k_1), (i_2, k_2), ..., (i_{q_n}, k_{q_n}) \}.
\]

For ease, we will change our notation such that

\[
X_{i_1, i_2, ..., i_q}^{k_1, k_2, ..., k_q} \rightarrow X_U.
\]
In this approach, which we define as the *multiple flight column (MFC)* approach, we are removing columns associated with a group of flight legs from the traditional fleet assignment model and replacing them with columns that represent all feasible assignment combinations of them.

### 5.7.3 Developing Partitions: A Set Partitioning Problem

#### The Formulation

The key to the MFC approach is the determination of a good partition of flight legs. Let us consider only the set of flight legs that are capacitated, which we will denote as $\overline{L}$. We can do this since the assignment of any fleet type to an uncapacitated flight leg will not cause spill on other flight legs. We define the *connectivity* between two capacitated flight legs $i_1$ and $i_2$ as the number of passengers whose itinerary includes both flight legs. The connectivity, $C_{i_1, i_2}$, between two flight legs can be found by the following equation:

$$C_{i_1, i_2} = \sum_{p \in P} \delta_{i_1}^p \delta_{i_2}^p D_p. \quad (5.25)$$

It stands to reason that if the connectivity between two flight legs is high, then the assignment of a fleet type to one of the flight legs will have an impact on the spill costs of the other flight leg. Therefore, we want to include flight legs with high connectivity with one another in the same set. We can formulate this as a set partitioning problem.

Let’s assume that we have only $m$ different subsets of $\overline{L}$, denoted by $L^1, L^2, \ldots, L^m$. The weight of subset $L^j$ is calculated by:

$$w_j = \frac{\sum_{i_1 \in L^j} \sum_{i_2 \in L^j \setminus \{i_1\}} C_{i_1, i_2}}{2}, \quad (5.26)$$

with the division by 2 to correct the double counting associated with the sum in the numerator. We will define decision variable $u_j$ and indicator parameter $a^j_i$ as follows:

$$u_j: = \begin{cases} 
1 & \text{if subset } L^j \text{ is used in the partition.} \\
0 & \text{otherwise.}
\end{cases}$$

141
\( a^j_i = \begin{cases} 
1 & \text{if flight leg } i \text{ is a member of subset } L^j \\
0 & \text{otherwise.} 
\end{cases} \)

We now formulate our problem as such:

\[
(MFC - SP) \quad \text{maximize} \quad \sum_{j=1,2,\ldots,m} w_j u_j \\
\text{subject to:} \quad \sum_{j=1,2,\ldots,m} a^j_i u_j = 1 \quad \forall i \in \bar{L} \tag{5.27}
\]

\[u_j = \{0, 1\} \quad \forall j = 1, \ldots, m.\]

If we can include all potential subsets of flight legs, then the (not necessarily unique) optimal solution to \( MFC - SP \) is the subset that includes all capacitated flight legs. The objective value of this solution is the sum of all of the connectivity values for every pair of flight legs, i.e.,

\[
\frac{\sum_{i_1 \in \bar{L}} \sum_{i_2 \in \bar{L} \setminus \{i_1\}} C_{i_1,i_2}}{2}. \tag{5.28}
\]

Another partition is optimal if the following condition holds: for every flight leg in a subset, it has zero connectivity with all other flight legs not in its subset.

The reason for solving this set partitioning problem is to group flight legs to define multiple flight columns. Since the optimal solutions to \( MFC - SP \) could have a large collection of flight legs, we might want to limit the number of flight legs in any partition to \( q \). This restricted problem can be formulated as in 5.27; however, we must limit the different subsets to those that have no more than \( q \) flight legs, which means that no column of \( MFC - SP \) has more than \( q \) non-zero entries.

Another approach is to limit the number of variables in the resulting \( MFC \) approach, instead of limiting the number of flight legs that can be in a subset. For example, we might want to have at most some pre-set limit number of variables, \( N^v \). The number of variables
associated with subset $L^m$, which we denote as $n_m$, can be found by:

\[ n_m = \prod_{i \in L^m} |K(i)|. \]  \hfill (5.29)

Then we add the following constraint to $MFC - SP$

\[ \sum_m n_m u_m \leq N^v. \]  \hfill (5.30)

One can experiment to find an appropriate value of $N^v$. In the case where each flight leg is its own subset, the number of flight decision variables (excluding ground arc variables) is $\sum_{i \in L} K(i)$. Therefore a reasonable value of $N^v$ may be $\sum_{i \in L} \alpha K(i)$, that is linear in the number of flight legs, like the set of variables associated with the traditional fleet assignment model.

We may also relax Equation 5.30. This is equivalent to changing the problem back to the form of a general set partitioning problem with the new objective coefficient of $w_m - \mu n_m$ for subset $L^m$ where $\mu$ is a price associated with violating the constraint.

**Flight Leg Partition Determination: A Heuristic Approach**

The general set partitioning problem is NP-hard [13]. Therefore, we propose the following iterative heuristic to solve the problem. Let $LIST$ be a queue data structure which is a set of sets.
Step 0: Set $LIST = \{L\}$.
Step 1: Select the first subset from $LIST$, and remove it from $LIST$. Denote this subset $M$.
Step 2: Determine the all-pairs minimum cut in the following network:

- **Nodes**: Set of nodes = set of flight legs in $M$.
- **Arches**: If connectivity between two nodes $i$ and $j$ is strictly positive, set arc capacity to the connectivity of the flights associated with $i$ and $j$.

Step 3: Put two new subsets defined by cut into $LIST$.
Step 4: If the termination criteria is met, STOP!

Otherwise, go to Step 1.

For the termination criteria we can choose either that each subset has less than $q$ elements or that 5.30 is satisfied.

The standard minimum cut problem is for a pre-defined source and sink node. The all pairs minimum cut problem finds the minimum cut for every pair of nodes. Ahuja et al. [2] show that the solution of the all-pairs minimum cut problem can be determined by solving $|M - 1|$ standard minimum cut problems.

### 5.7.4 The Multiple Flight Column Formulation

A column in the traditional fleet assignment model represents one arc in a fleet-specific timeline network (Chapter 2). The decision variable is whether or not we use this arc. A column in this new formulation represents whether or not a set of arcs is used. These arcs may or may not be adjacent in the timeline network. However, they are connected based on passenger flows, i.e., for every two flight legs in a set of arcs, there are passengers that have both flight legs in their itinerary. In the traditional fleet assignment model, since we have only one arc per decision variable, we know that it can cross the timeline only once. However, with our new decision variable, the set of arcs might cross the timeline numerous times for different fleet types. This
might be the case for the flow balance of aircraft as well. Therefore, we define two sets of parameters that are associated with each assignment combination variable

\( v^U_k \): the number of times that the flight leg arcs in the assignment combination \( U \) cross the timeline for fleet type \( k \). This value can be any integer greater than or equal to zero.

\( w^U_{\{k,o,t\}} \): the number of flight arcs in assignment combination \( U \) that are incident to node \( \{k,o,t\} \) minus the number of flight arcs in the assignment combination \( U \) that are incident from node \( \{k,o,t\} \). This can be any positive or negative integer or zero.

The formulation for the \( MFC \) approach is as follows:

\[(FAM - MFC)\]

**maximize**

\[
\sum_{n=1}^{N} \sum_{U \in U(L_n)} C_U X_U
\]

**subject to:**

\[
\sum_{U \in U(L_n)} X_U = 1 \quad \forall n = 1, \ldots, N
\]

\[
\sum_{n=1}^{N} \sum_{U \in U(L_n)} v^U_k X_U + \sum_{o \in A} Y^+_k,o,t_m^+ \leq n_k \quad \forall k \in K
\]

\[
\sum_{n=1}^{N} \sum_{U \in U(L_n)} w^U_{\{k,o,t\}} X_U + Y^+_k,o,t_j^+ - Y^-_k,o,t_j^+ = 0 \quad \forall \{k,o,t_j\} \in N
\]

\( X_U \in \{0,1\} \forall U \in U(L_n), n = 1, \ldots, N \)

\( Y^+_k,o,t_j, Y^-_k,o,t_j \geq 0 \) and integer

\( \forall \{k,o,t_j\} \in N \)

The objective cost coefficient is found by the following relationship:

\[
C_U = Z^P_{U} + \sum_{(i,k) \in U} \tilde{c}_{k,i}, \quad (5.32)
\]

where the first term is an estimate of the spill costs when the different assignments in the assignment combination \( U \) are in place and \( \tilde{c}_{k,i} \) is the operating cost of flying flight leg \( i \) with
flees type \( k \). The estimate of spill costs is the optimal objective value to the following passenger mix subproblem:

\[
\text{minimize} \quad \sum_{p \in P(U)} \text{fare}_p t_p \\
\text{subject to:} \quad \sum_{p \in P} \delta^p_i t_p \geq Q_i - \text{SEATS}_k \quad \forall (i, k) \in U \\
0 \leq t_p \leq D_p \quad \forall p \in P,
\]  

(5.33)

where \( P(U) \) is the set of itineraries that contain the flight legs that are in assignment combination \( U \).

Let's look at the extremes of this approach to solving the fleet assignment model. In the case where each subset contains exactly one flight leg, the problem reduces to the traditional fleet assignment. The objective function reduces to the ISM using the full fare allocation scheme. When all of the flight legs are in one subset, the problem becomes an exhaustive search of all feasible fleet assignments, as each column is an entire fleeting. The minimization program would simply select the least cost fleeting decision. This is obviously not a practical approach for even relatively small problems. However, this approach does capture all of the network effects. The thrust of this research is to find some middle ground between the two approaches that captures most of the network effects without exploding the problem size.
5.7.5 Problem Modification Routine

The problem modification routine is similar to the capacity based fare allocation schemes (Section 5.6.2), i.e., we reclassify flight legs based on the LP relaxation.

<table>
<thead>
<tr>
<th>Step 0:</th>
<th>Set ( j = 1 ). Set ( \text{CAP}_i ) and ( \overline{\text{CAP}}_i ) as usual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1:</td>
<td>Let ( X^{j,LP}<em>i ) be the fleeting decision at iteration ( j ) of the LP relaxation. Let ( K^j(i) ) be the set of fleet types ( k ) such that ( X^{j,LP}</em>{k,i} &gt; 0 )</td>
</tr>
<tr>
<td>Step 2:</td>
<td>Let ( \text{CAP}<em>i = \min</em>{k \in K^j(i)} {SEATS_k} ), ( \overline{\text{CAP}}<em>i = \max</em>{k \in K^j(i)} {SEATS_k} ), and reclassify flight legs and passengers.</td>
</tr>
<tr>
<td>Step 3:</td>
<td>Set ( j = j + 1 ).</td>
</tr>
<tr>
<td>Step 4:</td>
<td>If the termination criteria is met, STOP! Otherwise, go to Step 1.</td>
</tr>
</tbody>
</table>

By reclassifying flight legs, some may no longer be capacitated and will thus be removed from the set of capacitated flight legs, \( \overline{L} \).

5.8 Generalizations to Demand Uncertainty and Recapture

This chapter emphasizes the situation in which there is no recapture and deterministic demand. However, we present some thoughts in this section on the generalization of some ideas as they relate to the practical issues of demand uncertainty and recapture.

5.8.1 Fare Allocation and Recapture

To incorporate recapture in traditional fleet assignment solvers, the airlines currently use a simple model. For each flight leg (or market or itinerary), a recapture rate, \( b_i \), is determined. If a passenger is spilled from a flight leg, this quantity represents an estimate by the airline of the probability that this passenger is recaptured somewhere else in the system. The actual mechanics of determining if there is capacity on these alternatives is not considered. The spill
fare for this flight leg is then calculated by:

\[ \text{fare}_{i}^{op} = \kappa \cdot \overline{\text{fare}}_{i}(1 - b_i). \]  

(5.34)

Another approach is to assign some of the fare of a passenger to some possible alternative itineraries. If the probability a specific passenger is spilled from an itinerary is \( \text{prob}_p \) (Equation 5.19), then one could create \( \text{prob}_p \overline{x}_p D_p \) “phantom” passengers on itinerary \( r \) to represent the potential of capturing spilled passengers.

### 5.8.2 Flight Leg Classification

When determining the classification of flight legs, we compared the potential capacity of the flight leg to the demand. If the demand is a distribution and not a single number, then we must clarify our definition of “capacitated.” Assume that each itinerary has estimates of the mean number of passengers, \( \overline{D}_p \), and the variance, \( \sigma_p^2 \). If we further assume that the demand on each itinerary is independent, then the estimate for the mean number of passengers on flight leg \( i \), \( Q_i \), is \( \sum_{p \in P} \delta_i^p \overline{D}_p \) and the estimate for the variance, \( \sigma_i^2 \), is \( \sum_{p \in P} \delta_i^p \sigma_p^2 \).

Although we could define an uncapacitated flight leg as \( \overline{Q}_i \leq \text{CAP}_i \), a better definition of uncapacitated might be \( \overline{Q}_i + \sigma_i \leq \text{CAP}_i \). If \( Q_i \sim N\left( \overline{Q}_i, \sigma_i^2 \right) \), we are guaranteed that at least 83% of the time the flight leg will not cause spill. Likewise an overcapacitated flight leg could be defined as a flight leg such that \( \overline{Q}_i - \sigma_i > \text{CAP}_i \).

### 5.8.3 Integrated Spill Model

The integrated spill model as stated only works in the case when there is deterministic demand. However, the generalization of the model to uncertain demand becomes a single flight leg seat inventory control problem. Belobaba and Farkas [10] propose this for spill models.

### 5.9 Conclusions

In this chapter, we study the specific problem of determining the optimal fleeting given a specific instance of demand. To model this problem exactly requires a new large formulation (Chapter
6) which is difficult to solve. We compare commonly used techniques in the old formulation to determine which performs better. We also present a number of new approaches that perform well.

There are a number of important conclusions that result from the research in this chapter, namely:

1. To determine the quality of a fleeting, it is important that the fleetings are measured by some post-processing procedure. Whether it is a linear program such as the passenger mix model or an extensive simulation, comparing different fleetings based on these post-processing contribution solvers is crucial to determining good assignments. Logically, the more accurate the attainable contribution solver is to reality, the better the assignments.

2. The representative fare spill model that is commonly in use by the airlines performs relatively poorly. Consistently, the integrated spill model outperforms the RFSM, sometimes on the order of over $30M/year.

3. When using the ISM, the choice of fare allocation schemes is still quite important. Allocation schemes that assign at least all of the fare to the capacitated legs performs well, and those that assign some fare to all of the capacitated legs in a passenger itinerary perform the best.

4. The commonly used practice of pro-rating the fare on different flight legs of a passenger itinerary by using mileage as an estimate for costs performed quite poorly, especially in the case when using the ISM. The benefits of discontinuing this practice tied to accounting principles are quite clear.

5. There is some relationship between the characteristics of the demand data and the performance of various allocation schemes. Specifically, the ISM outperforms the RFSM especially at high demand factors and low passenger connectivity ratios.

6. An interesting new formulation of the fleet assignment model deserves some attention. One can more accurately capture the network effects by grouping flight legs into variables. Instead of the decision variable being a fleet type to an arc, the decision variable is a
combination of many fleet types to many flight arcs that are connected based on passenger itineraries. This model has some use in modeling difficult operational constraints as well.
Chapter 6

The Combined Fleet Assignment and Passenger Mix Model

6.1 Overview

In Chapter 5, we develop methods to solve the fleet assignment problem by sequentially (and sometimes iteratively) running a standard fleet assignment solver and an attainable contribution solver. If the attainable contribution problem is a linear program (such as our chosen passenger mix model) then we can combine it with the fleet assignment model. The result is a very large mixed integer program with some interesting properties. In this chapter, we describe some ways to solve this combined model. Also, we use this model to perform analysis in order to gain insights about the relationship between the fleet assignment model and passenger itineraries.

6.2 Itinerary-Based Airline Fleet Assignment

We define the following classifications of airline fleet assignment. The problem of solving an airline fleet assignment that maximizes contribution when the unconstrained demand estimates are broken out by itineraries is called the itinerary-based airline fleet assignment problem. We assume that the unconstrained demand for each itinerary is independent of the fleet assignment decision.

Let’s consider the passenger choice model presented in Figure 3-1. A more accurate (and
much more difficult) model considers unconstrained demand for an entire market broken out by airline instead of itineraries, i.e., we incorporate the itinerary choice model of passengers as well (the passengers have made their choice of airlines). We define this model as *market-based airline fleet assignment* or *OD-based airline fleet assignment*. If we accurately capture the network effects associated with recapture, the itinerary-based airline fleet assignment model and the market-based airline fleet assignment model converge to similar solutions.

A further enhancement uses demand estimates for a market instead of an itinerary. To accurately solve this problem, we must consider game theory. For example, if we believe that a competitor is assigning small fleet types in a specific market, we might want to counter this with large aircraft to steal demand. By incorporating an airline choice model into the fleet assignment problem, we approach the globally optimal solution. We define this model as *competition-based airline fleet assignment*. The amount of information needed for modeling this situation is large, not only in terms of passenger behavior, but in terms of the behavior of competitors. Even if this information is accessible, the chance of solving such a model with today’s technology is quite small.

The basic idea of itinerary-based airline fleet assignment is discussed by many in the airline industry, however, few have attempted to actually solve the problem. Parkas [21] proposes a linear programming formulation that combines the traditional fleet assignment model and a passenger mix model with no recapture. He presents two different procedures for solving the model. The first procedure is a column generation approach. However, the sub-problem that he uses is the basic fleet assignment problem, therefore each column is a specific fleeting. The second approach is to partition the flight leg sets into sub-networks where the network effects of multi-leg itineraries are present. No computational experiments on real airline networks are performed with either of these two models.

Erdmann et al. [20] present a sequential fleet assignment, passenger mix model similar to that presented in Chapter 5. Their work is done with Lufthansa airlines for whom difficulties arise that do not appear in the United States daily domestic problem. Specifically, weekly flight plans and the large number of fleet types at Lufthansa result in a fleet assignment problem with over 8,000 flights, 35 fleets and a half million itineraries. Their work proposes various solution approaches, including Lagrangean relaxation.

152
6.3 The Formulation

In this section, we present the general formulation of itinerary-based airline fleet assignment. We propose a mixed integer linear program that has deterministic demand estimates. By combining FAM and PMIX, we formulate the itinerary-based airline fleet assignment model, which is similar to what Farkas does in his thesis [21]. However, this formulation includes the partial recapture of passengers, where Farkas does not consider passenger recapture.

(IFAM)

\[
\text{maximize} \quad - \sum_{i \in L} \sum_{k \in K} \tilde{c}_{k,i} X_{k,i} + \sum_{(r,g) \in (P,F)} \text{fare}_{r,g} \sum_{(p,f) \in (P,F)} x_{p,f}^{r,g}
\]

subject to:

(cover) \quad \sum_{k \in K} X_{k,i} = 1 \quad \forall i \in L

(count) \quad \sum_{r \in C(k)} X_{k,i} + \sum_{o \in A} Y_{k,o,t}^+ \leq n_k \quad \forall k \in K

(balance) \quad \sum_{i \in O(k,o,t)} X_{k,i} + Y_{k,o,t^+} - \sum_{i \in I(k,o,t)} X_{k,i} - Y_{k,o,t^-} = 0 \quad \forall \{k, o, t\} \in N

(capacity) \quad - \sum_{k \in K} CAP_{k,j} X_{k,i} + \sum_{(r,g) \in (P,F)} \sum_{(p,f) \in (P,F)} \delta_{r}^{r,g} x_{p,f}^{r,g} \leq 0 \quad \forall (i, j) \in (L, J)

(limit) \quad \sum_{(r,g) \in (P,F)} x_{p,f}^{r,g} \leq D_{p,f} \quad \forall (p, f) \in (P, F)

\[X_{k,i} \in \{0, 1\}; Y_{k,o,t^+}, x_{p,f}^{r,g} \geq 0.\]

6.3.1 An Upper Bound

We prove that the objective value of IFAM is an upper bound to the contribution that the airline can achieve under the same assumptions. The proof is similar to that of Theorem 4 from Chapter 4.
Let’s assume that there is no recapture and, therefore, all of the $b$-values are zero except $b_{p,f}^J = 1$. As in Chapter 4, we can assume that $D_{p,f}$ is a random variable which represents the demand for the IFC $(p,f)$, with $d_{p,f}$ being a realization of this demand. Likewise $\mathbf{D}$ is the vector of these random variables, and $\mathbf{d}$ denotes a vector of realization for one instance of demand.

**Theorem 5** Let $Z_{IFAM}(E[\mathbf{D}])$ be the optimal objective value of IFAM, when we use average demand. This value is an upper bound to the expected contribution that can be achieved by the airline independent of the fleeting decision.

**Proof:** Let $F$ be the set of all possible feasible fleetings. For fleeting $X$, let $Z_{PMIX}(E[\mathbf{D}], X)$ be the optimal objective value of PMIX when we use fleeting $X$ and the average demand. Also, let $Z_{IC}(\mathbf{D}, X)$ be the random variable associated with the contribution from an arbitrary seat inventory control policy, IC, for fleeting $X$ with the vector of demand random variables $\mathbf{D}$. From Theorem 4 in Chapter 4, we see that $Z_{PMIX}(E[\mathbf{D}], X) \geq E[Z_{IC}(\mathbf{D}, X)]$ for all fleetings $X$ and any control policy IC. Also, $Z_{PMIX}(E[\mathbf{D}], X)$ is just the objective function of $(IFAM^*)$ (excluding the operating costs) for the fleeting $X$. Therefore

$$Z_{PMIX}(E[\mathbf{D}], X) - \sum_{i \in L} \sum_{k \in K} \tilde{c}_{k,i} X_{k,i} \geq E[Z_{IC}(\mathbf{D}, X)] - \sum_{i \in L} \sum_{k \in K} \tilde{c}_{k,i} X_{k,i}, \quad (6.2)$$

for all fleetings $X \in F$. Let’s take the maximum of both sides over the set $F$.

$$Z_{IFAM}(E[\mathbf{D}]) = \max_{X \in F} \left\{ Z_{PMIX}(E[\mathbf{D}], X) - \sum_{i \in L} \sum_{k \in K} \tilde{c}_{k,i} X_{k,i} \right\}, \quad (6.3)$$

$$\geq \max_{X \in F} \left\{ E[Z_{IC}(\mathbf{D}, X)] - \sum_{i \in L} \sum_{k \in K} \tilde{c}_{k,i} X_{k,i} \right\}. \quad (6.4)$$

The far right side is the best expected contribution that can be achieved.□

**6.3.2 Solution Methods**

In the next few sections, we present some different ways that the combined model can be solved. We begin with a direct solution approach where we solve the LP relaxation of the problem,
then use a branch and bound enumeration strategy to achieve the optimal solution to the mixed integer program. We also present techniques to achieve an upper bound and solve the problem using Lagrange multipliers.

6.4 The Direct Solution Approach

Unless otherwise noted, we consider the case where all of the assumptions expressed at the beginning of Chapter 5 are in place, i.e., deterministic demand, no recapture, and one fare class. We present the formulation using the keypath variables from Chapter 3.

\[(IFAM^*)\]

\[
\text{minimize } \sum_{i \in L} \sum_{k \in K} c_{k,i} x_{k,i} + \sum_{p \in P} \text{fare}_p t_p^- 
\]

subject to:

(cover) \[\sum_{k \in K} x_{k,i} = 1 \quad \forall i \in L\]

(count) \[\sum_{i \in C(k)} x_{k,i} + \sum_{o \in A} y_{k,o,t+} \leq n_k \quad \forall k \in K\] \hspace{1cm} (6.5)

(balance) \[\sum_{i \in O(k,o,t)} x_{k,i} + y_{k,o,t+} - \sum_{i \in I(k,o,t)} x_{k,i} - y_{k,o,t-} = 0 \quad \forall \{k,o,t\} \in N\]

(capacity) \[\sum_{k \in K} \text{SEATS}_k x_{k,i} + \sum_{p \in P} \delta_{i}^- t_p^- \geq Q_i \quad \forall i \in L\]

\[x_{k,i} \in \{0,1\}; y_{k,o,t+}, 0 \leq t_p^- \leq D_p\]

6.4.1 The LP Relaxation

Given the fast run times of traditional FAM and PMIX that have been accomplished for realistically sized problems, it follows that solving the LP relaxation of IFAM\(^*\) is a possibility. Table 6.1 presents the solution times for the LP relaxation of IFAM\(^*\). For a basis of comparison,
we show the LP relaxation solution times for traditional FAM on the same machine. We use the EF allocation scheme with the optimal spill fare to average fare ratio for that scenario as the basis.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Total # of Flights</th>
<th># of Integral Flights</th>
<th>Solution Time (CPU sec)</th>
<th># of Integral Flights</th>
<th>Solution Time (CPU sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>157</td>
<td>151</td>
<td>0.11</td>
<td>121</td>
<td>1.49</td>
</tr>
<tr>
<td>F97-4A</td>
<td>431</td>
<td>431</td>
<td>2.24</td>
<td>275</td>
<td>11.98</td>
</tr>
<tr>
<td>F97-6A</td>
<td>823</td>
<td>677</td>
<td>32.57</td>
<td>530</td>
<td>116.58</td>
</tr>
<tr>
<td>F97-9</td>
<td>2,044</td>
<td>1,634</td>
<td>722.24</td>
<td>1,385</td>
<td>2,989.14</td>
</tr>
<tr>
<td>A97-3A</td>
<td>173</td>
<td>160</td>
<td>0.15</td>
<td>83</td>
<td>1.66</td>
</tr>
<tr>
<td>A97-4A</td>
<td>485</td>
<td>437</td>
<td>1.79</td>
<td>293</td>
<td>15.78</td>
</tr>
<tr>
<td>A97-6A</td>
<td>877</td>
<td>857</td>
<td>29.85</td>
<td>450</td>
<td>171.33</td>
</tr>
<tr>
<td>A97-9</td>
<td>1,888</td>
<td>1,595</td>
<td>442.37</td>
<td>1,173</td>
<td>3,076.22</td>
</tr>
</tbody>
</table>

Table 6.1: The number of integral flights and the LP relaxation solution time for traditional FAM model and the itinerary-based FAM

In Table 6.1, we also present the number of non-integral flights, i.e., flights that have been partially assigned to more than one fleet type. Typically, the number of integer variables that are fractional in the optimal solution of the LP relaxation is a good indication of the amount of time that is spent in the branch and bound tree. This case is no exception. For FAM, the amount of time in the branch and bound tree ranges from a couple of seconds for the smaller problems to a couple of minutes for the full-sized problems. However, for IFAM* the time spent in the branch and bound tree is on the order of a couple of CPU days for just the small problems. In the next section, we present the reasons for the increase in the number of fractional variable values followed by methods to improve this situation.

### 6.4.2 Causes of Fractionality

The increased fractionality in the solution — beyond that of traditional FAM — of the IFAM* formulation is caused by the combination of the following sets of constraints:

\[
\sum_{k \in K} X_{k,t} = 1, \quad (6.6)
\]

\[
\sum_{k \in K} SEATS_k X_{k,t} + \sum_{p \in P} \delta_{t_p} \geq Q_t, \quad (6.7)
\]
For the analysis in this section, we ignore the fleet size and the aircraft flow balance constraints, i.e., we look at the optimal fleet assignment for each specific leg in isolation of other flight legs. In the traditional fleet assignment model, an optimal fleet assignment in isolation of other flight legs is always integral. However, we show that there is a tendency in the combined fleet assignment model to assign fractional fleet types to achieve the optimal solution.

Matching of Capacity to Demand

We examine the case when a flight leg is potentially capacitated, i.e., there are fleet types $k_1$ and $k_2$ such that $SEATS_{k_1} < Q_i < SEATS_{k_2}$. Typically, there is a positive correlation between the operating cost of an aircraft and its seating capacity, so we assume $\tilde{c}_{k_1,i} < \tilde{c}_{k_2,i}$. In isolation of other flight legs, the optimal assignment is fractional so that the capacity constraint is binding:

$$X_{k_1,i}^* = \frac{SEATS_{k_2} - Q_i}{SEATS_{k_2} - SEATS_{k_1}}, \text{ and}$$

$$X_{k_2,i}^* = \frac{Q_i - SEATS_{k_1}}{SEATS_{k_2} - SEATS_{k_1}}.$$  \hspace{1cm} (6.8)

The Optimal Aircraft Frontier and the Marginal Cost of Capacity

Fractional assignments also occur when the flight leg is overcapacitated. One source of fractionality is caused by the marginal cost of capacity. In the two-fleet problem, with fleet types $k_1$ and $k_2$, the marginal cost of capacity on each flight leg is measured by the following expression:

$$\frac{\tilde{c}_{k_2,i} - \tilde{c}_{k_1,i}}{SEATS_{k_2} - SEATS_{k_1}}.$$  \hspace{1cm} (6.10)

To minimize costs, we spill passengers that have a fare less than the above expression.

<table>
<thead>
<tr>
<th>Fleet Type</th>
<th>Number of Seats</th>
<th>Operation Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>108</td>
<td>$11,250</td>
</tr>
<tr>
<td>B</td>
<td>120</td>
<td>$12,000</td>
</tr>
<tr>
<td>C</td>
<td>144</td>
<td>$16,250</td>
</tr>
<tr>
<td>D</td>
<td>188</td>
<td>$19,750</td>
</tr>
</tbody>
</table>

Table 6.2: The operation cost and the number of seats for a specific flight leg

When there are more than two fleet types, the marginal cost of capacity continues to play
a role in causing fractional solutions. As we have more fleet types, a piecewise linear function outlines a frontier. This frontier represents the cheapest method to “create” an airplane of desired capacity. Using the operation costs and capacities in Table 6.2, we create Figure 6-1, which shows all of the possible two fleet type convex combinations of capacity. The minimal linear pieces, which represent the convex combination between fleet types A and B and the convex combination between fleet types B and D, are the optimal aircraft frontier. To create an airplane of capacity 144, the optimal solution is not to assign fleet type C, but to have a convex combination of B and D. The slope of the lines in the optimal aircraft frontier is the marginal cost of capacity. Therefore, the marginal cost of capacity from 108 to 120 seats is \((\$12,000.16 - \$11,250)/(120 - 108) = 62.50\) and from 120 to 188 seats is \((\$19,750 - \$12,000)/(188 - 120) = 113.97\).

![Figure 6-1: The optimal aircraft frontier](image)

In Figure 6-1, we consider only convex combinations of two fleet types. We do not need to
consider a convex combination of three fleet types.

**Theorem 6** For each desired capacity, there is always a point on the optimal aircraft frontier which has no more than two fleet types.

**Proof:** We define the optimal aircraft frontier as a function \( OAF \). Let \( OAF(CAP_i) \) be the operation cost of the point on the optimal aircraft frontier for the desired capacity, \( CAP_i \). The value \( OAF(CAP_i) \) is the objective value of the following program:

\[
\begin{align*}
\text{minimize} & \quad \sum_{i \in L} \sum_{k \in K} c_{k,i} X_{k,i} \\
\text{subject to:} & \\
& \sum_{k \in K} X_{k,i} = 1 \quad (6.11) \\
& \sum_{k \in K} SEAT_{k} X_{k,i} - s = CAP_i \\
& s, X_{k,i} \geq 0 \quad \forall (k, i)
\end{align*}
\]

The \( s \) variable is a slack variable that is used in situations when a smaller aircraft actually has a higher operating cost. For example, if the operation costs of fleet types A and B are switched, then the optimal aircraft of 108 seats is fleet type B. Any vector that is a basic feasible solution to Program 6.11, which is in standard form, has at most two(2) non-zero elements (an application of Theorem 2.4 in Bertsimas et al.[13]).

The Convexity of the Optimal Aircraft Frontier

**Theorem 7** The optimal aircraft frontier, \( OAF \), is a convex function.

**Proof:** We prove by contradiction. If \( OAF \) is not a convex function, then we can find three arbitrary points, \( CAP \), \( a \) and, \( b \), (\( a < CAP < b \)) such that the following two conditions hold:
\[ \tau a + (1 - \tau)b = \text{CAP}, \text{ where } \tau \in (0, 1) \]

\[ OAF(\text{CAP}) > \tau OAF(a) + (1 - \tau)OAF(b) \]

We select \( a \) as the convex combination of two fleet types (by Theorem 6), \( m \) and \( n \), so that

\[ a = \mu_1SEATS_m + (1 - \mu_1)SEATS_n, \quad (6.12) \]

and

\[ OAF(a) = \mu_1c_m + (1 - \mu_1)c_n. \quad (6.13) \]

Likewise, we select \( b \) as the convex combination of two fleet types (by Theorem 6), \( p \) and \( q \), so that

\[ b = \mu_2SEATS_p + (1 - \mu_2)SEATS_q, \quad (6.14) \]

and

\[ OAF(b) = \mu_2c_p + (1 - \mu_2)c_q. \quad (6.15) \]

Note that \( \mu_1, \mu_2 \in [0, 1] \) and it is possible that \( m = p \) and \( q = n \) or \( n = p \).

A valid solution for Program 6.11 given capacity \( \text{CAP} \), is \( X_m = \tau \mu_1, X_n = \tau (1 - \mu_1), \)
\( X_p = \tau \mu_2, \) and \( X_q = \tau (1 - \mu_2). \) The objective value of this solution is \( \tau OAF(a) + (1 - \tau)OAF(b). \) Since this is a valid solution and not necessarily the optimal solution,

\[ OAF(\text{CAP}) \leq \tau OAF(a) + (1 - \tau)OAF(b), \quad (6.16) \]

contradicting our second condition. Therefore the optimal aircraft frontier is a convex function. □
Spill Causing Fractionality

We now see how the aircraft optimal frontier plays a part in causing fractional assignments. In our example from Table 6.2, assume that the demand consists of 100 passengers paying a fare of $200, 60 passengers paying a fare of $125, and 40 passengers paying a fare of $75, for a total of 200 passengers. Let’s initially assign the largest, most expensive aircraft (fleet type D). We have a capacity of 188 seats and \( X_D = 1.0 \). First, we spill twelve people (the $75 passengers) so that we do not violate the capacity constraint. If we move down the optimal aircraft frontier to 187 seats, we can decrease operating costs by $113.97 by assigning a convex combination of fleet types C and D, i.e., \( X_D = .9853 \), and \( X_B = .0147 \). Since we are “removing” one seat, we must spill another passenger (another $75 passenger), losing $75 in revenue. Therefore, we save $38.97 for each seat of capacity that is removed. We continue doing this until we remove all 40 $75 passengers. The optimal solution is to partially assign fleet types B and D (\( X_B = .4118 \) and \( X_D = .5882 \)) such that we have a capacity of 160 seats and spill all 40 $75 passengers.

However, spill doesn’t always cause fractionality. Suppose the middle group of 60 passengers pay a fare of $100 instead of $125. We will spill these passengers until the cost of removing seats goes from $113.97 to $62.50. Specifically, this point is when we have fully assigned fleet type B. If this is the case, the optimal solution is to assign fleet type B fully to the flight leg (\( X_B = 1.0 \)) and spill all 40 $75 passengers and 40 $100 passengers.

In the real problem with aircraft balance and fleet size constraints, this analysis is much more difficult as we consider the entire network. The goal of balancing aircraft flow plays a crucial part in this analysis. Also, the multi-leg nature of passenger itineraries has some impact. However, this analysis shows why the number of integral flights decreases drastically when combining the fleet assignment model with the passenger mix model.

6.4.3 Achieving an Integer Solution

In Figure 6-2, we present a flowchart of the direct solution procedure. We first formulate the combined fleet assignment model consisting of the traditional fleet assignment model variables and the passenger mix variables. After the model has been formulated, we reduce coefficient in the formulation. We then solve the LP relaxation of the model. The two dashed boxes in Figure 6-2 represent procedures that are employed if we are considering recapture. These
Figure 6-2: The direct solution approach for the combined fleet assignment and passenger mix model

...procedures are similar to the addition of violated rows and negatively reduced columns for the passenger mix model, described in Chapter 3. There is a series of cuts that are added iteratively which help tighten the LP relaxation. Once all cuts are added, the process enters a branch & bound enumeration procedure to find an integer solution. After entering the branch and bound enumeration step, we no longer add any cuts. The next sections discuss the cuts in more detail.

Coefficient Reduction (Pre-processing)

One way to alleviate some of the fractionality caused by matching capacity to demand is to reduce all coefficients in which $SEATS_k > Q_i$ in Inequality 6.7. This method does not remove any integer solutions, but does eliminate some fractional solutions. Notice that this reduction is much weaker if we consider recapture. With recapture, to ensure that no integer solutions are removed, we compare the capacity to $Q_i + \sum_{p \in P} \sum_{q \in P} \delta_{p,q} t_p t_q$ instead of just $Q_i$. Typically the number of passengers that potentially can be recaptured onto a flight leg is even greater than the unconstrained demand.

After we have reduced the coefficients, if no spill occurs on the flight leg after the LP relaxation is solved, i.e., all $t_p$ variables are at zero, then this constraint no longer causes
fractionality. In isolation of other flight legs, the optimal assignment is the fleet type with capacity $SEATs_k$ greater than or equal to the unconstrained demand, $Q_i$, with the smallest operating cost, $c_{ik}$, $c_{k,i}$. Table 6.3 compares the number of integral flights in the optimal solution of the LP relaxation in the cases with and without this pre-processing coefficient reduction. The reduction has increased the number of integral flights at the end of the LP relaxation.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Total Number of Flight Legs</th>
<th>IFAM* without reduction</th>
<th>IFAM* with reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>F97-3A</td>
<td>157</td>
<td>121</td>
<td>134</td>
</tr>
<tr>
<td>F97-4A</td>
<td>431</td>
<td>275</td>
<td>375</td>
</tr>
<tr>
<td>F97-6A</td>
<td>823</td>
<td>530</td>
<td>620</td>
</tr>
<tr>
<td>F97-9</td>
<td>2,044</td>
<td>1,385</td>
<td>1,470</td>
</tr>
<tr>
<td>A97-3A</td>
<td>173</td>
<td>83</td>
<td>101</td>
</tr>
<tr>
<td>A97-4A</td>
<td>485</td>
<td>293</td>
<td>331</td>
</tr>
<tr>
<td>A97-6A</td>
<td>877</td>
<td>450</td>
<td>514</td>
</tr>
<tr>
<td>A97-9</td>
<td>1,888</td>
<td>1,173</td>
<td>1,274</td>
</tr>
</tbody>
</table>

Table 6.3: The number of integral flights after the initial LP relaxation with and without coefficient reduction

This reduction has some interesting ramifications to the optimal aircraft frontier. Let's examine the example above. Suppose that there are no $75 passengers, so the demand is 100 $200 passengers and 60 $125 passengers. Since the maximum number of passengers is 160 on this flight leg, we reduce the number of seats of fleet type D from 188 to 160. Figure 6-3 shows the new optimal aircraft frontier. Notice that the optimal aircraft frontier has changed shape. The marginal cost of capacity is now $218.75 from 144 to 160 and $177.08 from 120 to 144. From 108 to 120 seats, it remains $62.50. In this case, the optimal solution is to assign fleet type B fully, while spilling 40 $125 passengers. If we did not perform coefficient reduction, the optimal solution would be the same as previously, partially assigning fleet type B and fleet type D. So, even in the case where there is spill, this cut might increase the integrality of the LP relaxation.

Referring back to Figure 6-2, we ignore the iterative cut phase. Once the LP relaxation is solved, the solution procedure uses branch and bound to find an integer solution. We find an optimal integer solution for the 3 fleet problems. For all other problems, no integer solutions are found except poor quality initial integer solutions found at the start of branch and bound,
Figure 6-3: The optimal aircraft frontier after coefficient reduction.

when an initial depth first search is performed. In the next section, we present a set of cuts that we add iteratively that allows us to find optimal (or at least near-optimal) solutions with branch and bound.

**Additional Cuts**

Table 6.3 shows that using coefficient reduction can greatly increase the number of integral flights in the LP relaxation. However, because of the marginal cost of capacity and aircraft balance, there might be numerous fractional valued flight legs. We impose iteratively some cuts to counter fractionality when spill occurs. After a run of the LP relaxation, we can rewrite a capacity constraint as:

\[
SEATS_1 X_{1,i} + SEATS_2 X_{2,i} + \sum_{t=0}^{\delta_i} \delta t_p + \sum_{t>0}^{\delta_i} \delta t_p \geq Q_i,
\]  

(6.17)
where the spill variables are separated into two sets: those that are strictly positive (i.e., \( t_p > 0 \)) and those that are equal to zero (\( t_p = 0 \)). Say that \( SEATS_2 \) is equal to \( Q_i \) due to coefficient reduction, then in isolation of other flight legs, the optimal assignment \((X_{1,i}^*, X_{2,i}^*)\) satisfies the following two equations:

\[
SEATS_1 X_{1,i}^* + Q_i X_{2,i}^* = \max \left\{ SEATS_1, Q_i - \sum_{t>0} \delta_{i,t} t_p \right\}, \quad \text{and} \quad (6.18) \\
X_{1,i}^* + X_{2,i}^* = 1.0. \quad (6.19)
\]

The term \( \sum_{t>0} \delta_{i,t} t_p \) is strictly positive. If the spill is such that

\[
\sum_{t>0} \delta_{i,t} t_p < Q_i - SEATS_1, \quad (6.20)
\]

then there is a fractional assignment. To eliminate this fractional solution, we want to ensure that we require more spill if the fleeting variable of fleet type 1 is strictly positive. The following cut will exclude the fractional solution defined by Equation 6.18:

\[
\left( SEATS_1 - Q_i + \sum_{t>0} \delta_{i,t} D_p \right) X_{1,i} + \sum_{t=0} \delta_{i,t} t_p \geq 0.0. \quad (6.21)
\]

The cut defined in Inequality 6.21 says the following: the total amount of potential spill from a flight leg for a set of spilled itineraries and a utilized fleet type on that leg must exceed or equal the differential between unconstrained demand on the leg and the number of seats for that fleet type. The reason that we don’t include these cuts in pre-processing is that there are a huge number of them. There are a number of ways to partition the spill variables, so we let the LP relaxation guide us in selecting the partition. In Table 6.4, we present the changes in the number of integral flights when these cuts are used and when they are omitted and the number of iterations that are performed before we stop adding these cuts. As the problem size gets larger, empirically, we notice a tailing effect with the addition of more cuts. There are a small number of flight legs with numerous itineraries, and many cuts are generated for these legs. However, these additional cuts do little to improve the performance. We heuristically impose a
termination criterion that no additional cuts be generated when the improvement between two successive iterations is less than $1/day.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Total Number of Flight Legs</th>
<th>Number of Integral Flight Legs</th>
<th>Number of Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>no cuts reduction all cuts</td>
<td></td>
</tr>
<tr>
<td>F97-3A</td>
<td>157</td>
<td>121 134</td>
<td>151 2</td>
</tr>
<tr>
<td>F97-4A</td>
<td>431</td>
<td>275 375</td>
<td>401 7</td>
</tr>
<tr>
<td>F97-6A</td>
<td>823</td>
<td>530 637</td>
<td>712 9</td>
</tr>
<tr>
<td>F97-9</td>
<td>2,044</td>
<td>1,385 1,470</td>
<td>1,624 8</td>
</tr>
<tr>
<td>A97-3A</td>
<td>173</td>
<td>83 101</td>
<td>120 6</td>
</tr>
<tr>
<td>A97-4A</td>
<td>485</td>
<td>293 331</td>
<td>375 10</td>
</tr>
<tr>
<td>A97-6A</td>
<td>877</td>
<td>450 514</td>
<td>608 11</td>
</tr>
<tr>
<td>A97-9</td>
<td>1,888</td>
<td>1,173 1,274</td>
<td>1,377 21</td>
</tr>
</tbody>
</table>

Table 6.4: The number of integral flights after the initial LP relaxation with both pre-processing and post-processing cuts

The Branch and Bound Tree

Once the LP relaxation with added cuts is solved, we use branch and bound to find an integer solution. The first phase of tree exploration is a depth first search to find a feasible solution. After some experimentation, the strong branching variable selection strategy in CPLEX achieves the best initial integer solution. The SOS branching strategy from Chapter 2 is used to create the different branches. Results of the branch and bound procedure are presented in Table 6.5.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Best Solution</th>
<th>1st Integer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time (CPU sec)</td>
<td>Contribution ($/day)</td>
</tr>
<tr>
<td>F97-3A</td>
<td>6.37</td>
<td>3,325,628</td>
</tr>
<tr>
<td>F97-4A</td>
<td>43.42</td>
<td>5,141,876</td>
</tr>
<tr>
<td>F97-6A</td>
<td>3438.57</td>
<td>7,216,325</td>
</tr>
<tr>
<td>F97-9</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>A97-3A</td>
<td>41.58</td>
<td>1,448,465</td>
</tr>
<tr>
<td>A97-4A</td>
<td>157.44</td>
<td>4,913,692</td>
</tr>
<tr>
<td>A97-6A</td>
<td>3408.92</td>
<td>8,044,039</td>
</tr>
<tr>
<td>A97-9</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 6.5: The performance of the branch and bound procedure

Due to the large physical size of the problem in the full-sized data set, the procedure does
not have enough memory to create the SOS data sets and enter the branch and bound tree. Therefore, we are unable to get integer solutions from the $IFAM^*$ formulation for the A97-9 and F97-9 data sets. However, the objective value of the LP relaxation provides an upper bound to the contribution.

### 6.4.4 Analysis of the Optimal Solution and Contribution

We compare the contribution using the combined approach presented in this chapter with the current state of practice typically used by airlines today (traditional $FAM$ with the $MF$ allocation scheme). This comparison will give an estimate of the maximum amount of contribution improvement that can be achieved by considering the network effects of passenger itineraries. In Table 6.6, we present the contribution for the $MF$ allocation scheme, the best $ISM$ contribution and the combined model, with boldface text denoting the best solution.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>$\tilde{Z}_{MF}$ ($/\text{day}$)</th>
<th>$Z_{S}^*$ ($/\text{day}$)</th>
<th>Best Method</th>
<th>$IFAM^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Best Solution</td>
</tr>
<tr>
<td>F97-3A</td>
<td>3,325,628</td>
<td>3,325,628</td>
<td>Many</td>
<td>3,325,628</td>
</tr>
<tr>
<td>F97-4A</td>
<td>5,136,779</td>
<td>5,141,876</td>
<td>SF,PF,GF</td>
<td>5,141,876</td>
</tr>
<tr>
<td>F97-6A</td>
<td>7,176,229</td>
<td>7,216,325</td>
<td>FS</td>
<td>7,216,325</td>
</tr>
<tr>
<td>F97-9</td>
<td>8,199,869</td>
<td>8,224,292</td>
<td>FS</td>
<td>NA</td>
</tr>
<tr>
<td>A97-3A</td>
<td>1,448,092</td>
<td>1,448,465</td>
<td>GS</td>
<td>1,448,465</td>
</tr>
<tr>
<td>A97-4A</td>
<td>4,901,197</td>
<td>4,913,692</td>
<td>PS</td>
<td>4,913,692</td>
</tr>
<tr>
<td>A97-6A</td>
<td>8,014,395</td>
<td>8,043,632</td>
<td>CES</td>
<td>8,044,039</td>
</tr>
<tr>
<td>A97-9</td>
<td>11,155,890</td>
<td><strong>11,247,114</strong></td>
<td>CES</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 6.6: A comparison of the contribution achieved with the state of the practice and the combined model

The first thing to notice from Table 6.6 is the effectiveness of the methods presented in Chapter 5. These methods achieve the true optimal solution in the six smallest scenarios. Using the LP relaxation (with added cuts) of the $IFAM^*$ formulation as an upper bound to the contribution, we see that the gap between the best method from Chapter 5 and the upper bound is $2,326$/day ($849$K/year) and $3,360$/day ($1.2M/year) for the F97-9 and A97-9 data set respectively. Since the $FS$, $PS$, and $CES$ methods were all within $750$/day of each other for the full-sized data sets, using any of these results in a solution that is within $4,110$/day ($1.5M/year) of the optimal solution. We check how these solutions compare when there is
demand uncertainty in Section 6.7.2.
Also, Table 6.6 shows us that the LP relaxation of the IFAM$^*$ formulation provides a good upper bound for itinerary-based fleet assignment. However, using this formulation to determine the fleeting raises some issues. For the full-sized problem, the physical size of the problem is on the order of 70MB of computer memory. As we include recapture, operational constraints, etc., the physical size of the problem will become a limiting factor on many workstations. Including recapture drastically increases the time required to solve the LP relaxation, similar to the increased time to solve just the passenger mix problem with recapture, presented in Chapter 3. Therefore, this formulation has many computational issues associated with it.

6.5 Lagrangean Methods

With large-scale problems, in the preceding sections we observe that solving the entire program directly is not always practical. Notice how much longer it took to solve the LP relaxation of IFAM$^*$ compared to traditional FAM in the preceding section. We are concerned with getting both upper and lower bounds to the expected contribution, so that we know how well our particular approach is doing. An obvious lower bound is the contribution from a fleeting found using the traditional fleet assignment model. One possible method of finding an upper bound is to relax some of the difficult constraints using Lagrangean relaxation. If we relax the capacity constraints, we have two separable problems.

1. The first problem is a traditional FAM with the objective coefficient for $X_{k,i}$ being:

   \[
   \tilde{c}_{k,i} - \lambda_i \text{SEATS}_k, \tag{6.22}
   \]

   where $\lambda_i$ is the dual price of the capacity constraints for flight leg $i$.

2. The second problem becomes a trivial problem which is solved by a simple greedy algorithm, since the new objective coefficient for a passenger flow variable, $t_p$ becomes:

   \[
   \text{fare}_p + \sum_{i \in L} \delta_i^p \lambda_i, \tag{6.23}
   \]

   168
Therefore, we set all passenger variables with negative coefficients to their upper bound and set passenger variables with nonnegative coefficients to their lower bound.

The flow chart of the solution procedure using Lagrangean relaxation is presented in Figure 6-4. The first step is to have some initial estimates of the Lagrangean multiplier, $\lambda_i$, for each flight leg $i$. For capacitated flight legs, we initially set the Lagrange multipliers to the average fare on the flight leg, while we set Lagrange multipliers associated with uncapacitated flight legs to zero. We then solve the two problems stated above. Combining the objective value of these two problems gives the Lagrangean dual value. The Lagrange multipliers are updated using subgradient optimization. The subgradient is the slack (or surplus) of the capacity constraints, i.e., the Lagrange multipliers for the $(j + 1)^{st}$ iteration is:

$$
\lambda_i^{j+1} = \max \left\{ 0, \lambda_i^j + \alpha^j \left( \sum_{k \in K} SEATS_k X_{k,i}^j + \sum_{p \in P} \delta_{i,p}^j \right) - Q_i \right\}, \quad (6.24)
$$

where $\alpha^j > 0$ is a multiplier for iteration $j$. The multiplier $\alpha^j$ has the following properties:

$$
\lim_{j \to \infty} \alpha^j = 0, \quad \text{and} \quad \sum_{j=1}^{\infty} \alpha^j = \infty. \quad (6.25)
$$

When using Lagrangean relaxation to solve problems, the resulting solution is not guaranteed to be feasible. In this instance, the resulting fleet assignment and collection of traffic variables may lead to capacity constraints being violated. At each iteration, however, a feasible fleet assignment is determined. We determine the quality of this feasible assignment by running the attainable contribution solver (the passenger mix model). For problems in air traffic management, Stock [40] has shown that this Lagrangean dual-primal approach may be an effective solution technique for certain problems.

In Figure 6-5, the first vertical line represents the time when the LP relaxation of IFAM* was solved for data set A97-4A. At this time, there is still a large gap between the best Lagrangean dual and primal values. As the number of iterations increases, we see that the dual
value cannot match the LP relaxation of \( IFAM^* \) with cuts. The problem that we experience is a common problem with Lagrangean relaxation. The multipliers take too long to converge to good values. This problem increases as we have a large number of multipliers, especially when there are many interactions between them as in our case.

Unfortunately, the feasible fleetings do not converge to good solution values either. The primal solution values display a chaotic behavior as the number of iterations grows; all of which are worse than the optimal solution. A near-optimal solution isn’t achieved until the 30th iteration, well after the time in which \( IFAM^* \) is solved to optimality.

Lagrangean relaxation does not seem to perform well in this case. Experiments on other data sets display the same pattern found in Figure 6-5. The LP relaxation of \( IFAM^* \) provides a much better upper bound sooner. However, as recapture and other constraints are added to the problem, Lagrangean relaxation may be the only way to find upper bounds if the LP relaxation of \( IFAM^* \) proves to be unsolvable due to memory limitations.
6.6 Special Case Optimality

In the previous sections, we describe the difficulties associated with solving IFAM* in the general case. If we incorporate multiple fare classes, recapture, and operational constraints, the size of the formulation presents a challenge for solving even the linear programming relaxation of the full-size problems typically encountered at major airlines. We notice that even in fairly small problems, achieving integer solutions might be difficult, caused by the simultaneous assignment of capacity and determination of traffic flow.

Empirically, the traditional fleet assignment model is easily solved, because we are not
explicitly considering the passenger flow. In this section, we investigate the special cases when
the traditional fleet assignment model along with the developments described in Chapter 5
guarantee the optimal assignment to the \( IFAM^* \) formulation.

6.6.1 One Capacitated Leg Passenger Itineraries

We first present the following theorem:

**Theorem 8** Under the assumptions of deterministic demand and no recapture, if every pas-
senger itinerary variable, \( t_p \), appears in at most one capacity constraint, then the optimal as-
signment of the traditional fleet assignment formulation FAM using spill costs calculated by the
integrated spill model (ISM) is the same as the optimal assignment of the \( IFAM^* \) formulation,
with identical optimal objective values.

**Proof:** To prove the theorem, it suffices to show that the following two conditions hold:

1. a feasible assignment of FAM is also a feasible assignment of \( IFAM^* \) and vice versa.

2. the objective value of each feasible fleet assignment in FAM is equal to the minimum
   objective value of each feasible fleet assignment of \( IFAM^* \).

Let \( \mathbf{X} \) be a feasible solution to FAM. Since the balance, count and cover constraints
are exactly the same in both formulations, the only way that \( \mathbf{X} \) is not a feasible solution is that somehow the capacity constraints are violated. However, we can always satisfy the capacity constraints by spilling all passengers (i.e., setting all \( t_p \)
variables to their upper bound \( D_p \)). Therefore, if \( \mathbf{X} \) is a feasible solution of FAM,
then it is also a feasible solution of \( IFAM^* \). The converse is true since a feasible assignment of \( IFAM^* \) satisfies all of the constraints in FAM, plus the capacity constrains of \( IFAM^* \). We have shown that condition 1 holds.

Let's examine a specific assignment of a fleet type to a flight leg, \( \mathbf{X}_{k,i} \), of a feasible
assignment, \( \mathbf{X} \). The objective coefficient of \( \mathbf{X}_{k,i} \) is equal to \( \bar{c}_{k,i} + c_{k,i}^{sp} \) where the first
term is the operation cost of the assignment and the second term is the approx-
imation of the spill costs. This approximation, \( c_{k,i}^{sp} \), is calculated by the following
program when using spill integration:

\[
\text{minimize} \quad \sum_{p \in P} fare_p t_p \\
\text{subject to:} \quad \sum_{p \in P} \delta_i^p t_p \geq Q_i - SEATS_k \\
0 \leq t_p \leq D_p \quad \forall p \in P.
\]  

(6.27)

If \( \bar{X} \) is the assignment for \( IFAM^* \) the rest of the problem becomes the passenger mix problem which is

\[
\text{minimize} \quad \sum_{p \in P} fare_p t_p \\
\text{subject to:} \quad \sum_{p \in P} \delta_i^p t_p \geq Q_i - \sum_{k \in K(i)} SEATS_k \bar{X}_{k,i} \quad \forall i \in L \\
0 \leq t_p \leq D_p \quad \forall p \in P.
\]  

(6.28)

However, since each \( t_p \) is only in one constraint, this can be solved as \(|L|\) different programs. The program is associated with flight leg \( i \) is the same program that is used to find \( c_{k,i}^{sp} \) when \( fare_i^p = fare_p \). Therefore, the optimal assignment of \( (FAM) \) using spill integration is equal to the optimal assignment of \( IFAM^* \) and they will have the same costs. \( \square \)

This preceding Theorem leads directly to the following corollary. As a reminder, Type 0 passengers have no capacitated flight legs in their itinerary, while Type 1 passengers have exactly one capacitated flight leg in their itinerary.

**Corollary 9** Under the assumptions of no recapture and deterministic demand, if all passengers are either Type 0 or Type 1, then the optimal assignment of the traditional fleet assignment model \( FAM \) with spill costs calculated using capacity-based fare allocation and the integrated spill model is the optimal assignment of the itinerary-based fleet assignment model \( IFAM^* \).

**Proof:** Consider a new formulation \( IFAM^{**} \) which is the same as \( IFAM^* \) except that all constraints associated with uncapacitated flight legs are removed. Since
these constraints are redundant, $IFAM^{**}$ is an equivalent formulation to $IFAM^*$. Since the formulation $IFAM^{**}$ only has capacitated flight legs in the capacity constraints, and each passenger itinerary has only one capacitated flight leg, each traffic variable is only in one constraint. Therefore, by Theorem 8, the optimal assignment of $IFAM^{**}$ is the same as the optimal assignment of $FAM$ with the same costs. Therefore, the optimal assignment of $FAM$ is equal to the optimal assignment of $IFAM^{**}$ which is equal to the optimal assignment of $IFAM^*$. □

If all passengers are either Type 0 or Type 1, then we can find the optimal solution of the itinerary-based airline fleet assignment, by simply solving the traditional fleet assignment model. The evidence in Chapter 5 suggests that roughly 5% of the passengers are not Type 0 nor Type 1. This theorem is further evidence of the importance of distributing passenger revenue among the flight legs based on capacity rather than costs.

From the above corollary, we can also state that if all passenger itineraries have one leg, then all passengers must be either Type 0 or Type 1, therefore traditional $FAM$ will give an optimal assignment in this case as well. If there are no connecting passengers, then the itinerary-based fleet assignment problem reduces to the traditional fleet assignment problem using the integrated spill model. However, notice that even with this simple problem with deterministic demand, we cannot make any guarantee of optimality using the representative fare spill model. There will be cases when just using the representative fare spill model will not give the optimal solution. As the spread of fares on each flight leg increases, i.e., the bigger the difference between the highest fare and the lowest fare, the worse the representative spill fare method becomes.

The deterministic linear program (6.27) calculates the spill costs, $c_{k,l}^{SP}$, for the specific assignment of fleet type $k$ to flight leg $i$. If demand is uncertain and all itineraries have one leg, then (6.27) is replaced by the leg-independent multiple fare class revenue management problem (see Chapter 4). Therefore, if we can solve the revenue management problem to optimality such that all estimated expected spill costs are accurate, then we can solve for the optimal fleet assignment in the case with uncertain demand and one leg itineraries.
6.6.2 The Optimality of the Multiple Flight Column Approach

Using the traditional fleet assignment model with modified objective costs can not guarantee optimality for many problem instances. By using the multiple flight column approach described in Chapter 5, we are striking a compromise between the traditional fleet assignment model and the itinerary-based fleet assignment model.

Recalling from Chapter 5 that the connectivity of two flight legs is defined as

\[
C_{i_1,i_2} = \sum_{p \in P} \delta_{i_1}^p \delta_{i_2}^p D_p,
\]

we present the following theorem.

**Theorem 10** The MFC approach guarantees the optimal solution of \(IFAM^*\) if the partition of flight legs has the following property: for any pair of capacititated flight legs, \(i_1\) and \(i_2\), if their connectivity is positive, then the two flight legs are in the same subset.

**Proof:** The proof is similar to that of Theorem 8. We will show that the following two conditions hold:

1. Every feasible assignment for \(FAM - MFC\) is also a feasible assignment for \(IFAM^*\) and conversely every feasible assignment for \(IFAM^*\) is also a feasible assignment for \(FAM - MFC\).

2. The objective value of each feasible fleet assignment in \(FAM - MFC\) is equal to the minimum objective value of each feasible fleet assignment of \(IFAM^*\).

The structure of the timeline network is the same in both \(FAM - MFC\) and \(FAM\). In both formulations we are selecting arcs to be used such that the count, balance and cover constraints are satisfied. In the \(FAM\) formulation, a decision variable represents selecting one arc, while in the \(FAM - MFC\) we are selecting a set of arcs. If a selection of arcs is feasible, then whether we selected these arcs one at a time or in sets has no bearing on feasibility. Therefore, condition 1 holds.

Let \(\hat{X}\) and \(\bar{X}\) be equivalent assignments from the \(FAM - MFC\) and \(IFAM^*\) approach, respectively. Let \(U\) be an arbitrary permutation from \(FAM - MFC\)
such that $\tilde{X}_U = 1.0$. Since $\tilde{X}$ and $\bar{X}$ are equivalent assignments, for every pair $(i,k) \in U$, then $\bar{X}_{k,i} = 1.0$ in the $IFAM^*$ formulation. The objective coefficient of $\tilde{X}_U$ is

$$Z_U^{PMIX} + \sum_{(k,i) \in U} \tilde{c}_{k,i}. \tag{6.30}$$

Obviously, $\tilde{X}_U \left( \sum_{(i,k) \in U} \tilde{c}_{k,i} \right) = \sum_{(i,k) \in U} \tilde{c}_{k,i} \bar{X}_{k,i}$.

Given $\bar{X}$, the remaining program for $IFAM^*$ is (removing all uncapacitated flight leg constraints):

$$\begin{align*}
\text{minimize} & \quad \sum_{p \in P} fare_p t_p \\
\text{subject to:} & \quad \sum_{p \in P} \delta^i_p t_p \geq Q_i - \sum_{k \in K(i)} SEATS_k \bar{X}_{k,i}, \quad \forall i \in \bar{L} \\
& \quad 0 \leq t_p \leq D_p, \quad \forall p \in P.
\end{align*} \tag{6.31}$$

Let's rewrite the above program in the following form:

$$\begin{align*}
\text{minimize} & \quad \sum_{p \in P} fare_p t_p \\
\text{subject to:} & \quad \sum_{p \in P} \delta^i_p t_p \geq Q_i - \sum_{k \in K(i)} SEATS_k \bar{X}_{k,i}, \quad \forall i \in L^1 \\
& \quad \sum_{p \in P} \delta^i_p t_p \geq Q_i - \sum_{k \in K(i)} SEATS_k \bar{X}_{k,i}, \quad \forall i \in L^2 \tag{6.32} \\
& \quad \vdots \\
& \quad \sum_{p \in P} \delta^i_p t_p \geq Q_i - \sum_{k \in K(i)} SEATS_k \bar{X}_{k,i}, \quad \forall i \in L^m \\
& \quad 0 \leq t_p \leq D_p, \quad \forall p \in P.
\end{align*}$$

Because of the property above, two flight legs that are connected will be in the same subset. Conversely, two flight legs in different subsets will not be connected.
Therefore we can solve each of the following programs separately:

\[
\text{minimize} \quad \sum_{p \in P(L^m)} f_\text{are}_p t_p \\
\text{subject to:} \quad \sum_{p \in P(L^m)} \delta_i^p t_p \geq Q_i - \sum_{k \in K(i)} S_\text{EATS}_k \bar{X}_k,i \quad \forall i \in L^m \\
0 \leq t_p \leq D_p \quad \forall p \in P(L^m),
\]  
(6.33)

for every subset, \(L^m\). The objective value of the program above is equal to

\[
\sum_{U \in U(L^m)} Z_U^{P_\text{MIX}} \hat{X}_U,
\]  
(6.34)

and summing up

\[
\sum_m \sum_{U \in U(L^m)} Z_U^{P_\text{MIX}} \hat{X}_U = \sum_{p \in P} f_\text{are}_p t_p^*.
\]  
(6.35)

so condition 2 above holds. \(\square\)

We can construct a partition that has the property of all capacititated connected flights are in the same subset. Obviously, the partition in which all flight legs are in the same subset has this property. However, we can use a modification of the heuristic described in Section 5.7.3. If we alter the heuristic such that we only perform a cut when the all-pairs minimum cut is equal to zero.

### 6.7 The Relative Importance of Network Effects and Demand Uncertainty

The conventional wisdom in regards to airline fleet assignment is that the uncertainty of demand is the most important attribute to capture in the decision process. While potentially important, capturing network effects is secondary. In this section, we will test this hypothesis. We will accomplish this by comparing how various fleetings perform in a simulation environment. Some
of the fleetings assume leg independence but capture uncertain demand, others consider network effects with deterministic demand.

6.7.1 The Experiment Description

Given the flight leg network and a set of characteristics of the demand patterns, four fleet assignments are determined and their performance is compared based on simulating a large number of realized demands and calculating which fleeting decision did better based on the mean contribution received. We do this for a number of different scenarios where we vary the level of network effects and the level of variance.

The Flight Network and Demand Patterns

We use the flight leg network and the base demand data estimates from the data set A97-3A. We create a number of scenarios by varying the percentage of connecting passengers and the level of variance. The percentage of the number of connecting passengers is changed by multiplying the demand of connecting passengers by one scalar and the demand of local passengers by another scalar. The desire is to keep the demand factor, \( DF \), at the same level. This process was described in Section 3.7.3. To change the level of variance we modify the standard deviation by a universal constant. Specifically for each itinerary \( p \), the expected demand, \( \overline{D}_p \) is given. For this experiment, we model the standard deviation as a linear function with respect to the square root of the mean of demand \([10]\), i.e.,

\[ \sigma_p = z \cdot \sqrt{\overline{D}_p}, \]  

(6.36)

where \( z \) is some chosen universal constant. The model that we use for the distribution of demand for each itinerary is an integral gamma density function (See Appendix A).

Determine the Fleetings

The first step of the test is to determine the four different fleetings. We will compare the following methods: \( IFAM^* \), \( PS \), \( MF \), and \( EF \). The \( IFAM^* \) method gives the optimal solution for the instance when demand is deterministic and equal to the expected demand for every
itinerary \( p \). The \( PS \) method gives near-optimal solutions for most of our experimental problems when we consider deterministic demand. The \( MF \) method is commonly used at the major airlines. Finally, the \( EF \) method is the best method in our previous experiments when limited to the \( RFSM \).

The generalization of Equation 5.3, which calculates the spill costs of assigning fleet type \( k \) to flight leg \( i \) for the \( RFSM \) model, is:

\[
e_{k,i}^{sp} = \kappa \cdot \text{fare}_{i} \cdot \sum_{n=\text{SEATS}_k}^{\infty} P(Q_i = n) \cdot \max \{0, (n - \text{SEATS}_k)\},
\]

where \( Q_i \) is the realized demand on flight leg \( i \). We approximate the realized demand by a Gaussian distribution with mean, \( \bar{Q}_i \), and variance, \( \sigma_i^2 \), found by the following equations:

\[
\bar{Q}_i = \sum_{p \in P} \delta_i^p \bar{D}_p,
\]
\[
\sigma_i^2 = \sum_{p \in P} \delta_i^p \sigma_p^2.
\]

Notice that Equation 6.37 simplifies to Equation 5.3, since in the deterministic case we assume that

\[
P(Q_i = x) = \begin{cases} 1 & x = \bar{Q}_i, \text{ and} \\ 0 & \text{otherwise}. \end{cases}
\]

Since the Gaussian is a continuous distribution, we approximate \( P(Q_i = n) \) by the following expression:

\[
P(Q_i = n) \approx \int_{n-0.5}^{n+0.5} \frac{e^{-|x-Q_i|^2/2\sigma_i^2}}{\sqrt{2\pi}\sigma_i} dx \approx \left( \frac{e^{-|n-Q_i-0.5|^2/2\sigma_i^2} + e^{-|n-Q_i+0.5|^2/2\sigma_i^2}}{2\sqrt{2\pi}\sigma_i} \right),
\]

where the last relationship comes from the trapezoidal approximation.

The average fare, \( \text{fare}_{i} \), is calculated the same as presented in Chapter 5, however, we substitute \( \bar{D}_p \) in place of \( D_p \), i.e., we determine a priori the average fare based on the expected
demand, e.g.,

$$\text{fare}_i^{MF} = \frac{\sum_{p \in P} m_p \text{fare}_{p_i} \sigma_i D_p}{\sum_{p \in P} \sigma_i D_p}.$$  \hspace{1cm} (6.42)

We use the optimal value of $\kappa$ for the data set A97-3A found in Chapter 5, i.e., 1.08 and 0.70 for the $EF$ and $MF$ methods, respectively.

The $SSP$ and the $IFAM^*$ fleetings are determined by the methods presented before with no changes; these two methods represent solving fleet assignment with regard for just the network effects, but not considering the effects associated with demand uncertainty. Notice that the fleetings for both methods will not change as we vary the level of variance, only as we vary the level of network effects. On the other hand, the $EF$ and $MF$ methods capture demand uncertainty, but do little to consider the leg dependent network effects.

**Comparison of the Fleetings**

To compare the different fleetings, we use simulation to create a large collection of different demand instances. The creation of each instance is described in Appendix A. The number of instances for each scenario varies, so that the mean square error associated with our estimate for the expected contribution is on the order of $\$1,000/day. Therefore, scenarios with high demand uncertainty will have more instances. Within each scenario, we use the same random number seed for each fleeting, therefore the performance of each fleeting decision is based on the same collection of demand instances. The contribution for each demand instance is calculated by the chosen attainable contribution solver, which in our case is the passenger mix model.

**Different Demand Characteristics**

We create nine distinct scenarios. Each scenario has a low, medium, and high level associated with the demand uncertainty and the network effects. To change the uncertainty of demand, we vary the Z-factor, i.e., the constant in Equation 6.36. To change the level of network effects for a data set, we vary the $PCR$. Table 6.7 shows the numerical levels associated with each characteristic. The $DF$ is kept constant at 0.70 for all scenarios.
6.7.2 Computational Experience

In Table 6.8, we show the mean contribution of each scenario for all four fleetings, which is also an estimate for the expected value of the contribution when it is viewed as a random variable. The number in parentheses is the mean squared error of the estimate of the true expected contribution.

Let's concentrate on just the \( EF \) and the \( IFAM ^ * \) solution since the \( EF \) solution always dominates the \( MF \) solution and the \( IFAM ^ * \) solution always dominates the \( PS \) solution. Let's define the relative performance of the \( EF \) and \( IFAM ^ * \) scheme, denoted \( RP ( EF, IFAM ^ * ) \), as:

\[
RP^j = Z^j_{EF} - Z^j_{IFAM^*}.
\] (6.43)

Table 6.9 shows the mean value of the relative performance and the mean squared error for that estimate of the expected relative performance. We are able to eliminate much of the variance associated with the estimates since the performance of each method is highly correlated for a specific instance of demand \( j \).

In terms of relative performance, a few trends are apparent:

1. As we increase the level of network effects by the increasing the \( PCR \), the \( IFAM ^ * \) solution performs better than the \( EF \) solution.

2. As we increase the variance in the demand data, the \( EF \) solution performs better than the \( IFAM ^ * \) solution. However, this effect seems minor compared to the trend associated with varying the \( PCR \) level.

Surprisingly, we find that the relative performance of the various methods is more dependent on the level of network effects than on the level of uncertainty. As we increase the level of

<table>
<thead>
<tr>
<th>Level</th>
<th>Demand Uncertainty (Z-factor)</th>
<th>Network Effects (PCR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.50</td>
<td>.10</td>
</tr>
<tr>
<td>Medium</td>
<td>1.25</td>
<td>.25</td>
</tr>
<tr>
<td>High</td>
<td>2.00</td>
<td>.40</td>
</tr>
</tbody>
</table>

Table 6.7: The values used for the different levels of the demand data characteristics
<table>
<thead>
<tr>
<th>Z-factor</th>
<th>Method</th>
<th>Low ($/day)</th>
<th>Medium ($/day)</th>
<th>High ($/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>EF</td>
<td>1,808,068 (1,070)</td>
<td>1,006,790 (1,097)</td>
<td>340,436 (1,042)</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>1,800,191 (1,075)</td>
<td>1,005,056 (1,104)</td>
<td>335,665 (1,050)</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>1,797,097 (1,045)</td>
<td>1,004,543 (1,067)</td>
<td>341,314 (1,024)</td>
</tr>
<tr>
<td></td>
<td>IFAM*</td>
<td>1,802,815 (1,040)</td>
<td>1,009,592 (1,097)</td>
<td><strong>346,286 (1,030)</strong></td>
</tr>
<tr>
<td>Med.</td>
<td>EF</td>
<td><strong>1,785,049 (973)</strong></td>
<td>994,894 (1,028)</td>
<td>330,541 (1,060)</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>1,776,855 (982)</td>
<td>993,810 (1,035)</td>
<td>326,210 (1,069)</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>1,774,257 (954)</td>
<td>991,905 (1,008)</td>
<td>331,768 (1,046)</td>
</tr>
<tr>
<td></td>
<td>IFAM*</td>
<td>1,779,199 (949)</td>
<td><strong>997,694 (1,029)</strong></td>
<td><strong>336,018 (1,049)</strong></td>
</tr>
<tr>
<td>High</td>
<td>EF</td>
<td><strong>1,767,812 (1,181)</strong></td>
<td>981,451 (1,115)</td>
<td>322,913 (1,071)</td>
</tr>
<tr>
<td></td>
<td>MF</td>
<td>1,758,954 (1,195)</td>
<td>980,948 (1,123)</td>
<td>318,940 (1,078)</td>
</tr>
<tr>
<td></td>
<td>PS</td>
<td>1,757,001 (1,160)</td>
<td>977,458 (1,092)</td>
<td>324,128 (1,055)</td>
</tr>
<tr>
<td></td>
<td>IFAM*</td>
<td>1,761,318 (1,158)</td>
<td><strong>984,336 (1,116)</strong></td>
<td><strong>327,426 (1,057)</strong></td>
</tr>
</tbody>
</table>

Table 6.8: The mean data set for A97-3A

<table>
<thead>
<tr>
<th>Z-factor</th>
<th>Low ($/day)</th>
<th>Medium ($/day)</th>
<th>High ($/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td></td>
<td>PCR Level</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>5,253 (110)</td>
<td>-2,802 (32)</td>
<td>-5,850 (115)</td>
</tr>
<tr>
<td>Medium</td>
<td>5,850 (100)</td>
<td>-2,800 (33)</td>
<td>-5,477 (120)</td>
</tr>
<tr>
<td>High</td>
<td>6,494 (114)</td>
<td>-2,885 (42)</td>
<td>-4,513 (112)</td>
</tr>
</tbody>
</table>

Table 6.9: The mean and mean squared error for the relative performance of EF and IFAM*
uncertainty in the demand, there is a slight increase in the performance of \( EF \) over \( IFAM^* \). This is more apparent at the low and high level of \( PCR \). For these two \( PCR \) levels, the increase is on the order of $400/day from low to medium level of demand uncertainty. The increase is roughly $800/day from the medium level to the high level of demand uncertainty. For the medium level of \( PCR \), there is very little change in the relative performance as we increase the variance of the demand. However, there is a large change in relative performance as the level of network effects change. For example, at the low level of network effects, the \( EF \) solution performs better than the \( IFAM^* \) solution on the order of $5000/day. However, with the medium level of network effects, the \( IFAM^* \) solution performs better than the \( EF \) solution by $2800/day and increases to $5,000/day as we increase amount of connecting passengers to the high level. While not necessarily conclusive, this experiment does suggest that network effects are important and that method used to determine the fleet assignment should incorporate these effects.

There is another interesting point to note from Table 6.8. Within each scenario, the same number of iterations are performed for each fleeting decision. Therefore, the mean squared error is proportional to the estimated standard deviation of the contribution. The \( PS \) and \( IFAM^* \) methods consistently have lower standard deviations than the \( RFSM \) methods. This is another unexpected result, since intuition suggests that a fleeting decision that considers demand uncertainty should have a lower variability in performance than one that does not.

### 6.8 Conclusion

In this chapter, we formulate the problem of determining the optimal fleet assignment given a specific instance of demand. By doing this, we develop a number of important conclusions:

- The optimal objective function value of this problem is an upper bound to the expected contribution determined by any other method.

- We demonstrate that this problem is similar to a number of difficult problems in the transportation industry. Certain aspects of the problem are also similar to the general network design problem and the knapsack problem.
• By solving the LP relaxation of this model, we show that some of the methods developed in Chapter 5 lead to either optimal or near-optimal solutions.

• We also demonstrate the difficulty of solving the problem with binary variables. We provide an analysis of a number of reasons why this formulation suffers from a large amount of fractionality as compared to the traditional formulation of the fleet assignment problem. We present a number of different approaches to solving this problem, including a direct solution approach, Lagrangean relaxation, and the heuristic methods in Chapter 5. We presented a number of cuts that make a direct solution approach effective.

• We prove that the traditional fleet assignment model guarantees optimal solutions to the IFAM* formulation in some special circumstances. Specifically, when the demand data only has passengers that travel on one capacitated leg. Also, we show that the multiple flight column formulation of the fleet assignment model guarantees the optimal solution to the IFAM* formulation, however, the number of variables required may be quite large.

• Finally, we have compared the relative importance of network effects and demand uncertainty. For the data set that was used, the solutions show that considering network effects may be just as important as considering demand uncertainty in the fleet assignment model. In many circumstances, capturing flight leg dependencies may be more important than capturing the effects associated with demand stochasticity.
Chapter 7

Conclusions

In this chapter, we summarize the main contributions of this thesis, present some caveats to consider when reading our results, and describe various future research directions that are logical extensions of this work.

7.1 Contributions

In this section, we categorize our contributions based on the model with which they are associated. One significant contribution, however, is prevalent throughout the thesis. That is, the computational results are performed mostly on actual airline flight networks and not simple toy problems.

7.1.1 The Passenger Mix Model

One of the central problems in various airline decision processes is determining how passengers are accommodated (or flow) on the flight leg network. We refer to this problem as the attainable contribution problem. We begin Chapter 3 by developing a taxonomy which classifies previous research on the attainable contribution problem based on the uncertainty of the demand data and the proposed solution approach. Next, we develop a linear programming formulation of the attainable contribution problem, called the passenger mix model. Our formulation is significant for two reasons, namely:

1. This is one of the first formulations of the attainable contribution that models the partial
recapture of passengers if the desired itinerary is not available. Almost all other models either assume no recapture or perfect recapture.

2. Using a change of variable strategy and employing column and constraint generation, a realistically-sized problem with no recapture is solved in seconds; one with partial recapture is solved in just a few minutes on a workstation class computer.

With these fast solution times, the application areas of our passenger mix model can be operational, tactical, or strategic in scope. We present modifications that allow our model and algorithm to be used operationally for passenger re-routing during recovery. The passenger mix model can also be used for revenue management in a number of capacities, including:

- It provides an upper bound to the expected contribution for any seat inventory control.
- Its optimal dual solution can be used as displacement costs in a bid price network revenue management system.
- It determines passengers who should be re-routed to less capacitated alternative itineraries.

Lastly, the passenger mix model is a valuable tool for assigning fleets to a flight schedule. These benefits are addressed in the following sections.

7.1.2 The Basic Fleet Assignment Model

In Chapter 5, we show the importance of measuring the quality of a fleeting, whether it is a linear program such as the passenger mix model or a comprehensive simulation. Comparing different fleetings based on an accurate post-processing contribution solver is crucial to determining good assignments. Logically, the more accurate the attainable contribution solver is to reality, the better the assignments.

We present two different models to represent spill. The first is the representative fare spill model (RFSM), which is currently employed at most major airlines. It assumes a constant spill fare for all passengers who are spilled. The second is the integrated spill model (ISM), which calculates spill costs by determining which passengers are spilled and summing their fares. Consistently, the ISM outperforms the RFSM, sometimes on the order of $30M/year.
The ISM is actually the special case of the single flight leg, multi-fare class, seat inventory control problem when demands are deterministic. Belobaba and Farkas [10] conjecture that using this seat inventory control model could lead to improved fleet assignment decisions. The computational results from Chapter 5 affirm this conjecture for all problem sizes including the full size daily domestic fleet assignment problem.

Using the ISM blindly does not guarantee improved fleet assignments. The method by which fare is allocated to the different flight legs in a multi-leg itinerary plays a crucial role. A common practice at airlines is to pro-rate the fare based on the flying miles of each flight leg in an itinerary. In Chapter 5, we show that this fare allocation scheme performs poorly, especially when using the ISM to model spill costs. Using allocation schemes that pro-rate the fare among all of the capacitated flight legs in an itinerary consistently perform the best.

7.1.3 The Multiple Flight Column Approach

We also present a new formulation of the fleet assignment that could potentially capture the network effects of multi-leg passengers more accurately by grouping subsets of flight legs into variables. Instead of the decision variable being a single fleet type to a single flight leg, the decision variable is an assignment of many fleet types to many flight arcs that are connected based on passenger itineraries.

7.1.4 The Combined Model

The optimal objective value of the combined fleet assignment and passenger mix model is an upper bound on the expected contribution that can be obtained. Using both the LP relaxation solution and initial integer solutions of the combined model, we can get bounds that show the solutions from the methods developed in Chapter 5 are either optimal or near-optimal. Empirically, the LP relaxation solution of the combined model is much more fractional than the traditional fleet assignment model. We analyze the reasons why and present techniques to decrease this fractionality. We discuss a number of approaches for solving the combined model, including a direct solution approach, Lagrangean relaxation, and the heuristic methods in Chapter 5.

In some circumstances the optimal solution of the traditional fleet assignment model is
also the optimal solution to the combined model, such as when all passengers travel on at most one capacitated leg. Also, we prove that the optimal assignment of the multiple flight column formulation is also the optimal solution to the combined model, however, the number of variables required may be quite large.

Traditionally, planners have assumed that capturing the stochasticity of the demand is primary, while considering network effects is secondary. For the data set used, the evidence suggests that capturing network effects might be just as important as demand uncertainty in the fleet assignment model, if not more important.

7.2 Some Caveats

7.2.1 Itinerary Demand

Forecasting demand for itineraries raises a number of issues, within the context of itinerary-based fleet assignment. Our research suggests that the benefits of using better spill models, which require itinerary demand data, are on the order of $30M/year. Yet, if the increased error associated with forecasting itinerary demand forecasts is greater than this value, the benefits may not be fully realized. Also, we do not know how other factors, such as multiple fare classes and the additional side constraints, might affect the potential benefits.

A more fundamental issue is the model of breaking out market demand by itineraries, and then considering this recapture. Typically, a passenger does not know the itinerary he wants, but prefers to go from his/her origin to his/her destination around some specified time. A better method might include a model that more accurately captures the passenger decision process.

7.2.2 Recapture

Currently, the airlines use very simple models to account for recapture in their fleet assignment models (and revenue management systems). In Chapter 3, using a rough estimate for recovery rates, we estimated that the magnitude of recapture is on the order of $100M/year. Minor perturbations in these recovery rates affected the contribution by $10M/year. This evidence suggests that considering recapture could have a significant impact on fleeting decisions and
this relationship should be investigated further.

7.2.3 The Attainable Contribution Solver

Using the passenger mix model as our attainable contribution solver might lead to bias in some of our results. The ISM approximates spill costs by essentially solving a passenger mix model on one flight leg, therefore, the selection of the passenger mix model as the attainable contribution solver in our framework might favor the use of the ISM. This might also be the case in the simulation results at the end of Chapter 6. If a simulation model or a stochastic passenger flow model is used as the basis in comparing different fleetings, the ISM model might not perform quite as well.

7.3 Areas of Future Research

This research has contributed to the understanding of the relationship of passenger itineraries and the fleet assignment model, however, this has still raised numerous questions and areas of further investigation. We present a number of directions of future research in the areas of further validation and extensions of the models. In this section, when we refer to the different methods of fleet assignment, we are referring collectively to all combinations of fare allocation schemes and the two spill models presented plus the combined fleet assignment model.

7.3.1 Further Validation

The major implementation hurdle of the ideas presented is the switch from aggregated demand on a flight leg to forecasted demand for itineraries. Since accurately forecasting demand for an itinerary is more difficult than for a flight leg, the sensitivity of the model to the itinerary demand must be determined. An effort should be made to measure which of the fleet assignment methods does a better job in the face of uncertainty and errors in the itinerary data. There are a number of issues that should be addressed to determine the sensitivity of the different fleet assignment models.

1. At the end of Chapter 6, we showed how fleetings obtained from different methods performed with stochastic demand for a 3 fleet problem. The results of this experiment
should be validated for a number of different scenarios, especially the full size problem.

2. Ideally, we wish to use a fleet assignment model that is *stable*. If small perturbations in the itinerary demand forecasts radically alter the fleet assignment, then it will be difficult to have a lot of confidence in the model and the resulting fleetings.

3. Also, the performance of the fleet assignment model when the itinerary demand forecasts are wrong should be considered.

4. While it can be modeled, we have not explicitly considered multiple fare classes due to the resulting size of the problem and the increased difficulty in accurately determining *IFC* forecasts. The stability and performance of the various methods should be investigated with multiple fare classes.

It is possible that including a number of additional constraints might alter the performance of the various fleet assignment models. Also, there are a number of models in which the fleet assignment model is integrated with another aspect of the airline scheduling process, such as maintenance ([4], [26]), crew pairing ([32], [14]), or schedule generation ([1], [34]). A thorough investigation of how the new models work with these additional constraints should be validated. Finally, the effects of using a different attainable contribution solver should be considered.

### 7.3.2 Extensions

In Chapter 3, we present a simple model of recapture to get an approximate representation. We show that recapture can be as large as 1% of contribution for a given fleeting. This increase in contribution is larger than the benefits that we estimate using the methods presented in this thesis. We also show that recapture can have a significant impact on dual costs which are used as displacement costs in bid price revenue management systems. Another area of future research would be to examine the concept of recapture, develop a more sophisticated model, and see how this will affect such things as schedule generation, fleet assignment, and revenue management.

In Chapter 5, we develop a fleet assignment model with decision variables that represent an assignment of many fleet types to many flight legs in order to more accurately consider spill. One
of the keys to this model is a partitioning of the flight legs into subsets. To accomplish this, we also developed a set partitioning heuristic based on all-pairs minimum cut problems. Research should be conducted to measure the effectiveness of this multiple flight column fleet assignment model and determine empirically the solution quality of the set partitioning heuristic.

For itinerary-based airline fleet assignment, we show that for small and large problems, optimal and near-optimal solutions can be obtained, respectively. In itinerary-based airline fleet assignment, we assume that the airline and itinerary choices have been made by the passenger and that the resulting fleet assignment does not affect the itinerary choice. A logical next step is to incorporate an itinerary choice model into the fleet assignment process, which would result in a market-based airline fleet assignment model. Since this model would be hard to solve, a first step would be to see the difference in fleetings that occur between the market-based and the itinerary-based fleet assignment models for small problems. If there isn't much difference, then itinerary-based airline fleet assignment might be adequate.

7.4 Final Thoughts

We present evidence that suggests the importance of considering network effects associated with leg-dependent spill and recapture in the scheduling process, specifically the fleet assignment problem. The key point of this thesis is that the current models used to measure spilled revenue in the traditional fleet assignment are somewhat arbitrary and, many times, bad approximations. One of the improvements that airlines can make to their fleet assignment decision process immediately is to include some type of procedure that more accurately measures the performance of a given fleetting. Doing so, a true comparison of different fleetings can be made. We present a deterministic linear programming model for the attainable contribution problem that can be solved in seconds when recapture is not considered and in a few minutes when recapture is considered. Using this model in an interactive scheduling environment would allow human planners to quickly determine fleetting quality after they have altered the fleet assignment.

We also describe methods that show how to achieve better initial fleetings. These methods adjust the objective value function of the traditional fleet assignment models, and achieve
better fleeting without increases in solution times. We formulate the combined fleet assignment, passenger mix model to show that these methods of improving the traditional fleet assignments lead to near-optimal solutions. While most of this is conducted in a deterministic framework, we present evidence that suggests that capturing network effects is just as important if not more important than capturing the stochastic nature of passenger demand.
Bibliography


Appendix A

Simulation Methods

It is not always possible to analytically get an estimate, especially when many random variables are interacting, as is the case with airline fleet assignment. In this situation, one has no choice but to use simulation. Many times there may be some assumptions or approximations that must be used for tractability. In simulation we can relax some of these assumptions and potentially get more accurate estimates.

A.1 Assumptions

In this thesis, we need to simulate many different demand instances for the entire network. The following assumptions and models were chosen:

1. For each itinerary, an estimate of the mean is known, $\overline{D}_p$, and assumed accurate.

2. For each itinerary, the standard deviation is a linear function of the square root of the mean, i.e., $\sigma_X = z \cdot \sqrt{\overline{D}_p}$, where the constant $z$ is leg-independent within each scenario.

3. The chosen distribution for itinerary demand is the gamma distribution.

4. The demand for each itinerary is independent of the demands of other itineraries.
A.2 The Gamma Distribution

The gamma distribution is a two parameter model with the following probability density function:

\[
f_X(x) = \begin{cases} 
\frac{x^{\alpha - 1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} & x \geq 0 \\
0 & \text{otherwise.}
\end{cases}
\]  
(A.1)

The mean and variance of the gamma distribution in regards to the shape parameter, \( \alpha \), and scale parameter, \( \beta \), are as follows:

\[
E[X] = \alpha \beta \quad \text{(A.2)}
\]
\[
\sigma_X^2 = \alpha \beta^2. \quad \text{(A.3)}
\]

Since we desire \( E[X] = \overline{D}_p \) and \( \sigma_X^2 = z^2 \overline{D}_p \), we wish to sample from the gamma distribution with the following parameters:

\[
\alpha = \frac{\overline{D}_p}{z^2} \quad \text{(A.4)}
\]
\[
\beta = z^2. \quad \text{(A.5)}
\]

We use the sampling techniques for the gamma distribution found in Dagpunar [17]. If \( \alpha > 1 \) then we use a \( t \)-distribution method; if \( \alpha \leq 1 \), we use a switching method. To determine a random number uniformly distributed from 0 to 1, we use a linear congruential pseudorandom number generator with constant 16807 and modulus \( 2^{31} - 1 \).

The above methods give us a positive number based on the distribution. We wish, however, to have strictly integer values for our simulations. Given the output of the random variate generation, \( x \), we compare a uniform random number from 0 to 1 to the fractional part of \( x \). If it is less than the fractional part, we round \( x \) up, otherwise we round down. The resulting random number will have the same mean, \( \overline{D}_p \), however, the variance will be higher than the desired \( z^2 \overline{D}_p \). Let \( \tilde{z} \) be the resulting Z-factor after using this rounding technique. The error \( (\tilde{z} - z) \) is higher at small values of \( \overline{D}_p \), and very small at high values of \( \overline{D}_p \).
The total algorithm is as follows:

<table>
<thead>
<tr>
<th>For each ( p \in P ), which has mean ( \bar{D}_p ) and variance ( \chi^2 \bar{D}_p ), do the following:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1:</strong> Generate ( x \sim G \left( \frac{D_p}{\chi^2}, \chi^2 \right) ) using the methods described above.</td>
</tr>
<tr>
<td><strong>Step 2:</strong> Generate ( y \sim U(0, 1) ).</td>
</tr>
<tr>
<td>If ( y &lt; (x - \lfloor x \rfloor) ), then ( d_p = \lfloor x \rfloor )</td>
</tr>
<tr>
<td>else ( d_p = \lfloor x \rfloor ).</td>
</tr>
</tbody>
</table>
Appendix B

Glossary

B.1 Definitions

aircraft family: a grouping of aircraft types that is based upon similar operating characteristics from the flight crew’s perspective.

aircraft flow balance: the set of constraints in the fleet assignment problem that ensure that during the course of a day equal number of aircraft of the same type depart from and arrive at an airport.

aircraft maintenance routing: the set of decisions to route aircraft such that every aircraft in the fleet overnights at a maintenance station within the standards set by the Federal Aviation Administration.

airline choice model: a model for the behavior of passengers determining the airline that they wish to fly.

airline recovery problem: the problem of returning to the schedule plan after a major disruption caused by irregular operations has occurred.

arrival event: a node on the fleet assignment timeline network associated with the arrival of an aircraft of fleet type \( k \), at airport \( o \), at time \( t \).

assignment combination: a specific assignment of fleet types to a subset of flight legs used in the multiple flight column fleet assignment model.
**attractiveness factor**: a quantitative number assigned to itineraries that measures the desirability in relation to other itineraries in the same market. The Quantitative Share Index is an example of an attractiveness factor.

**availability event**: a node on the fleet assignment timeline network associated with the availability of an aircraft of fleet type $k$, at airport $o$, at time $t$. This will be typically be the time of arrival plus the minimum ground time of fleet type $k$.

**base recovery rate**: the recovery rate from one itinerary to another based solely on $QSI$ values.

**bid pricing**: a network revenue management seat inventory control in which a passenger is accepted if his fare is greater than the sum of the displacement costs on all of the flight legs in the itinerary.

**block time**: the time difference between the removal of blocks for departure and the placement of blocks upon arrival.

**blocked itinerary**: an itinerary in which one of the flight legs is full in sequential passenger flow models.

**booking limits**: the maximum number of passengers that are accepted in specific fare classes.

**choke leg**: the flight leg that is full causing blocked itineraries in sequential passenger flow models.

**capacitated equal fare allocation scheme (CEF)**: a method of allocating fare to flight legs in an itinerary by equally dividing the fare among all capacitated flight legs.

**capacity-based fare allocation scheme (CF)**: a method of allocating to flight legs in an itinerary based on the comparison of the demand and the possible assignments of aircraft to that flight leg.

**cold-start fleet assignment**: developing a new fleeting from scratch based on the various models presented in this book.
column generation: a Dantzig-Wolfe decomposition approach in which variables are only included in the problem as needed.

competition-based fleet assignment: a fleet assignment model in which demand data is not broken out by airlines and itineraries.

connection networks: a time-space network where connection arcs are used. While larger than a timeline network, a connection network maintains the identity of specific aircraft, thus allowing a greater modeling flexibility.

connectivity: a quantitative measure between two flight legs based on the number of passengers that have both flight legs in their itinerary.

contribution: either the revenue generated minus the carrying costs for the passenger mix model or the revenue generated minus the carrying and operating costs for the fleet assignment model (see fleeting contribution).

cover constraints: the constraints in the fleet assignment problem which state that exactly one fleet type can be assigned to a flight leg.

daily domestic fleet assignment problem: assigning fleet types to flight legs for the main flight schedule of major U.S. airlines. The schedule repeats every day of the week and only includes flight legs that both take off and land in North America.

demand factor (DF): a measure of the congestion in the flight network which is the ratio of the number of seats requested throughout the system versus the supply of seats.

departure event: a node on the fleet assignment timeline network associated with the departure of an aircraft of fleet type \( k \), at airport \( o \), at time \( t \).

displacement costs: the marginal cost of a seat on a flight leg in a bid pricing network revenue management system.

equal fare allocation scheme (EF): a method of allocating fare to flight legs in an itinerary by equally dividing the fare among all capacitated flight legs.
fare class: a distinction between different passengers on an itinerary or a flight leg based on the restrictions that the passenger has purchased with his ticket.

fences: restrictions in fare products set up in revenue management systems that maintain segmentation of the market.

fleet assignment: the process by which an airline determines the assignment of fleet types to flight legs.

fleet planning: the strategic decision process in which airlines determine which aircraft types to buy for and which to remove from their fleet.

fleet types: a grouping of aircraft types that is based on similar operating costs and seating capacity. This is not to be confused with the aircraft family grouping.

fleeting: the output a fleet assignment model. The assignment of fleet types to flight legs.

fleeting contribution: the revenue generated minus the carrying and operating costs for a specific fleeting.

flight: a set of flight legs with the same flight number.

flight leg: the basic revenue generating aircraft operation which consists of exactly one take-off and exactly one landing.

full fare allocation scheme (FF): a method of allocating fare to flight legs in an itinerary by allocating the full fare to all flight legs in an itinerary.

greatest probability fare allocation scheme (GF): a method of allocating fare to flight legs in an itinerary by allocating the full fare to the one flight leg that is most likely to spill.

ground arcs: an arc in the fleet assignment timeline network which measures the number of aircraft of a fleet type, \( k \), on the ground at airport \( o \) between two successive events.

horizontal aggregation: the process in which airlines aggregate demand data across the different days of the week.
**hub and spoke**: an airline network structure in which a majority of the flights originate or terminate at a small number of airports. This type of structure allows the service in numerous markets with relatively few flight legs.

**integrated spill model (ISM)**: a spill model in which lower fare passengers are spilled before high fare passengers.

**irregular operations**: the operation of an airline due to unforeseen events such as bad weather and mechanical failure.

**itinerary**: a series of connected flight legs by which passengers travel in a market.

**itinerary-based fleet assignment**: a fleet assignment model in which demand data estimates are given for each possible itinerary.

**itinerary choice model**: a model of passenger behavior in choosing a specific itinerary for their journey.

**K-factor**: the constant used in a model of demand variability where the standard deviation is linearly proportional to the mean of the demand.

**keypaths**: the most likely path on which a commodity will flow.

**level sets**: in a minimization problem, the collection of points in the feasible set for which the objective function value is less than or equal to some number.

**marginal cost of capacity**: the change in operating costs along the optimal aircraft frontier for increasing the capacity by one seat.

**market**: an ordered origin airport and destination airport pair grouping of passengers. BOS-LAX and LAX-BOS are distinct and different markets.

**market-based fleet assignment**: a fleet assignment model in which passenger demand is broken out by markets and airline choices.

**mileage-based pro-rated fare allocation scheme (MF)**: a method of allocating fare to flight legs in an itinerary by pro-rating the fare based on the flying miles of each leg in the itinerary.
**minimal fleeting solution**: given a specific assignment of fleet type to aircraft, this solution contains the minimum number of aircraft on the ground at all airports.

**minimum turn time**: the required time to prepare an aircraft for the next flight in its rotation.

**multiple flight column approach (MFC)**: a fleet assignment model in which decision variables represent whether or not to use a set of arcs in the timeline network. The decision variable of a traditional fleet assignment model represents the use of a single arc in the timeline network.

**network effects**: the effects associated with the multi-leg passenger itineraries. Because of these effects, one cannot guarantee an optimal fleeting by assuming that spill costs are leg independent.

**no-shows**: passengers who make reservations on a flight leg, then do not show up at the time of departure.

**node consolidation**: a method used in fleet assignment models of decreasing the size of the timeline network by consolidating redundant nodes.

**null itinerary**: a dummy itinerary associated with spilling and not offering an alternative itinerary.

**optimal aircraft frontier (OAF)**: a function which represents the aircraft of cheapest cost for a specific number of seats. Aircraft may be "created" using a convex combination of fleet types.

**overbook**: the procedure in which airlines sell more seats then are available to counter the effects of no-shows.

**overcapacitated**: a flight leg in which the unconstrained is larger than the number of seats on the largest aircraft that can be assigned.

**overnight arcs**: a special ground arc in the fleet assignment timeline network that represents aircraft of fleet type \( k \) staying on the ground overnight at airport \( o \).
parallel markets: markets in which a passenger could potentially travel to get from his ultimate origin to ultimate destination. For example, LGA (New York Laguardia)-DCA (Washington National) and JFK (New York JFK)-IAD (Washington Dulles) are parallel markets since a passenger may look at itineraries in both markets to get from Manhattan to downtown Washington DC.

passenger connectivity ratio (PCR): a ratio of the number of passengers

passenger flow model: a model that determines the resulting number of passengers on each flight leg for after spill and recapture is considered.

passenger mix model: a specific passenger flow model which is solved by a deterministic linear programming model.

passenger sell-up: a phenomenon of passenger recapture in which a passenger makes a reservation in a fare class that is more expensive than desired.

point to point: an airline network structure in which there are either flights to and from most airport pairs, or a series of flight legs. The opposite of a hub and spoke network.

potentially capacitated: a flight leg that is neither overcapacitated or undercapacitated, i.e., the demand is between the number of seats of the smallest and largest aircraft that can be assigned.

pricing subproblem: the subproblem used to determine columns to add to the restricted master problem in a column generation approach.

probabilistic pro-rated fare allocation scheme (PF): a method of allocating fare to flight legs in an itinerary by pro-rating the fare based on the probability of being spilled because of the flight legs in the itinerary.

Quantitative Share Index (QSI): an industry standard index for determining the attractiveness of itineraries.

quasiconcave: a maximization function in which the level sets form convex sets.
recapture: the process in which passengers are accommodated on another itinerary when their desired itinerary is full.

representative fare spill model (RFSM): a model in which all spilled passengers are assumed to have the same representative fare. This fare is the same regardless of the number of passengers spilled. The representative is usually some function of the average fare on the flight leg.

restricted master problem: the problem to be solved at each stage of a column generation approach. It is restricted since only a subset of all possible variables are included at any one time.

revenue management: a five step process in which a supplier segments the market for shared resources, sets up restrictions to maintain segmentation, prices the different segments, forecasts demand for each segment, and determines a control procedure for accepting requests for the resources.

rotations: a sequence of flight legs that are flown by the same aircraft that begin and end at a maintenance station.

schedule generation: the decision process in which the airline determines where and when to fly.

seat inventory control: a method that airlines use to control the supply of seats for different segments of the market.

sit arcs: see ground arcs.

special ordered set (SOS): a set of variables that have special dependencies. For example, a Type 3 SOS requires that exactly one variable be set to one and all other variables in the set are set to zero.

spill: the process by which passengers are not accommodated on their desired itinerary due to capacity limitations.

spill calculation module: a method that calculates the spill that occurs for a specific fleeting. This can be done by either passenger flow models or simulation.
spill costs: the amount of revenue that is not accommodated by an airline due to the fleeting decision.

spill inaccuracy ($SI$): the difference between the estimated spill of the fleet assignment model and the spill that is measured by a more accurate spill calculator.

system load factor ($SLF$): the ratio of the number of seats filled to the number of available seats.

system optimization: a model in which optimization is performed from the viewpoint of the system.

tailing: an effect in iterative algorithms in which many iterations at the end of the process are performed without improving the solution.

through flights: flights that are considered direct (not non-stop) with the same flight number.

timeline network: the network structure that is the basis of fleet assignment models.

uncapacitated: a flight leg in which the unconstrained demand is less than the number of seats on the smallest aircraft that can be assigned.

unconstrained demand: the number of passengers that would travel on either an itinerary or flight leg if there was unlimited capacity.

unconstrained revenue: the revenue attained if there was infinite capacity in the system.

user optimization: a model in which optimization is performed from the viewpoint of the user (see system optimization).

vertical aggregation: the process in which airlines aggregate demand data across different fare classes.

warm-start fleet assignment: a fleet assignment procedure in which minor adjustments are made to an input fleet assignment.

wrap-around arcs: see overnight arcs.
yield management: see revenue management.

Z-factor: the constant used in a model of demand variability where the standard deviation is linearly proportional to the mean of the demand.

zero-based fleet assignment: see cold-start fleet assignment.

B.2 Notations

B.2.1 Roman Symbols

$A$: the set of airports indexed by $o$.

$q^m_i$: 1 if flight leg $i$ is a member of subset $L^m$

0 otherwise.

$b_{p,f}^{r,g}$: the recovery rate of a passenger desiring itinerary $p$ under fare class $f$ to itinerary $r$ under fare class $g$. This is roughly equivalent to the probability of a passenger flying on itinerary $r$ under fare class $g$ given that there is no availability on itinerary $p$ under fare class $f$.

$c_{i}^{sp}$: the spill costs on flight leg $i$.

$c_{k,i}^{sp}$: the spill costs on flight leg $i$ when fleet type $k$ is assigned.

$\bar{c}_{k,i}$: the operating cost of fleet type $k$ when assigned to flight leg $i$.

$\bar{c}_{p,f}^{r,g}$: the reduced cost associated with the variable $t_{p,f}^{r,g}$.

$C(k)$: the set of flights that cross the countline when flown by fleet type $k$.

$C_{i_1,i_2}$: the connectivity between flight legs $i_1$ and $i_2$.

$CAP_{i,j}$: the assigned capacity to flight leg $i$ for cabin $j$.

$CAP_{i}$: the minimum feasible capacity of flight leg $i$.

$\overline{CAP}_{i}$: the maximum feasible capacity of flight leg $i$.

$D_{p,f}$: the unconstrained demand for itinerary $p$ under fare class $f$. 

210
\(d(p, r)\): the arrival time difference between itinerary \(p\) and \(r\).

\(F\): the set of all fare classes, indexed by \(f, g,\) or \(h\). The set of all fare classes assigned to a specific cabin \(j\) is denoted by \(F_j\).

\(fare_i\): the fare of a passenger on flight leg \(i\). The fare of a passenger for itinerary \(p\) under fare class \(f\) is denoted \(fare_{p, f}\).

\(FC_{sp}\): the parameter for fare spread to model multiple fare classes with vertically aggregated demand.

\(fc_p\): the number of fare classes for itinerary \(p\).

\(I(k, o, t)\): the set of inbound flights to node \(\{k, o, t_j\}\).

\(IC\): an arbitrary seat inventory control scheme.

\(K\): the set of different fleet types indexed by \(k\).

\(K(i)\): the set of fleet types that may be assigned to flight leg \(i\). This will be a subset of all fleet types \(K\). The two sets might not be equal due to noise restrictions, flying range restrictions, etc.

\(J\): the set of cabins (e.g., first-class, coach), indexed by \(j\).

\(J(f)\): the mapping from fare class \(f\) to the cabin with which it is associated.

\(L\): the set of flight legs in the flight schedule indexed by \(i\).

\(\bar{L}\): the set of all capacitated flight legs.

\(L^\prime\): an arbitrary subset of flight legs.

\(l_i\): the number of passengers that are spilled on flight leg \(i\).

\(legs_p\): the number of valid legs on itinerary \(p\).

\(M\): the set of all markets, indexed by \(m\)

\(N\): the set of nodes in the timeline network indexed by \(\{k, o, t_j\}\)
$N_{HA}$: the number of different demand instances for horizontal disaggregation

$N^v$: the pre-set maximum number of variables in the MFC approach.

$n_k$: the number of aircraft in fleet type $k$.

$n_m$: the number of itineraries in market $m$ or the number of variables in flight leg subset $m$.

$O(k,o,t)$: the set of outbound flights from node $\{k,o,t\}$.

$P$: the set of all itineraries (including the null itinerary), indexed by $p$, $r$, or $q$.

$pr_i$: an estimate of the probability of an arbitrary passenger being spilled on flight leg $i$.

$prob_p$: the sum of all $pr_i$ for all flight legs in itinerary $p$.

$Q_i$: the unconstrained demand on flight leg $i$.

$Q_m$: the sum of all QSI in market $m$. This is equal to 1.0 if the airline has a monopoly in a market.

$q_p$: the QSI value for itinerary $p$.

$S$: an arbitrary fare allocation scheme.

$S(p,f)$: the number of demand segments for $(p,f)$ indexed by $s$.

$SEATS_k$: the number of passenger seats on fleet type $k$.

$SP_{HA}$: a parameter for the variance of demand when horizontally disaggregating demand data.

$T$: the sorted set of all event (departure or availability) times at all airports, indexed by $t_j$.

The event at time $t_j$ occurs before the event at time $t_{j+1}$. Also, $|T| = m$; therefore $t_1$ is the time associated with the first event after the count time and $t_m$ is the time associated with the last event before the count time.

$t_{p,f}^{r,g}$: the number of passengers who desire travel on itinerary $p$ under fare class $f$, but whom the airline attempts to redirect onto itinerary $r$ under fare class $g$.

$U$: an arbitrary assignment combination of fleet types and flight legs.
$U(L^n)$: the set of all assignment combinations of fleet type-flight leg assignments for the subset of flight legs in $L^n$.

$u_m: \begin{cases} 1 & \text{if subset } L^m \text{ is used in the partition.} \\ 0 & \text{otherwise.} \end{cases}$

$\psi^U_k$: the number of times that the flight leg arcs in the assignment combination $U$ cross the timeline for fleet type $k$. This value can be any integer greater than or equal to zero.

$X_{k,i}: \begin{cases} 1 & \text{if flight leg } i \text{ is assigned to fleet type } k. \\ 0 & \text{otherwise.} \end{cases}$

$\pi^r_{p,f,g}$: the number of passengers (possibly fractional) that desire to travel on itinerary $p$ under fare class $f$, but instead travel on itinerary $r$ under fare class $g$.

$w_m$: the weight of a subset of flights defined as the sum of the connectivity of all flight leg pairs where both flight legs are in the subset.

$w^U_{\{k,o,t\}}$: the number of flight arcs in assignment combination $U$ that are incident to node $\{k,o,t\}$ minus the number of flight arcs in the assignment combination $U$ that are incident from node $\{k,o,t\}$. This can be any positive or negative integer or zero.

$Y^+_{k,o,t_j}:$ the number of fleet type $k$ aircraft that are on the ground at airport $o$ immediately after time $t_j$.

$Y^-_{k,o,t_j}:$ the number of fleet type $k$ aircraft that are on the ground at airport $o$ immediately before time $t_j$. If $t_1$ and $t_2$ are the times associated with adjacent events, then $Y^+_{k,o,t_j} = Y^-_{k,o,t_{j+1}}$.

$Z(X)$: the contribution function associated with the fleeting decision $X$.

**B.2.2 Greek Symbols**

$\delta^p_i: \begin{cases} 1 & \text{if flight leg } i \text{ is on itinerary } p \\ 0 & \text{otherwise} \end{cases}$

$\alpha^j$: the subgradient discount at iteration $j$ when using Lagrangean relaxation.

$\beta$: the percentage change in recovery rate to determine sensitivity.
$\beta_{p,f}^s$: a factor to discount the fare for segment $s$ of $(p,f)$ to model the expected marginal seat revenue.

$\gamma_j$: multiplier for demand instance $j$ when doing horizontal disaggregation.

$\varepsilon$: a parameter used for tolerance levels.

$\theta_L$ and $\theta_C$: multipliers used to adjust the demand on local and connecting passengers, respectively.

$\kappa$: the spill fare to average fare ratio in the representative fare spill model.

$\lambda_i$: the dual cost associated with flight leg $i$.

$\mu$: a convex combination parameter.

$\pi_{p,f}$: the dual cost associated with the capacity constraint of itinerary $p$, fare class $f$ in the passenger mix problem.

$\sigma_i$: the standard deviation on flight leg $i$. Also, $\sigma_p$ denotes the standard deviation on itinerary $p$.

$\tau$: a convex combination parameter.

$\psi$: the penalty for not accommodating a passenger in irregular operations passenger re-routing.

$\omega$: a parameter used to limit the percentage change in recovery rate analysis.