Dynamic Airline Scheduling and Robust Airline Schedule De-Peaking

by

Hai Jiang

B.S., Tsinghua University (2001)
M.S., Massachusetts Institute of Technology (2003)

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2006

Author .....
Department of Civil and Environmental Engineering
August 11, 2006

Certified by .....
Cynthia Barnhart
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by .....
Andrew J. Whittle
Chairman, Department Committee on Graduate Students

Dynamic Airline Scheduling and Robust Airline Schedule De-Peaking

by

Hai Jiang

Submitted to the Department of Civil and Environmental Engineering
on August 11, 2006, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Transportation

Abstract

Demand stochasticity is a major challenge for the airlines in their quest to produce profit maximizing schedules. Even with an *optimized* schedule, many flights have empty seats at departure, while others suffer a lack of seats to accommodate passengers who desire to travel. Recognizing that demand forecast quality for a particular departure date improves as the date comes close, we tackle this challenge by developing a *dynamic scheduling* approach that re-optimizes elements of the flight schedule during the passenger booking period. The goal is to match capacity to demand, given the many operational constraints that restrict possible assignments. We introduce flight re-timing as a dynamic scheduling mechanism and develop a re-optimization model that combines both flight re-timing and flight re-fleeting. Our re-optimization approach, re-designing the flight schedule at regular intervals, utilizes information from both revealed booking data and improved forecasts available at later re-optimizations. Experiments are conducted using data from a major U.S. airline. We demonstrate that significant potential profitability improvements are achievable using this approach.

We complement this dynamic re-optimization approach with models and algorithms to de-peak existing hub-and-spoke flight schedules so as to maximize future dynamic scheduling capabilities. In our robust de-peaking approach, we begin by solving a *basic* de-peaking model to provide a basis for comparison of the robust de-peaked schedule we later generate. We then present our *robust* de-peaking model to produce a schedule that maximizes the weighted sum of *potentially connecting itineraries* and attains at least the same profitability as the schedule produced by the basic de-peaking model. We provide several reformulations of the robust de-peaking model and analyze their properties. To address the tractability issue, we construct a restricted model through an approximate treatment of the profitability requirement. The restricted model is solved by a decomposition based solution approach involving a variable reduction technique and a new form of column generation. We demonstrate, through experiments using data from a major U.S. airline, that the schedule generated by our robust de-peaking approach achieves improved profitability.

Acknowledgments

I am fortunate and blessed to have Professor Cynthia Barnhart as my advisor during my doctoral study. I am deeply grateful to Cindy for her constant encouragement and tremendous support through the years. She has given me the widest freedom to pursue my research directions, and provided me the most valuable advice on how to enhance this work. She also showed great care and understanding to me and always stated that the student's interests are primary while the advisor's are secondary. Cindy made herself available to me during weekends and evenings. We sometimes even talked over the phone about research when she was driving between Boston and Vermont. She is indeed a great advisor.

I would like to thank my doctoral committee members, Professor Nigel Wilson and Professor John-Paul Clarke for their insightful comments and suggestions.

I would like to thank my colleagues in the Large-Scale Optimization Group for their friendship. I would like to extend my thanks to Ginny Siggia and Maria Marangiello for their wonderful assistance in all matters.

Many thanks go to the United Parcel Service and the Sloan Foundation for providing financial support for my doctoral study.

Finally, I want to thank my family for their love and confidence in my ability to succeed.

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

List of Tables

Chapter 1

Introduction

1.1 Motivation

It has been a major challenge for airlines to design a flight schedule (timetable and fleeting) to match fluctuating passenger demand. The flight schedule, that is, the supply side of the passenger air transportation system, has to be determined well in advance due to contractual and operational requirements in the industry. Examples of the schedule planning process can be found in Goodstein (1997), Jarrah (2000), Barnhart et al. (2002), and Frank et al. (2005). The steps and timelines of the schedule planning process may differ slightly from one airline to another, yet it typically starts 12 months prior to departure and lasts approximately 9 months. To generate a schedule that has the most revenue potential, the airline's scheduling department compiles input data to the planning process that reflects a macro-forecast of the economy, the airline's strategic objectives, forecasted passenger demands, average fares, available resources (aircraft, personnel, gates), and so on. Then, the flight schedule is constructed and published through different distribution channels. Despite the fact that scheduling decisions are made at a time when demand is highly uncertain, the schedule is anticipated to be carried out on the date of departure.

In the *booking period*, that is, the time between the schedule being published and the departure date, airlines employ sophisticated revenue management techniques to sell as many seats as possible in the flight schedule, while maximizing revenue. In peak

demand situations, prices are raised, thereby reserving the scarce seats for high fare passengers. When demands are low, prices are reduced to stimulate bookings. Hence, revenue management systems help to shape demand to fit the fixed supply of seats in the schedule. Nevertheless, no matter how sophisticated these systems are, the stochastic nature of passenger demand still results in many flight legs having empty seats upon departure, while others suffer a lack of seats to accommodate passengers who desire to travel. Figure 1-1 shows the histogram of *load factors (LFs)* (the ratio of the number of paid passenger seats on a flight to the seating capacity of the flight) for all flights operated by a major U.S. airline in a month. The load factors range widely, with a mean of 0.76. The quantiles of the histogram show that more than 25% of the flights are highly demanded and are sold close to capacity (LFs are greater than 0.93), while at the same time another 25% of the flights have significant numbers of empty seats (LFs are less than 0.66, that is, about 1 out of every 3 seats remain unsold at the time of departure). The imbalanced load factors across flight legs indicate that the actual demand at the departure date is far from being accurately captured by the demand forecasts used in developing the schedule. If those empty seats appeared on flight legs with excess demand, significant revenue gains would be possible.

Because forecast quality improves dramatically as the time of departure approaches (see Berge and Hopperstad 1993, Feldman 2000, Bish et al. 2004, and Sherali et al. 2005), an interesting question is how an airline can utilize improved demand forecasts to *re-optimize* the original schedule and move excess capacity to flight legs with a shortage of capacity. We focus on developing mechanisms to be used in matching supply and (fluctuating) demand, by making only small changes to the schedule so as to minimize the complication in operations. In doing so, capacity will be shifted from flight legs with *expected* excess capacity to those with *expected* capacity shortfalls, and hence, forecast quality will play a central role in making wise *capacity shift* decisions. Of particular interest is how the re-optimized schedule performs when varying degrees of forecast quality are available. After evaluating the effects of schedule re-optimization, we consider the question of how to design a schedule that allows maximal flexibility for adjustment during the booking process in response to unex-

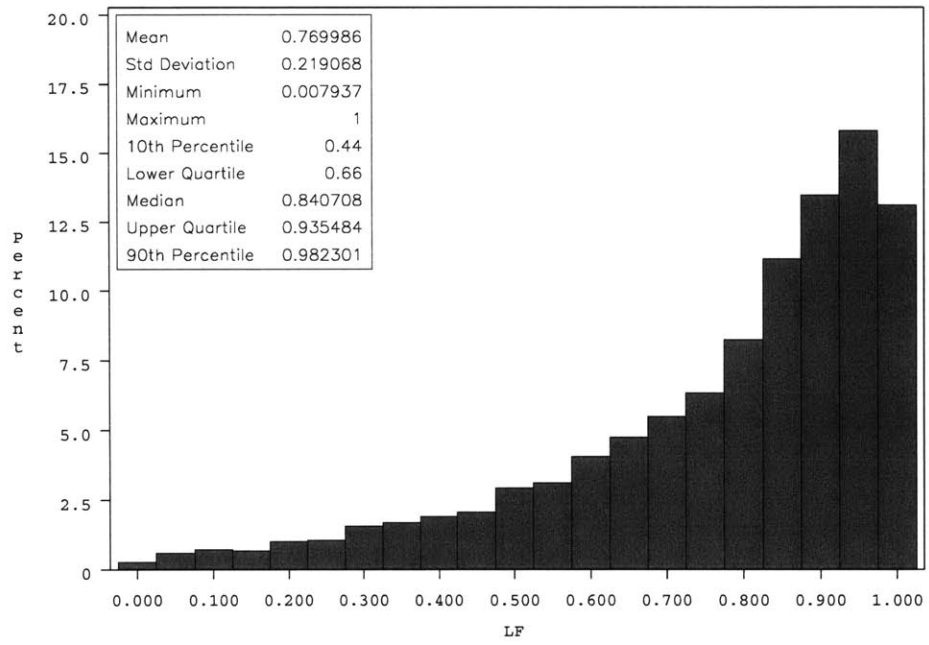


Figure 1-1: Histogram of load factors for all flights operated by a major U.S. airline in a 4-week period

pected levels of demand. Because this is a largely unexplored question, we begin by recognizing the need for metrics to gauge *schedule flexibility*, and then the need for new models and algorithms to determine schedules that maximize these schedule flexibility metrics, while simultaneously increasing realized schedule profitability.

1.2 Research Summary

In this thesis we organize the questions mentioned in the previous section into two research topics: one focuses on *schedule re-optimization*, or *dynamic airline scheduling*; and the other aims at developing *robust schedules*, that is, flexible schedules achieving maximal profitability when the schedules are allowed to be altered dynamically as passenger demands materialize during the booking process. We focus on the effects of dynamic and robust scheduling in airline networks in which one or more hubs are *de-peaked*, that is, the set of incoming and outgoing flights are interspersed, unlike *peaked* or *banked* hubs in which flight leg arrivals typically occur first, followed by period of aircraft inactivity and then flight leg departures occur. The period of inac-

tivity between arrival and departure banks allows connecting passengers to disembark from their aircraft, walk to their connecting gates and embark on their departing aircraft. The recent trend of schedule de-peaking by hub-and-spoke carriers provides us with an opportunity to develop and utilize dynamic and robust scheduling approaches that adjust the number of seats in various markets to match passenger demands at departure and maximize final schedule profitability.

1.2.1 Dynamic Airline Scheduling

The concept of dynamic airline scheduling, in which elements of the schedule are re-optimized in the booking period to reflect improved knowledge of passenger demands, dates back to the work by Etschmaier and Mathaisel (1984) and Peterson (1986), where *re-fleeting*, or *aircraft swapping*, is proposed as a dynamic scheduling mechanism. Using improved demand forecasts, the fleeting of flights in the schedule are adjusted later in the booking process to match improved demand forecasts. Berge and Hopperstad (1993) are the first to provide an in-depth presentation of this concept, providing implementation and performance evaluation details. Bish et al. (2004) and Sherali et al. (2005) later provide additional insights regarding dynamic re-fleeting approaches.

We begin with a review of the recent trend of schedule de-peaking by legacy hub-and-spoke carriers. Recognizing the high operating costs associated with peaked schedules, many legacy carriers have adopted de-peaked schedules in an attempt to cut costs through increased resource productivity. We observe that, in addition to cutting costs, de-peaked schedules can lead to increased revenues if small flight re-timings are allowed. These re-timings alter the set of connecting itineraries serving a market, and therefore, provide a mechanism for increasing the number of seats sold in markets with unexpectedly high demand, without utilizing more aircraft or crew resources. We develop a schedule re-optimization model that combines both flight leg re-fleeting and re-timing. In our dynamic airline scheduling approach, the re-optimization model is used to re-design the flight schedule at regular intervals, utilizing information from both revealed bookings and improved forecasts available

at the time of re-optimization. The re-optimization model is solved by a branch-and-bound method, aided with branching on *Special Ordered Sets*. Experiments are conducted under two forecast scenarios: one with perfect information and the other with simple averages calculated from historical demands. Two sets of experiments are performed. In the first set, a same-everyday schedule is assumed and experiments are carried out for a week of operations. We determine the resulting profit improvements and report the contributions of flight re-fleeting and re-timing when applied alone or jointly. We also study the effect of forecast quality to the benefits of dynamic scheduling. Because schedule changes, especially flight re-timings, complicate operations, we conduct sensitivity analyses to determine the degradation in schedule profitability when the number of changes are limited. We also evaluate the duration of the passenger connection times for the itineraries newly created through flight leg re-timings. In the second set of experiments, a same-every week schedule is assumed and experiments are carried out on the same weekday in seven consecutive weeks to assess the potential benefit of dynamic scheduling in the absence of day-of-week demand variations.

1.2.2 Robust Airline Schedule De-Peaking

The success of dynamic scheduling not only relies on improved demand forecasts, but also on the amount of flexibility to adjust capacity in the original schedule. In the second part of this thesis, we explore ways to imbed flexibility in the original schedule and increase its *robustness*. Such a schedule is robust in the sense that it has enhanced capability to handle demand variations through dynamic scheduling.

Expanding on the observations of Berge and Hopperstad (1993) that flight re-fleeting opportunities can be abundant in hub-and-spoke networks, we expend our efforts to identify metrics to measure the amount of flight re-timing opportunities in a schedule. We then develop a two step approach to construct robust de-peaked schedules. In the first step, a *basic de-peaking model*, ignoring the potential for altering the schedule dynamically during the booking period, is solved to obtain a baseline schedule and its associated profit. In the second step, a *robust de-peaking model*

is developed to maximize the potential for new connecting itineraries to be created through schedule adjustments in the booking period, while achieving similar profits as the baseline schedule. We present and compare several reformulations of the robust de-peaking model, each with distinct mathematical and computational properties. We next present an approximate model to reduce the size of the problem. The model is solved by branch-and-bound algorithms together with a decomposition approach involving a variable reduction technique and a new form of column generation, which result in dramatically reduced problem sizes, and greatly enhanced tractability. We compare the robust and baseline schedules and report that greater profitability is achieved by our robust schedules.

1.3 Thesis Contributions

In this thesis, our contributions to the knowledge base of dynamic airline scheduling include:

- We introduce a new dynamic scheduling mechanism, that of flight re-timing, and develop a schedule re-optimization model that integrates both flight re-fleeting and flight re-timing. Experiments are conducted using data from a major U.S. airline.

- We demonstrate that dynamic scheduling improves profitability by 2.5-5%, or $18-36 million annually. We study the effects of forecast quality on these benefits and show that considerable benefits remain even when simple forecasts calculated from historical data are used. We also report that the full benefit of re-timing is achieved even when the number of flight legs that are re-timed is strictly restricted.

- We compare and analyze the effectiveness of flight leg re-timing and re-fleeting, our two dynamic scheduling mechanisms, when applied alone under different forecast scenarios. Flight re-timing demonstrates less sensitivity to the deterioration of forecast quality and contributes a larger portion to the potential

benefit of dynamic scheduling.

- We show that benefits remain significant when dynamic scheduling is applied to weekly schedules, in which day-of-week demand variations are explicitly considered in constructing the schedules.

In this thesis, our contributions to the knowledge base of robust scheduling include the following:

- This work represents the first research effort of its kinds in which flexibility is built into the original schedule to facilitate later application of dynamic scheduling.

- We present a mathematical model and several reformulations to achieve this schedule robustness. By studying the mathematical and computational properties of these reformulations, we devise new solution algorithms and conduct experiments using data from a major U.S. airline. We show that a robust schedule further improves profitability of dynamic scheduling by an additional 1%.

1.4 Thesis Organization

We organize the remainder of this thesis as follows. In Chapters 2 and 3, we study the topics of dynamic airline scheduling and robust schedule de-peaking, respectively. Future research directions extending elements of this thesis are detailed in Chapter 4.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 2

Dynamic Airline Scheduling

2.1 Introduction

It has been a major challenge for airlines to design a flight schedule, that is, the timetable and the corresponding fleeting on each flight leg, to match fluctuating passenger demand. The flight schedule, defining the supply side of the passenger air transportation system, is designed well in advance, typically six months to one year prior to its implementation, due to contractual and operational requirements in the industry. The design process to generate a schedule that has the maximum profit potential utilizes macro-forecasts of the economy and the airline industry, forecasts of passenger demand, estimates of average fares, and estimates of available resources, such as aircraft, personnel, and gates. The resulting schedule is published through different distribution channels. Despite the fact that scheduling decisions are made at a time when demand is highly uncertain, the flight schedule is intended to remain unchanged once published. More often than not, the published schedule fails to allocate the *optimal* number of seats, that is, capacity, to where it is needed.

During the *booking period*, that is, the time between the date the schedule is published and the departure date, airlines employ different revenue management techniques to maximize the schedule's revenue. By increasing fares on highly demanded flights to decrease low fare demands and reducing fares on flights with excess capacity to stimulate travel, revenue management techniques help to smooth demand varia-

tions. Notwithstanding these techniques, the stochastic nature of passenger demand still results in some revenue being lost due to non-optimal allocations of capacity.

To achieve the goal of balancing supply and demand, researchers have started to focus their attention on the supply side and the concept of dynamic airline scheduling was born, that is, the flight schedule is re-optimized during the booking period using improved demand forecast. An early discussion of the concept of dynamic airline scheduling can be found in Etschmaier and Mathaisel (1984) and Peterson (1986). In a survey of aircraft scheduling problems, Etschmaier and Mathaisel (1984) mention dynamic scheduling as an emerging operating philosophy, where the exact schedule could be made as the total demand situation evolves. Peterson (1986) proposes the idea of re-fleeting the schedule during the booking period to better match updated forecasts. Such re-fleeting is allowed only within the same fleet *family*, which is a set of crew-compatible aircraft types. Hence, any pilot qualified to operate one fleet type within a family is, by definition, qualified to operate all fleet types in that family. The requirement of re-fleeting within families is critical because it ensures that the crew assignment based on the initial schedule can remain intact after re-fleeting. This idea is developed, implemented and tested by Berge and Hopperstad (1993) as "Demand Driven Dispatch" (D^3). The re-fleeting problem is formulated as a multi-commodity network flow problem and heuristics are developed to solve it. In the simulation study, several planning points are set in the booking period. At each planning point, the simulator gathers incremental booking information since the last planning point. Based on cumulative bookings received before the current planning point and historical information, an updated demand forecast is generated. Aircraft are re-assigned to all flight legs using the new forecast and leg capacities are updated in the reservation system. It is assumed that booking demand follows a normal distribution truncated at zero and is specified by flight leg and fare class. It is further assumed that there are no recapture of passengers and no cancelation of booked passengers. Berge and Hopperstad (1993) evaluate the approach with computational experiments performed on a network including 22 airports, 40 aircraft representing three models from the Boeing 737 family and 244 flights per day. An

improvement of 1-5% in operating profits using D^3 is reported.

Bish et al. (2004) further restrict re-fleeting to be aircraft swaps between two *swappable loops*, each of which consists of a round-trip originating and terminating at a common airport with similar time frames. Such a restriction ensures that the aircraft assigned to those routes can be swapped without violating aircraft flow balance. Two swapping strategies are analyzed, with one strategy allowing swapping only once prior to flight departure and the other allowing multiple swaps prior to departure. The conditions under which the different strategies are effective are studied.

Recently, Sherali et al. (2005) present a Demand Driven Re-fleeting (DDR) model for a single fleet family. The re-fleeting model developed is essentially the Itinerary-Based Fleet Assignment Model (see Barnhart et al., 2002) with additional constraints to limit fleeting decisions within a specific family. While the assumption of no re-capture in Berge and Hopperstad (1993) is maintained, the assumption of leg-fare class-based passenger demand is relaxed and path-fare class-based passenger demand is considered. Several reformulations and partial convex hull construction mechanisms are developed, together with various classes of valid inequalities to tighten the DDR formulation. Improvements in computational speed are reported, yet the benefit of DDR is not quantified.

To the best of our knowledge, all past research in the area of dynamic scheduling relied solely on flight re-fleeting. In this chapter, we introduce a new mechanism, referred to as flight re-timing, in which the timetable of the original schedule is altered slightly during the booking period to adjust the supply of seats provided in various markets. By shifting the arrival time of a flight leg inbound to a hub, and the departure time of an outbound flight from the hub, the set of feasible passenger connections may change, thereby increasing or decreasing the number of seats available in the affected markets.

A dynamic scheduling approach for airlines that integrates both flight re-fleeting (also referred to as aircraft swapping) and flight re-timing is developed. In this approach, the two dynamic scheduling mechanisms are carried out one or more times during the booking period. The goal is to adjust, for each day, the capacity provided

so that it will match the particular passenger demand realizations for that day more closely. These adjustments are made well in advance, perhaps 3-4 weeks prior to implementation of the schedule, to allow sufficient time for maintenance and crew planning. Another aspect that differentiates our approach from past research is that we relax the no recapture assumption and model partial recapture in our models.

The remainder of this chapter is organized as follows. In Section 2.2, we present the evolution of airline route networks and flight schedules in the U.S. to illustrate industry trend. The opportunity in a de-peaked hub-and-spoke network is discussed in Section 2.3. We detail our modeling architecture for dynamic airline scheduling in Section 2.4 with corresponding mathematical formulations in Section 2.5. Solution algorithm approach and computational experiences are presented in Section 2.6. In Sections 2.7 and 2.8, we present the setup of two case studies and the results of our computational experiments. Section 2.9 provides a review of important elements in this research. Finally, we conclude this discussion and summarize our findings in Section 2.10.

2.2 Airline Route Networks and Flight Schedules

In passenger air transportation, cities are connected to each other by flights. A *route network*, or *network* in short, describes how the cities are connected. *Flight schedule*, or *timetable*, describes the flight departure and arrival times in a route network. In this section, we review the evolution of airline route networks and flight schedules.

2.2.1 Linear Networks

Prior to deregulation in the U.S., many airlines operated over point-to-point networks in which passengers are transported directly from point of origin to point of destination without intermediate stops. This is because under regulation, there was pressure on the airlines from local communities and the Civil Aeronautics Board (CAB) to provide these direct, point-to-point services. Any airline that chose not to exercise its franchise for nonstop service in a particular market (or origin-destination pair) took

the risk of the CAB revoking that airline's permission to serve that market.

Many city pairs, however, did not have sufficient demand to cover the cost of nonstop service. Therefore, cities were often added on either end of a nonstop route to create *backup markets*. Demands in backup markets were used to fill seats on the nonstop leg. Revenue from these backup markets helped to make the nonstop service economically viable. The inclusion of these backup markets resulted in an evolved network structure, referred to as as a *linear network*. In linear networks, an aircraft begins at an origin airport and makes a number of intermediate stops along its route to a destination airport. The intermediate stops are made either to refuel or to pickup and discharge passengers.

2.2.2 Hub-and-Spoke Networks with Banked Flight Schedules

Deregulation in the U.S. has led to significant changes in the airline route networks. After deregulation, airlines quickly adopted *hub-and-spoke* networks. In hub-and-spoke networks, airlines designate several, typically large, cities as their hubs. Nonstop flights between smaller cities are substantially reduced; instead flight services for smaller cities are provided by connecting two nonstop flight legs to and from a hub airport. The major advantage of hub-and-spoke networks compared to other network structures is the disproportionately large number of city-pairs that can be served with a given number of aircraft miles operated (see Wells and Wensveen, 2004, chap. 12).

Besides providing non-stop services between hub and spoke cities, hub-and-spoke networks provide significant numbers of connecting, multiple-flight leg services between spoke cities. The result is that on a flight into or out of a hub, the airline can serve nonstop passengers between the spoke and the hub airport, and connecting passengers between two spoke cities. The consolidation of traffic leads to higher *load factors* (the portion of aircraft seating capacity that is actually sold and utilized), and, in some cases, makes it economically viable for airlines to increase frequency of flight legs in certain markets, or to operate larger aircraft with lower unit costs. The

result is lower ticket costs and/or increased frequency of service for passengers.

In hub-and-spoke networks, airlines typically operate *banked schedules*, in which are a set of arriving flights occurring in a relatively short period of time, followed by a set of departing flights also occurring within a short period of time. The amount of time separating the flight arrivals and the flight departures is defined by the amount of time needed for passengers to transfer, that is, connect between arriving and departing flight legs. Figure 2-1 shows the departure and arrival operations of a major U.S. airline at a banked hub. In this operation, there are 11 easily identifiable banks. A positive bar corresponds to the number of arrivals in a time interval, while a negative bar corresponds to the number of departures.

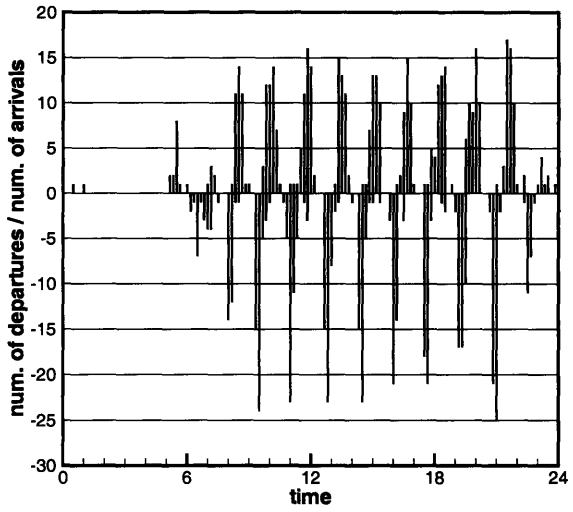


Figure 2-1: Departure and arrival activities at hub in a banked schedule

Banked schedules create departure and arrival peaks at the hub, with each peak planned to last about 45-60 minutes. Peaking has associated negative economic impacts. For example, in order to process passengers and baggage during peak operations, staffing requirements at the gate, on the apron, at the ticket counter, and in baggage handling, as well as infrastructure requirements (that is, runway capacity, gate capacity, baggage handling equipment, etc.) are at a maximum. Between adjacent banks, however, there are typically 60 to 90 minutes of *quiet* time. During this

period, staffing requirements are low and hence, labor, equipment and infrastructure are not fully utilized.

Banked operations, with their peak demands for infrastructure capacity, exacerbate the effects of congestion and delays. When bad weather conditions reduce airport capacity, two pronounced effects result. First, schedule delays are incurred, thereby increasing airline operating costs due to the extra crew and fuel costs when aircraft are queued to arrive or depart. Second, passenger travel times are increased, sometimes significantly if flight delays result in passengers missing their flight connections.

Yet another disadvantage of banked operations is the resulting reduction in aircraft productivity. First, because inbound flights in a common bank arrive at the hub at about the same time, aircraft operating flight legs with shorter flying times must wait at spoke cities and depart later than aircraft operating flight legs with longer flying times. This waiting time is non-productive time for aircraft, and very expensive to the airlines. Second, aircraft (especially those arriving early in the bank) must sit on the ground well beyond the minimum time needed to *turn* the aircraft, that is, the minimum time needed to re-fuel, disembark and embark passengers, and service the aircraft. Moreover, because the arrival times of inbound flights in a bank are coordinated, the departure times at spoke cities might not be convenient to passengers. Similarly, because the departure times of outbound flights are coordinated, the arrival times at spoke cities again might not be convenient to passengers.

2.2.3 Moving Toward De-Peaked Schedules

The airline industry in the U.S. has been negatively impacted since 2001 by terrorist attacks, overcapacity, soaring fuel costs, and stiff competition. As a result, airlines have been forced to look for new approaches and strategies to achieve profitability.

Banked operations at hubs are very expensive, requiring a surplus of labor, equipment and infrastructure. *De-peaked* operations at hubs (also referred to as continuous or rolling hubs) can continue to take advantage of the hub-and-spoke network structure, but with less intensive operations at the hub and therefore less cost (Donoghue, 2002; McDonald, 2002). In de-peaked schedules, arrival and departure operations

at the hub are smoothed, shaving the peaks and filling in the valleys of demand for resources. Flights are not coordinated in de-peaked operations to form connecting banks. Instead, the amount of time aircraft remain on the ground at the hub is not constrained by the need to provide passenger connections. Figure 2-2 shows the departure and arrival activities for the same airline shown in Figure 2-1 after the de-peaking of the same hub.

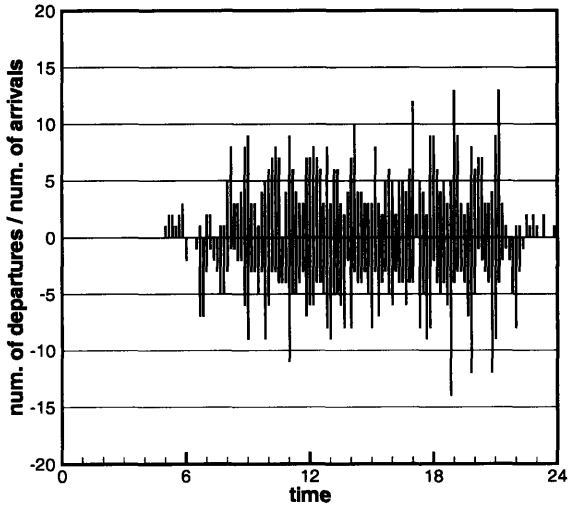


Figure 2-2: Departure and arrival activities at hub in a de-peaked schedule

In summary, the benefits resulting from de-peaking hub operations include:

- Hub staffing can be reduced because the maximum number of arrivals and departures occurring in a period of time for de-peaked operations is significantly smaller than that for peaked operations.
- Demands for infrastructure capacity, that is, airport runway and groundside capacity, gates, baggage handling equipment, etc., are similarly reduced in de-peaked operations.
- De-peaked schedules are more robust in the sense that airport capacity reductions caused by weather will have less of an impact than in the case of peaked operations with their high-levels of peak demand for capacity.

- In a de-peaked schedule, aircraft need not wait on the ground for connecting passengers. These reduced ground times for aircraft lead to increased aircraft utilization.

The major drawback of de-peaked operations is that the smoothed and spreaded-out arrivals and departures typically result in increased connection times for passengers. The effects of these increases on revenue are hard to quantify. Prior to the late 1990's, air tickets were mostly sold by travel agents using Global Distribution Systems (GDSs). Itineraries were displayed in increasing order of elapsed time, and the majority of bookings occurring in the first screen. Hence, increasing connection times, and thus, increasing elapsed travel time could displace an itinerary from being displayed in the first screen and result in significant reductions in the number of bookings for that itinerary. The adverse effect of this, however, is mitigated by recent changes in distribution: the Internet has been gaining great popularity among travelers. In 2004, more than 22% of U.S. airline tickets were sold through the Internet (Dorinson, 2004). In 2005, Alaska and its sister airline, Horizon, sold 34.6% of their tickets via the airlines' website and an additional 11% of the airlines' sales came via online travel sites (Gillie, 2006). Although travelers can choose to sort itineraries by schedule (departure time or elapsed time), the majority of air travel websites display search results by fare and researchers report that the fare display is most commonly used (Flint, 2002). GDSs also include fare display these days and fare display becomes the most commonly used method by agencies (Flint, 2002). Table 2.1 shows an example of the default displays of major U.S. airlines and leading online air ticket retailers. 11 out of 13 websites offer fare as the default display (or fare and schedule display simultaneously).

2.2.4 De-Peaking in Practice

This section discusses examples of schedule de-peaking in the airline industry. While it is not intended to be a comprehensive and thorough coverage of this topic, it does suffice the purpose of demonstrating this industry trend and its impact.

Website	Default Display Method
AirTran	fare and schedule simultaneously
Alaska	fare
America West	fare
American	schedule
Continental	fare
Delta	fare
Northwest	fare
Southwest	fare and schedule simultaneously
United	fare
US Airways	schedule
Expedia	fare
Orbitz	fare
Travelocity	fare

Table 2.1: Default display on the websites of major airlines and leading Internet air ticket retailers (obtained by visiting each website on March 6, 2006)

American Airlines de-peaked operations at its Chicago O'Hare (ORD) hub in April 2002. Flint (2002) reports that after de-peaking, the number of American Airlines flights remained the same as before, but arrivals and departures were more evenly spread throughout the day. In their de-peaked schedule, American Airlines restricted the number of arriving and departing flight legs per minute to no more than one.

Mean passenger connection times increased 10 minutes from 77 minutes to 87 minutes. Average aircraft turn times, however, reduced about 5 minutes, and by about 8 minutes at the spoke stations, resulting in less non-productive ground time for aircraft and, hence, increased aircraft utilization.

Benefits to American Airlines of de-peaking ORD include:

1. Fewer aircraft and gates were used to operate the same set of flight legs. It is reported that 3 mainline jets, 2 RJs, and 4 gates were saved after de-peaking ORD (Flint, 2002). These *saved* aircraft can be put to use in an expanded flight schedule, complementing the cost-saving attributes of de-peaking with revenue-gain potential.

2. On time performance was improved despite higher aircraft utilization. After spending years at the bottom of the Department of Transportation (DOT) on-time performance scorecard, American rose to second place in the second

quarter of 2002. With reductions in air traffic congestion resulting from de-peaking, block times were reduced by more than one minute at ORD, worth $4.5-5 million per year (Ott, 2003).

3. Labor efficiency increased. Less people were needed to handle the same amount of work, with each individual handling more flights per shift. The intensity of work per individual, however, does not increase because the workload is more evenly spread out through the day, unlike the peaked schedule in which periods of high levels of activity are followed by periods of little activity (Flint, 2002).

4. Revenues increased. American Airlines reported increased unit revenues resulting from a 1% improvement in the ratio of local (or, non-stop) to connecting traffic at ORD. Because flights are no longer coordinated to form banks at the hub, the departure and arrival times are set to convenient times for local markets, thereby attracting more non-stop passengers. Connecting revenues, however, declined as a result of longer connection times (Flint, 2002).

American airlines de-peaked its Dallas/Fort-Worth (DFW) hub in November, 2002. 9 mainline jets, 2 RJs, and 4 gates were saved (Flint, 2002). Because fewer gates are needed at DFW for their de-peaked schedule, American Airlines was able to move all mainline flights to Terminals A and C, both of which are on the same side of the airport. Previously, mainline flights operated in Terminals A, B, and C. In addition to the benefit to passengers of having all American Airlines flights on the same side of the airport, the airline estimates it will save at least $4.5 million annually from this consolidation (SL, 2002). American Airlines subsequently de-peaked its Miami (MIA) hub in May 2004.

Besides American Airlines, Continental Airlines de-peaked its schedule at Newark (Ott, 2003); United Airlines de-peaked its hub in Chicago in 2004, its hub in Los Angeles in 2005, and is expected to de-peak throughout the system, beginning with San Francisco (SFO) in the first quarter of 2006 (UAL, 2006); and Delta Airlines de-peaked their Atlanta hub in January, 2005, where about 65 airplanes an hour arrive and depart throughout the day. Daily departures for Delta Airlines grew to

1,051 a day under the de-peaked schedule from 970 a day prior to de-peaking, and the number of destinations served grew to 193 from 186. After de-peaking, average passenger connection times increased by about 3 minutes, up to 77 minutes from 74 minutes, and the amount of daily flying time per jet increased by about 8 percent, with the number of daily aircraft *turns* at each of the airline's gates increasing by up to 8.5 percent (Hirschman, 2004).

And, this de-peaking trend is not restricted to airlines the U.S.. Lufthansa Airlines de-peaked Frankfurt (FRA) in 2004, its biggest hub, as part of the effort to cut costs by EUR 300 million in the next two years (Flottau, 2003).

2.3 Opportunities in a De-Peaked Hub-and-Spoke Network

In a perfectly banked schedule, all inbound and outbound flights are scheduled to allow passengers to connect between any pair of arriving and departing flights in the same bank. Moreover, minor adjustments to flight arrival and departure times do not create, or eliminate, any connecting itineraries. In a de-peaked operation at a hub, however, minor adjustments to flight leg arrival and/or departure times can affect the set of connecting itineraries served through that hub. In fact, flight schedule re-timings can increase or decrease the supply of available seats in markets connecting at the hub. Figure 2-3 provides a schematic illustration of departure and arrival activities in a de-peaked hub-and-spoke network. We denote the minimum time needed for passengers to connect between flights at the hub as $MinCT$ and the maximum connection time acceptable to passengers as $MaxCT$. Note that inbound flight a cannot connect to outbound flights b and c because the associated connection times are not within the allowable limits. Re-timing flight leg b to b', however, creates a feasible connection between flight legs a and b', as shown in Figure 2-4. This re-timing has the effect of adding seats to the market served by flight legs a and b'. At the same time, re-timing flight b to b', however, may cause some connecting itineraries using b as the

outbound flight to violate the maximum connection requirement, thereby decreasing the number of seats offered in other markets. Similarly, re-timing flight c to c' creates a feasible connection between flight legs a and c' and may cause some connecting itineraries using c as the outbound flight to violate the minimum connection time requirement, thereby decreasing the number of seats offered in other markets. The re-timing of flight legs can thus be considered a a powerful mechanism, one capable of dynamically adjusting the supply of seats to better match demands as revealed through the booking period.

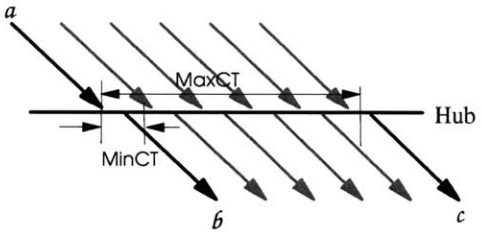


Figure 2-3: Original schedule

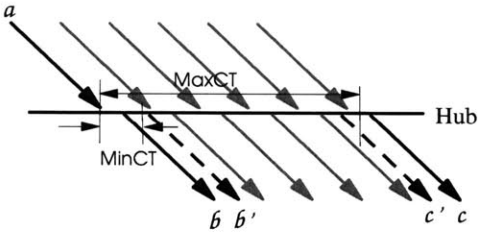


Figure 2-4: Re-timing creates new connecting itineraries

Flight re-timing has the added benefit that it can create more re-fleeting opportunities, as illustrated in the examples presented in Figures 2-5 and 2-6. Figure 2-5 depicts the original schedule, in which an Airbus A319 operates flight leg a and then flight leg c, and an Airbus A320 operates flight leg b followed by flight leg d. Suppose that more capacity is desired on flight leg d and less capacity is needed on flight leg c. The original schedule does not allow the same aircraft to operate both flight legs b and c due to insufficient turn time between the arrival of b and the departure of c. If b arrives earlier and c departs later, however, the A320 can operate b followed by c and the A319 can operate a followed by d, as depicted in Figure 2-6.

2.4 Modeling Architecture

In our dynamic scheduling approach, schedule re-optimization is performed for each departure date, thereby producing potentially different flight schedules, each of which is designed to capture the individual dynamics of passenger demand for that day. We

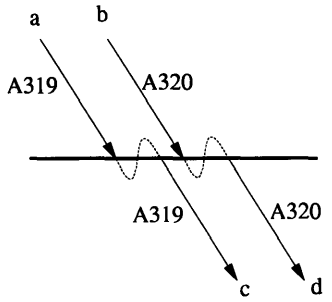


Figure 2-5: Original schedule

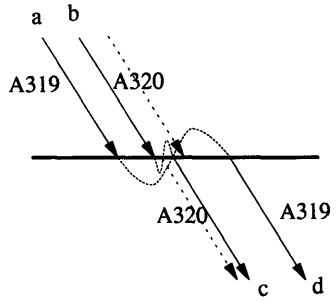


Figure 2-6: New schedule

refer to the modified schedule for each day as the *new schedule* and to the schedule produced by the initial planning process as the *original schedule.* We assume that the original schedule is a *daily schedule*, that is, the same schedule is repeated each day. *Re-optimization points* are points in time during the booking period when schedule re-optimization, that is, flight leg re-timings and re-fleetings, is performed. A portrayal of our dynamic scheduling process is shown in Figure 2-7.

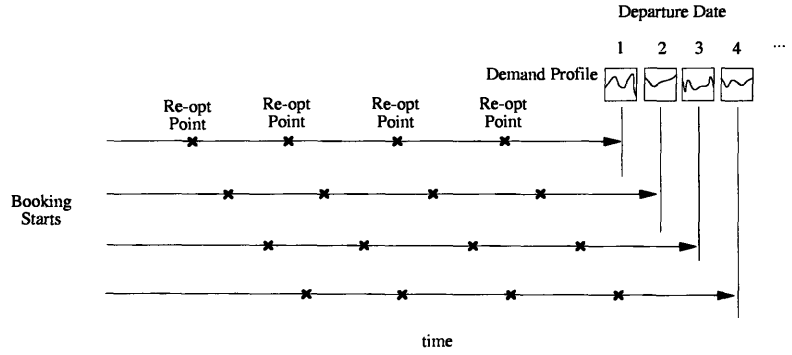


Figure 2-7: Dynamic scheduling process

For each day d included in the original schedule, we specify a few re-optimization points, with each re-optimization point earlier in the booking period of day d. Each schedule re-optimization results in a new schedule for a particular day, and it replaces the previous schedule, whether it is original or the result of re-optimization. At each re-optimization point for day d, three questions are answered, namely:

1. What are the current numbers of passenger bookings for each itinerary on day d?

2. What are the forecasted future itinerary bookings from the re-optimization point to day d?

3. What is the set of optimal flight leg re-timings and re-fleetings for day d given the current and forecasted future itinerary bookings?

Any solution to the third question must satisfy the following constraints:

1. Flight legs can be re-scheduled only to a time close to that of the original schedule. The set of allowable departure times for each flight leg defines that flight leg's feasible *time window*;

2. Allowable fleeting changes for a flight leg l are limited to fleet types in the same *family* as that of the original fleet assignment to l;

3. Service to passenger bookings made prior to the re-optimization point must be guaranteed in the new schedule;

4. At the start of the day, the number of aircraft of each type available at each airport is equal to the number positioned at that airport at the end of the preceding day. At the end of the day in the new schedule, the number of aircraft of each fleet type at each airport location must be no less than the number in the original schedule. Because a daily schedule re-optimization model is used, this is equivalent to constraining the number of aircraft for each fleet overnighted at each airport to be no more than that in the original schedule.

Constraints (1) limit the magnitude of schedule changes to minimize the impact to passengers who booked their itineraries before the re-optimization point and the possibility to disrupt aircraft maintenance routing and crew pairing plans if they are developed prior the last re-optimization point. Constraints (2) ensure that crew assignments remain feasible after re-fleeting. Constraints (3) guarantee service to all previously booked passengers in the new schedule. The exact meaning of "service guarantee" is explained in Section 2.4.1. Constraints (4) ensure that aircraft are appropriately positioned at the end of the day so that the re-optimized schedule for the next day can be implemented.

2.4.1 Service Guarantee to Previously Booked Passengers

In order to minimize passenger inconvenience, passengers booked before the current re-optimization point are to be accommodated on itineraries with the same flight numbers, but potentially with a slight change in flight departure and arrival times. Re-fleeting does not change the timetable, therefore passengers wouldn't even notice such changes. However, re-timing affects the flight times and could potentially make the connection times for previously booked connecting itineraries shorter than the minimum required.

For nonstop passengers booked on flights prior to re-timing, the effect is a slight deviation of time for their flights. For connecting passengers booked prior to re-timing, we make sure that the re-timing decisions do not disrupt their itineraries, that is, their new connection times after re-timing are no less than the minimum connection time. For example, suppose a passenger booked the itinerary shown in Table 2.2. After re-optimization, the passenger will still be traveling on flight 254 and then connecting to flight 487 as before. The only thing that changes is the flight departure and arrival times. An example of the new itinerary is shown in Table 2.3. For the itinerary shown in Table 2.3, the connection time at the hub is 30 minutes, which is still greater than the 25-minute minimum connection time. Please note that the connection times of booked connecting itineraries are allowed to exceed the maximum connection time. When flight re-timing is limited to a small magnitude, which we envision to be ± 15 minutes from the original timetable, the maximum possible increase in connection time is 30 minutes.

Flight	Origin	Departure Time	Destination	Arrival Time
254	BOS	9:00am	HUB	12:00noon
487	HUB	12:45pm	LAX	1:45pm

Table 2.2: Itinerary prior to re-timing

Flight	Origin	Departure Time	Destination	Arrival Time
254	BOS	9:10am	HUB	12:10noon
487	HUB	12:40pm	LAX	1:40pm

Table 2.3: Itinerary after re-timing

2.4.2 Frequency and Timing of Re-Optimization Points

While outlining the concept of dynamic scheduling in the previous section, we include several re-optimization points within the booking period. The question of when to re-optimize is, itself, an optimization problem. Because the objective of our research is to provide an estimate of the potential benefits of dynamic scheduling, we include only a single re-optimization point in the booking period. Even with this simplification, the question of when to perform this one-time schedule adjustment remains. The goal in selecting the re-optimization point is to balance *flexibility to modify the schedule* and *forecast quality.* Forecast quality improves as the re-optimization point is moved later in the booking period. Flexibility to modify the schedule, however, decreases as the re-optimization point is moved later in the booking period. This difficulty stems from passenger, crew and maintenance restrictions. Later in the booking period, in order to guarantee service to the large number of passengers who have already booked, the set of feasible re-optimization decisions are substantially constrained. With respect to crews and maintenance, if the re-optimization point occurs before the crew schedules and aircraft maintenance routings are constructed, crews and maintenance requirements need not be considered in the re-optimization. If, however, the re-optimization point occurs after crew and maintenance plans have been generated, the new schedule must maintain feasibility of these plans or generate feasible alternatives.

2.4.3 Flow Charts

Using a single re-optimization point for each day d in the original schedule, for example 21 days prior to departure, the booking period is divided into two periods. We refer to the time period beginning at the start of the booking period and ending at the re-optimization point as *Period 1*; and the time period from the re-optimization point to day d as *Period 2.* Empirical results show that around 50% of the passengers have booked their itineraries 21 days prior to departure, providing valuable revealed demand information but leaving flexibility in the system.

In Figure 2-8, we depict our modeling approach for the *static* case, that is, the case

in which the original flight schedule is not re-optimized. Booking limits are set for all flight legs in Period 1 to protect seats for Period 2 passenger demand. Passenger demand in Period 1 is assigned to the original schedule using a Passenger Mix Model (detailed in Section 2.5.3). Finally, Period 2 passenger demand is assigned to the remaining capacity on each flight leg in the original schedule, again using a passenger mix model.

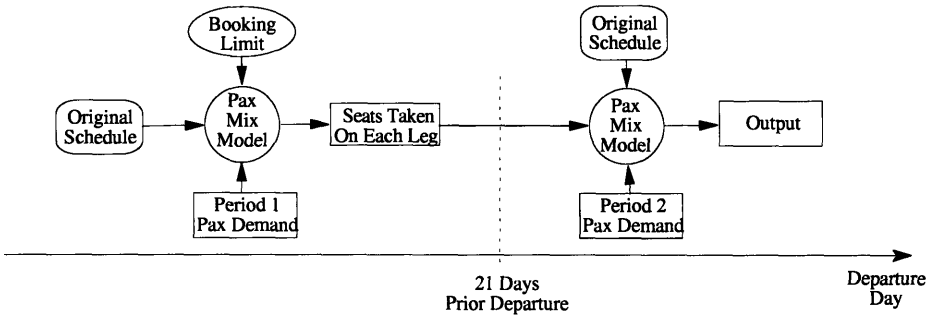


Figure 2-8: Static case

We illustrate our modeling approach for the dynamic scheduling case, or dynamic case, in Figure 2-9. A *re-optimization* module is inserted at the re-optimization point for any departure day d. The numbers of aircraft of each fleet type at each airport at the start and end of day d are computed from the original schedule. Additionally, for each flight leg, the number of seats sold in Period 1 is calculated, and the set of connecting itineraries that are booked in Period 1 are identified. These data, together with passenger demand forecasts for Period 2, are inputs to the re-optimization module that produces a new schedule to replace the original one.

2.5 Mathematical Models

In this section, we describe the network representations and the mathematical formulations of the models in our dynamic scheduling approach.

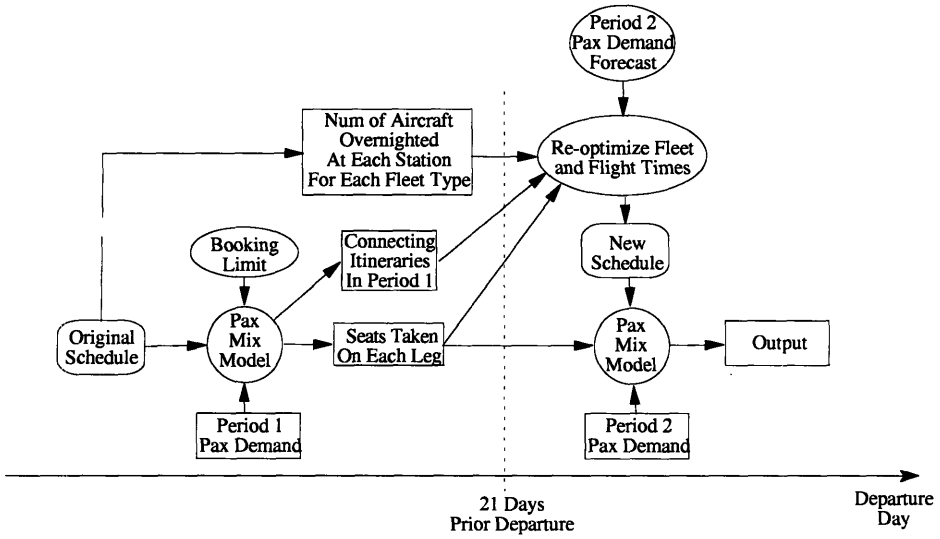


Figure 2-9: Dynamic scheduling case

2.5.1 Terminology

To facilitate our discussion, we define the following terms. A *flight leg* is a nonstop trip of an aircraft from an origin airport to a destination airport (one take-off and one landing). An *itinerary* consists of a specific sequence of scheduled flight legs in which the first leg originates from the origin airport at a particular time and the final leg terminates at the final destination airport at a later time. We model a round trip itinerary as two distinct one-way itineraries. The number of intermediate cities traversed is called the *number of stops* in this itinerary. A *non-stop itinerary* consists of only one flight leg, which originates from the origin and terminates at the destination. A *connecting itinerary* is an itinerary that has one or more stops. Although there do exist connecting itineraries with more than one intermediate stop, it is very rare. In the scope of our research, we assume all connecting itineraries have exactly one stop.

2.5.2 Network Representations

A *time-space network* is a mechanism designed for modeling dynamic networks. Nodes in a time-space network are associated with both time and place, and arcs represent scheduled movements between locations, or remaining at the same location for some

period of time.

To achieve schedule re-optimization, we create two tailored time-space networks: an *aircraft flow network* and a *passenger flow network.*

The aircraft flow network is used to model the flow of aircraft over a flight schedule, with a different flow network created for each fleet type. Each node in fleet k's network corresponds either to the departure time of a flight leg f, or its arrival time plus the minimum amount of time needed to turn a type k aircraft at the arrival station of leg f. Each arc in fleet k's network is classified as either a *flight arc* or a *ground arc.* Flight arcs represent scheduled flight legs, while *ground arcs* represent an aircraft's ability to remain on ground at the same place over time. A *wrap-around arc* is a ground arc that connects the first and last node at an airport station. The *count line* is an arbitrarily chosen point of time that is used to count the number of aircraft needed to operate a given flight schedule.

The passenger flow network is used to model the flow of passengers over a fleeted flight schedule. Each node in the passenger flow network corresponds either to the departure time, or the arrival time of leg f. Each arc is classified as either a *flight arc* or a *connection arc.* Flight arcs again represent scheduled flight legs, but *connection arcs* represent a passenger's ability to connect between two flight legs. Feasible passenger itineraries correspond to paths in this network; and flight arc capacities are determined by the type of aircraft assigned to the flight leg represented by that arc.

2.5.3 Passenger Mix Model

Given the set of *unconstrained passenger demand*, that is, the number of passengers wishing to travel, over its flight schedule with assigned fleet types, an airline's objective is to maximize their revenues by accommodating as many high fare passengers as possible. For some flight legs, unconstrained demand exceeds supply and passengers must be *spilled* either to another itinerary offered by the airline or a different airline, or to another transportation mode. When passenger demand exceeds seat capacity, the objective of the airline is to spill low fare passengers or passengers that can be easily recaptured on alternative itineraries offered by the airline.

Various difficulties arise in modeling and solving this problem of revenue maximization. Network effects, booking orders, itinerary cancellations by passengers, and the stochasticity of unconstrained demands all contribute to make this a very complex problem. In existing research, as well as the research presented here, some (but not all) of these complicating aspects are addressed. The complexities that are captured tend to be dictated by the methodology employed. For example, revenue management techniques are effective in modeling booking orders and stochasticity, while mathematical programming approaches are well-suited to model network interactions and the associated trade-offs.

Soumis et al. (1981) uses a mathematical programming model to assign passengers to itineraries. Dissatisfaction costs are assigned to unattractive itineraries and spill costs are assigned to overloaded itineraries. Glover et al. (1982) propose a minimum cost network flow model with side constraints to determine which passenger to spill. They refer to this problem as the *passenger mix problem*, which is solved by *Passenger Mix Models (PMMs)*. Other network approaches for PMMs are detailed in Dror et al. (1988), Phillips et al. (1991), Farkas (1995). Previous approaches assume either *no recapture*, that is, a passenger that cannot be accommodated on their preferred itinerary is not served by the airline; or *perfect recapture*, where each passenger is assumed to be indifferent to all itineraries serving his/her desired origin-destination city pair. Kniker (1998) represents one of the first efforts to model the more realistic case of *partial recapture* in which only a percentage of passengers who cannot be accommodated on their desired itineraries are willing to accept an alternative itinerary. These percentages, or *recapture rates*, vary for each alternative itinerary depending on its quality (that is, non-stop vs. connecting, time-of-day of departure and arrival, etc.). Hence, in this approach, passengers spilled from their desired itinerary p are assumed to be recaptured onto an alternative itinerary r at a recapture rate of b_p^r. Before we present the model by Kniker (1998), we introduce the following additional notation:

Data

P : set of all itineraries indexed by p or r.

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

$SEAT_l$: number of seats available on flight leg $l \in L$.

D_p : unconstrained demand for itinerary $p \in P$, that is, the number of passengers wishing to travel on itinerary p.

$$\delta_p^l = \begin{cases} 1, & \text{if itinerary } p \text{ traverses flight leg } l; \\ 0, & \text{otherwise.} \end{cases}$$

$fare_p$: average fare for itinerary p.

$A(p)$: set of alternative itineraries for passengers requesting itinerary p.

b_p^r : fraction of passengers spilled from itinerary p, and recaptured by itinerary $r \in A(p)$. $b_p^r \in [0,1), \forall r \in A(p), p \in P$ and $\sum_{r \in A(p)} b_p^r < 1$.

Decision Variables

x_p^r : the number of passengers requesting itinerary p who are served by itinerary $r \in A(p)$.

y_p : the number of passengers requesting itinerary p who are spilled and not recaptured by the airline.

The key-path formulation of PMM with partial recapture (Kniker (1998)) is as follows:

maximize

$$\sum_{p \in P} fare_p(D_p - \sum_{r \in A(p)} x_p^r - y_p) + \sum_{p \in P} \sum_{r \in A(p)} fare_r b_p^r x_p^r$$

Subject to:

$$\sum_{r \in A(p)} x_p^r + y_p \leq D_p, \forall p \in P \tag{2.1}$$

$$\sum_{p \in P} \delta_p^l (D_p - \sum_{r \in A(p)} x_p^r - y_p) + \sum_{p \in P} \sum_{r \in A(p)} \delta_r^l b_p^r x_p^r \leq SEAT_l, \forall l \in L \tag{2.2}$$

$$x_p^r \geq 0, \forall r \in A(p), p \in P \tag{2.3}$$

$$y_p \geq 0, \forall p \in P \tag{2.4}$$

Recognizing that $\sum_{p \in P} fare_p D_p$ is a constant, the objective can be re-written as

$$\sum_{p \in P} \sum_{r \in A(p)} \left(fare_r b_p^r - fare_p \right) x_p^r + \sum_{p \in P} (-fare_p) y_p.$$

Empirically the term $\left(fare_r b_p^r - fare_p \right)$ is usually negative, therefore in order to maximize the value of the objective function, x_p^r is pushed toward zero, if possible. The same is true for y_p. The result is that Constraints (2.1) are often redundant. This fact is exploited and a constraint generation approach is employed to solve the model.

Kniker (1998) points out that optimal PMM solutions can vary quite significantly, depending on the recapture rates used; and moreover, recapture rates are difficult to estimate. Due to PMM's sensitivity to recapture rate estimates, we adopt a simplifying approach in which we assume perfect recapture between selected subsets of itineraries, defined for each passenger. A passenger's subset includes all itineraries that: 1) depart within the passenger's window of acceptable departure times; and 2) provide a service level, measured as the number of stops in an itinerary, that is as good as that requested by the passenger. This subset of itineraries for each passenger define a *market*, with *average fare* equal to that of the itineraries in the subset. Recapture rates are 0 between itineraries in a market and itineraries serving this origin-destination city pair but not contained in the market. Before we present our PMM model, we introduce additional notation:

Data

M : set of markets.

$fare_m$: average fare in market $m \in M$.

D_m : unconstrained demand in market $m \in M$.

$R(m)$: set of itineraries serving market $m \in M$.

$$\delta_{mr}^l = \begin{cases} 1, & \text{if itinerary } r \in R(m) \text{ in market } m \in M \text{ traverses flight leg } l \in L; \\ 0, & \text{otherwise.} \end{cases}$$

Decision Variables

x_{mr} : number of passengers assigned to any itinerary $r \in R(m)$ in market $m \in M$.

Our PMM model is:

maximize

$$\sum_{m \in M} fare_m \sum_{r \in R(m)} x_{mr}$$

subject to:

$$\sum_{r \in R(m)} x_{mr} \leq D_m, \forall m \in M \tag{2.5}$$

$$\sum_{m \in M} \sum_{r \in R(m)} \delta_{mr}^{l} x_{mr} \leq SEAT_l, \forall l \in L \tag{2.6}$$

$$x_{mr} \geq 0, \forall m \in M, \forall r \in R(m) \tag{2.7}$$

2.5.4 Re-Optimization Model

The work of Berge and Hopperstad (1993), Bish et al. (2004), and Sherali et al. (2005) are most pertinent to our research in a dynamic scheduling context. Because the re-optimization model applies re-fleeting and re-timing simultaneously, a brief review of re-fleeting and re-timing in other areas of airline scheduling is presented below.

2.5.4.1 Literature Review

Research in the area of re-fleeting models and algorithms has focused largely on recovery from irregular operations, for instance, Jarrah et al. (1993), Rosenberger et al. (2002), Thengvall et al. (2000), and Yu and Luo (1997). Apart from these research, Talluri (1996) and Jarrah (2000) study models and algorithms to re-fleet a schedule that can be applied in dynamic scheduling.

Talluri (1996) addresses the problem of aircraft swapping in the context of both schedule planning and schedule recovery. Algorithms are presented to answer the following question: given a balanced and fleeted daily flight schedule, how do we change the fleet type assigned to a flight leg from Type A to Type B, while requiring

the smallest number of changes to the existing assignment. A flight network involving only Type A and Type B aircraft is constructed. Then the flight arcs and ground arcs corresponding to Type B are reversed. Each arc is assigned unit cost and a shortest paths algorithm is used to find the solution. The algorithm is able to find solutions quickly for problems involving two fleet types and the required swapping of fleet types on a single flight leg. Revenue and cost information are not considered in the algorithm.

Jarrah (2000) uses multi-type, multi-leg, re-fleeting models to modify planned fleet assignments incrementally. Several modules are defined, namely: 1) a popping module that can be used to find the least costly way to remove a specified number of aircraft from the schedule; 2) a change-of-gauge module that can be used to replace a user-specified number of Type A equipment with Type B equipment; 3) a swapping module that can be used to swap fleet assignments within a user-specified group of fleet types to maximize incremental profit; 4) a utilization module that can be used to adjust block time between two fleet types while maximizing incremental profit; and 5) a balancing module that can be used to find the most profitable set of swaps that will balance an *imbalanced schedule*, that is, a schedule in which the aircraft balance constraints are violated. A set of near-optimal solutions are determined by iteratively solving the model, with cuts added to the model at each iteration to preclude solutions from being re-generated. Let S represent the current solution; and x_i^k equal to 1 if fleet type k is assigned to leg i, and equal to 0 otherwise. The added constraints are of the form $\sum_{(i,k)\in S} x_i^k \leq |S| - 1$. The flight schedule is assumed to be fixed and passenger demands and associated fares are assumed to be flight leg specific.

Although flight re-timing has not been employed in dynamic scheduling approaches, it has been incorporated into schedule planning extensively. Levin (1971) is the first to propose a scheduling and fleet routing model with time windows. Time windows were created around the original scheduled flight departure times, with departures allowed to occur at discrete intervals within the time window. Later, Bandet (1994) and Desaulniers et al. (1997) present formulations for the fleet assignment and aircraft routing problem with time windows.

Recently, Rexing et al. (2000) present a generalized fleet assignment model to schedule flight departures and assign aircraft types to flight legs simultaneously. Two solution approaches are presented: one, the Direct Solution Approach (DST), is designed for speed and simplicity; while the other, the iterative solution approach (IST), is tailored to address large-scale problems in which memory management is an issue. In every test scenario, they produce fleet assignments with significantly lower costs than the basic fleet assignment model without time windows. Moreover, the approach is used to achieve more productive schedules for aircraft, indicating potential reductions in the number of aircraft needed to operate the flight schedule.

Klabjan et al. (2002) apply the idea of time windows to crew scheduling and develop a crew scheduling model that yields solutions with significantly lower planned costs than those obtained from conventional models. Lan et al. (2006) apply flight re-timing techniques to schedule planning models to reduce the number of occurrences of passengers missing their flight connections.

2.5.4.2 Model Statement

Our schedule re-optimization model is a generalization of models presented in Berge and Hopperstad (1993) and Bish et al. (2004) with enhanced modeling aspects. The major distinction is that in our model, *flight copies* are created and scheduled near the original flight departure time to allow for flight leg re-timing. In addition, we relax the assumption of leg demand independence and consider passenger recapture. Before introducing our schedule re-optimization model, we introduce the following additional notation:

Data and Parameters

Π	:	set of fleet types.
S	:	set of cities.
G^{π}	:	set of ground arcs in fleet $\pi \in \Pi$'s network.
$FML(\pi)$	:	set of fleet types in the same family as $\pi \in \Pi$.
D_m^F	:	demand forecast for market m.
$fare_m^F$	:	forecasted average fare for demand in market m.

$C(l)$: set of flight copies for flight leg $l \in L$.

$\langle l, k \rangle$: copy $k \in C(l)$ of flight leg $l \in L$.

$c_{lk\pi}$: cost to fly $\langle l, k \rangle$ with aircraft type $\pi \in \Pi$, where $k \in C(l), l \in L$.

c_π : fixed cost to have one additional aircraft of type $\pi \in \Pi$.

N^π : nodes in flight network of fleet type $\pi \in \Pi$.

n^π : number of aircraft available for fleet type $\pi \in \Pi$.

BKD_l : number of seats already booked on flight leg $l \in L$ before re-optimization.

$y_\pi^{i\,0}$: number of aircraft overnighted at city $i \in S$ for fleet type π in the original schedule.

π_l^0 : the fleet type used on leg $l \in L$ in the original schedule.

T : set of time interval at the hub, indexed by t.

MAX^{at} : maximum number of aircraft arrivals at the hub in interval $t \in T$.

MAX^{dt} : maximum number of aircraft departures from the hub in interval $t \in T$.

$$\zeta_{g\pi}^i = \begin{cases} 1, & \text{arc } g \in G^\pi \text{ is a wrap around arc at city } i \in S; \\ 0, & \text{otherwise.} \end{cases}$$

$$\overline{\alpha}_{lk\pi}^i = \begin{cases} 1, & \text{if } \langle l, k \rangle \text{ in fleet } \pi\text{'s network begins at node } i \in N^\pi; \\ -1, & \text{if } \langle l, k \rangle \text{ in fleet } \pi\text{'s network terminates at node } i \in N^\pi; \\ 0, & \text{otherwise.} \end{cases}$$

$$\widehat{\alpha}_{g\pi}^i = \begin{cases} 1, & \text{if ground arc } g \in G^\pi \text{ begins at node } i \in N^\pi; \\ -1, & \text{if ground arc } g \in G^\pi \text{ terminates at node } i \in N^\pi; \\ 0, & \text{otherwise.} \end{cases}$$

$$\overline{\beta}_{lk\pi} = \begin{cases} 1, & \text{if c in fleet } \pi\text{'s network crosses the count line;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\widehat{\beta}_{g\pi} = \begin{cases} 1, & \text{if ground arc } g \in G^\pi \text{ crosses the count line;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\gamma_{lk}^{at} = \begin{cases} 1, & \text{if } \langle l, k \rangle \text{ arrives at the hub during interval } t \in T; \\ 0, & \text{otherwise.} \end{cases}$$

$$\gamma_{lk}^{dt} = \begin{cases} 1, & \text{if } \langle l, k \rangle \text{ departs from the hub during interval } t \in T; \\ 0, & \text{otherwise.} \end{cases}$$

Decision Variables

$$f_{lk\pi} = \begin{cases} 1, & \text{fleet } \pi \in \Pi \text{ is used to fly flight copy}\langle l,k\rangle, \text{ where } k \in C(l), l \in L; \\ 0, & \text{otherwise.} \end{cases}$$

$y_{g\pi}$: number of aircraft on ground arc $g \in G^{\pi}$.

z_{π} : number of aircraft used for fleet type π.

The objective of the re-optimization model is to maximize future revenue less operating cost and fixed cost:

$$\sum_{m\in M}\sum_{r\in R(m)} x_{mr} fare_m^F - \sum_{l\in L}\sum_{k\in C(l)}\sum_{\pi\in\Pi} c_{lk\pi} f_{lk\pi} - \sum_{\pi\in\Pi} z_\pi c_\pi$$

The set of constraints are:

$$\sum_{k\in C(l)}\sum_{\pi\in\Pi} f_{lk\pi} = 1, \forall l \in L \tag{2.8}$$

$$\sum_{r\in R(m)} x_{mr} \leq D_m^F, \forall m \in M \tag{2.9}$$

$$\sum_{m\in M}\sum_{r\in R(m)} \delta_{mr}^{lk} x_{mr} \leq \sum_{\pi\in\Pi} f_{lk\pi}(CAP_\pi - BKD_l), \forall l \in L, k \in C(l) \tag{2.10}$$

$$\sum_{l\in L}\sum_{k\in C(l)} f_{lk\pi}\overline{\alpha}_{lk\pi}^i + \sum_{g\in G^\pi} y_{g\pi}\widehat{\alpha}_{g\pi}^i = 0, \forall i \in N^\pi, \pi \in \Pi \tag{2.11}$$

$$\sum_{l\in L}\sum_{k\in C(l)} f_{lk\pi}\overline{\beta}_{lk\pi} + \sum_{g\in G^\pi} y_{g\pi}\widehat{\beta}_{g\pi} = z_\pi, \forall \pi \in \Pi \tag{2.12}$$

$$z_\pi \leq n^\pi, \forall \pi \in \Pi \tag{2.13}$$

$$\sum_{l\in L}\sum_{k\in C(l)} \gamma_{lk}^{at}\sum_{\pi\in\Pi} f_{lk\pi} \leq MAX^{at}, \forall t \in T \tag{2.14}$$

$$\sum_{l\in L}\sum_{k\in C(l)} \gamma_{lk}^{dt}\sum_{\pi\in\Pi} f_{lk\pi} \leq MAX^{dt}, \forall t \in T \tag{2.15}$$

$$f_{lk\pi'} = 0, \forall l \in L, k \in C(L), \pi' \notin FML(\pi_l^0) \tag{2.16}$$

$$\sum_{g \in G^{\pi}} y_{g\pi} \zeta_{g\pi}^{i} \leq y_{\pi}^{i^0}, \forall i \in S, \forall \pi \in \Pi \tag{2.17}$$

$$f_{lk\pi} \in \{0,1\}, \forall l \in L, k \in C(l), \pi \in \Pi \tag{2.18}$$

$$x_{mr} \geq 0, \forall m \in M, r \in R(m) \tag{2.19}$$

$$y_{g\pi} \geq 0, \forall g \in G^{\pi}, \pi \in \Pi \tag{2.20}$$

$$z_{\pi} \geq 0, \forall \pi \in \Pi \tag{2.21}$$

Constraints (2.8) ensure that each flight leg is covered exactly once. Constraints (2.9) are passenger flow constraints limiting the number of passengers transported in each market to the value of that market's unconstrained demand. Constraints (2.10) limit the number of future passenger bookings to the remaining number of available seats. Constraints (2.11) ensure aircraft flow balance. Constraints (2.12) and (2.13) ensure that the number of aircraft of each fleet type used to operate the schedule is no more than that used in the original schedule. Constraints (2.14) and (2.15) limit the number of departure and arrival activities in each time interval to the maximum allowable. Constraints (2.16) enable re-fleeting within families. Constraints (2.17) enforces the requirement that aircraft are positioned as in the original schedule, at the start and end of each day. Constraints (2.18), (2.19), (2.20) and (2.21) are constraints on possible variable values. In the next section, we describe constraints to guarantee service to previously booked passengers.

2.5.4.3 Service Guarantee to Booked Passengers

All seats sold in Period 1 (prior to the re-optimization point) are *protected* during schedule re-optimization, that is, all nonstop passengers can be served in the new schedule on their original, although possibly re-timed, flight legs. To ensure that connecting passengers have sufficient time to connect between re-timed flights, we add constraints limiting the connection time for booked itineraries to be at least the minimum allowable. Let Q denote the set of connecting itineraries that are booked in Period 1. For the pair of flight legs $(l_1, l_2) \in Q$, let $CT(\langle l_1, k_1 \rangle, \langle l_2, k_2 \rangle)$ be the

connection time between copy k_1 of flight leg l_1 and copy k_2 of flight leg l_2. For a given $(l_1, l_2) \in Q$, let $C(l_1, l_2) = \{(k_1, k_2) | CT(\langle l_1, k_1 \rangle, \langle l_2, k_2 \rangle) \geq minCT, \forall k_1 \in C(l_1), \forall k_2 \in C(l_2)\}$. Assume we create $2a$ flight copies and index the flight copies as shown in Figure 2-10. The connection time between copy $+1$ of l_1 and copy 0 of l_2 is greater than $MinCT$, therefore the flight copy pair $(+1, 0) \in C(l_1, l_2)$. The connection time between copy $+a$ of l_1 and copy -a of l_2 is smaller than $MinCT$, therefore the flight copy pair $(+a, -a) \notin C(l_1, l_2)$.

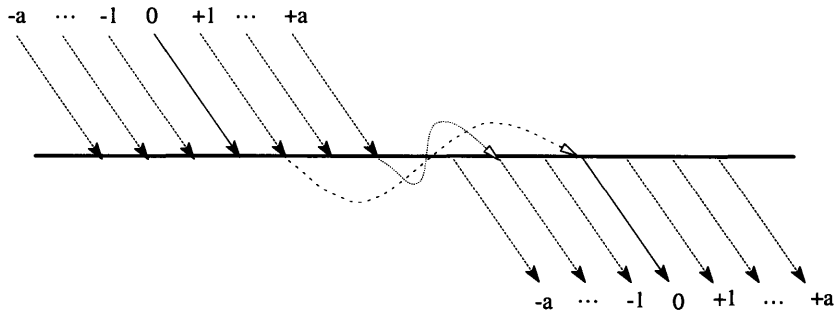


Figure 2-10: Flight copy indices

We introduce new binary decision variables $u_{l_1 k_1}^{l_2 k_2}$ for each $(l_1, l_2) \in Q$, and each $(k_1, k_2) \in C(l_1, l_2)$, where

$$u_{l_1 k_1}^{l_2 k_2} = \begin{cases} 1, & \text{if the connection formed by } \langle l_1, k_1 \rangle \text{ and } \langle l_2, k_2 \rangle \text{ is feasible after re-optimization;} \\ 0, & \text{otherwise.} \end{cases}$$

The following constraints ensure that, after schedule re-optimization, all leg pairs in Q can still form feasible connecting itineraries:

$$\sum_{(k_1, k_2) \in C(l_1, l_2)} u_{l_1 k_1}^{l_2 k_2} = 1, \forall (l_1, l_2) \in Q, \text{ and} \tag{2.22}$$

$$\sum_{\pi \in \Pi} f_{l_1 k_1 \pi} + \sum_{\pi \in \Pi} f_{l_2 k_2 \pi} \geq 2 u_{l_1 k_1}^{l_2 k_2}, \forall (l_1, l_2) \in Q, (k_1, k_2) \in C(l_1, l_2). \tag{2.23}$$

Constraints (2.22) ensure that among the feasible connections between copies of flight legs l_1 and l_2, exactly one must be enabled in the new schedule. Constraints

(2.23) guarantee that if the connection formed by $\langle l_1, k_1 \rangle$ and $\langle l_2, k_2 \rangle$ is in the new schedule, aircraft are assigned to $\langle l_1, k_1 \rangle$ and $\langle l_2, k_2 \rangle$.

Alternatively, instead of ensuring that at least one flight copy pair $(k_1, k_2) \in C(l_1, l_2)$ is enabled in the new schedule, it is also possible to forbid the selection of copy pairs not in $C(l_1, l_2)$. To achieve this goal, we only need the following constraints:

$$\sum_{\pi \in \Pi} f_{l_1 k_1 \pi} + \sum_{\pi \in \Pi} f_{l_2 k_2 \pi} \leq 1, \forall (l_1, l_2) \in Q, (k_1, k_2) \notin C(l_1, l_2). \tag{2.24}$$

Constraints (2.24) are preferred to constraints (2.22) and (2.23) because: 1) they introduce no new binary variables; and 2) they require the addition of fewer constraints to the schedule re-optimization model, as detailed in the following proposition.

Proposition 2.1 *There are fewer constraints of the form (2.24) than of the form (2.23).*

Proof: For each flight leg, equal numbers of flight copies are created, with departure times evenly spaced, both earlier and later around the departure time of the original flight leg. Assume there are $2a$ copies, indexed as shown in Figure 2-10. Consider two flight legs in the original schedule and let CT be the connection time between them. Denote the time displacement between flight copies as d, and let x be the copy index of l_1 and y be the *copy index* of l_2. We compute the connection times $CT(\angle l_1, x\rangle, \langle l_2, y\rangle) = CT - (x + y)d$ between copies of the inbound and outbound flight legs, as shown in Table 2.7. Rows represent the copy index of the inbound flight leg, while columns represent the copy index of the outbound flight leg.

For all connections in the table's lower-left triangle including the diagonal, connection times are greater than CT. Each of these entries, then, represent a constraint in Constraints (2.23). Let

$$m = \underset{b}{argmax}\{b | CT - bd \geq MinCT, b \in \overline{\mathbb{N}^-}\},$$

Index	-a	...	-1	0	+1	...	+a
-a	CT	...	CT-(a-1)d	CT-ad	CT-(a+1)d	...	CT-2ad
...	...	...	...	...	...	...	...
-1	CT+(a-1)d	...	CT	CT-d	CT-2d	...	CT-(a+1)d
0	CT+ad	...	CT+d	CT	CT-d	...	CT-ad
+1	CT+(a+1)d	...	CT+2d	CT+d	CT	...	CT-(a-1)d
...	...	...	...	...	...	...	...
+a	CT+2ad	...	CT+(a+1)d	CT+ad	CT+(a-1)d	...	CT

Table 2.7: Connection times between flight copies

then the total number of constraints in Constraints (2.24) is

$$p = \frac{1}{2}(2a + 1 - m)(2a - m),$$

and the total number of constraints in Constraints (2.23) is

$$q = (2a + 1)^2 - \frac{1}{2}(2a + 1 - m)(2a - m).$$

We have:

$$\begin{aligned} q - p =& (2a + 1)^2 - (2a + 1 - m)(2a - m) \\ \geq& (2a + 1)^2 - (2a + 1)(2a - m) \\ \geq& (2a + 1)(2a + 1 - 2a + m) \\ \geq& (2a + 1)(m + 1) \\ >& 0. \end{aligned}$$

■

We, thus, design our re-optimization model to include Constraints (2.24) and Constraints (2.8) through (2.21).

2.6 Solution Approach and Computational Experiences

The passenger mix model and the re-optimization model are both implemented in C using ILOG CPLEX 9.0. Computational experiments are conducted on a workstation equipped with one Intel Pentium 4 2.8 GHz processor and 1 GB RAM.

For a typical problem in our experiments, the passenger mix model has around 3,000 rows, 9,000 columns, and 20,000 non-zeros after CPLEX preprocessing (see ILOG, 2003, p. 322-324). The resulting formulation is solved easily, requiring only seconds of computation time, using the dual-simplex optimization routine in the CPLEX callable library.

The re-optimization model is solved by calling the mixed integer programming optimization routine in the CPLEX callable library. To improve tractability of our branch-and-bound algorithm for the schedule re-optimization model, we replace branching based on variable dichotomy with branching based on Type I Special Ordered Sets for Constraints (2.8) and assign CAP_π as weights to each variable $f_{lk\pi}$ (see ILOG, 2003). Consider one node in the branch-and-bound tree, and suppose that variables in the cover constraint for flight leg l_0, that is,

$$\sum_{k \in C(l_0)} \sum_{\pi \in \Pi} f_{l_0 k \pi} = 1,$$

take value $f^*_{l_0 k \pi}$ $(\forall \pi \in \Pi, k \in C(l_0))$, and some of these values are fractional. We compute $\overline{w} = \sum_{k \in C(l_0)} \sum_{\pi \in \Pi} f^*_{l_0 k \pi} CAP_\pi$ and partition the set of binary variables $f_{l_0 k \pi}$ $(\forall \pi \in \Pi, k \in C(l_0))$ into two groups, $G_1 = \{f_{l_0 k \pi} | CAP_\pi \leq \overline{w}, k \in C(l_0), \pi \in \Pi\}$ and $G_2 = \{f_{l_0 k \pi} | CAP_\pi > \overline{w}, k \in C(l_0), \pi \in \Pi\}$. On one branch in our special ordered set branching strategy, we impose the following restriction: $\sum_{f_{l_0 k \pi} \in G_1} f_{l_0 k \pi} = 1$; and on the second branch, we require: $\sum_{f_{l_0 k \pi} \in G_2} f_{l_0 k \pi} = 1$. Hane et al. (1995) illustrate the superiority of branching based on Type I special ordered sets over that of branching on individual variables for fleet assignment problems (although Hane et al. apply a slightly different method to create G_1 and G_2). The sizes and solution times of the

re-optimization model for a typical problem in our experiments are reported in Table 2.8. Instances A and B correspond to the two forecast scenarios introduced in Section 2.7.1.

	Instance A	Instance B
Num. of rows	16,731	21,262
Num. of columns	389,689	767,763
Num. of nonzeros	1,108,637	2,242,589
Num. of nodes searched	7,250	3,250
Solution time (hours)	10	20
Optimality gap	1.88%	2.34%

Table 2.8: Problem sizes and solution times

2.7 Case Study 1: Daily Schedules

In this section, we demonstrate the potential impact of dynamic airline scheduling, using data obtained from a major U.S. airline. The airline operates a hub-and-spoke network with a banked schedule, with about 1000 flights serving about 100 cities daily.

As a first step of our case study, we transform the banked hub-and-spoke schedule into a de-peaked daily hub-and-spoke schedule using a deterministic mathematical model (see Section 3.3). Seven copies of each flight leg are created at -30, -20, -10, 0, +10, +20, and +30 minutes offset from the leg's departure time in the banked schedule. Our model's objective is to maximize profit while satisfying flight cover constraints, aircraft balance constraints, aircraft count constraints, and de-peaking constraints (constraints that limit the number of departure and arrival activities per minute). We use the resulting de-peaked schedule as our original schedule, over which we carry out all experiments detailed in this section. The original schedule is a daily de-peaked schedule with about 300 flights originating from the major hub and the same number of flights arriving at that hub.

As shown in Table 2.9, the airline operates 7 fleets with two fleet families, namely {JET2,JET3} and {RJ1,RJ2}. During re-optimization, seven flight copies are created each at -15, -10, -5, 0, +5, +10, + 15 minutes offset from the original flight times.

Fleet Type	Capacity
JET1	190
JET2	150
JET3	124
JET4	132
RJ1	86
RJ2	50
RJ3	37

Table 2.9: Fleet composition and capacity

2.7.1 Assumptions on Unconstrained Demand and Forecast Quality

Past research in this field has assumed that unconstrained passenger demands conform to certain probabilistic distributions. Berge and Hopperstad (1993) assume that the mean and standard deviation of unconstrained demand is specified by flight leg and fare class in terms of independent normal distributions truncated at zero. In Bish et al. (2004), leg demand is considered as independent normal random variables.

The assumption that leg demands are independent is doubtful as is shown in Kniker (1998) and is inappropriate in our experiments, especially when we emphasize capturing connecting passengers through flight re-timing. However, it is difficult if not impossible to define the join probability distribution function of unconstrained market demands. Instead of making assumptions on the stochastic behavior of unconstrained market demands and drawing samples from the joint distribution for our experiments, we obtain unconstrained market demands for each individual day by running an unconstraining algorithm on *observed bookings*. Such unconstrained market demands in each day are treated as a sample from the joint market demand distribution function.

In our case study, we examine data detailing the airline's operations for one week indexed as Day 1 (a Sunday) to Day 7 (a Saturday). The cumulative demand curves for the 7 days and the average cumulative demand curve are shown in Figure 2-11, while cumulative demand curves as a fraction of total demand are depicted in Figure 2-12. At 21 days prior to departure, about $45 \sim 50\%$ of total passenger demand has

been realized. The shape of the curve reflects the booking behavior of passengers on each day: a flat curve indicates that a large portion of the passengers book well in advance; while a steep curve indicates that a larger portion of the passengers book close to departure. Day 1 (a Sunday) and Day 7 (a Saturday) have flatter curves because a large fraction of passengers traveling during the weekend are leisure travelers who tend to book early (Figure 2-12).

We study the effects of demand uncertainty on the quality of the solutions generated using our dynamic scheduling approach. We develop bounds on the potential benefits of dynamic scheduling by conducting experiments in which demand forecasts are achieved under one of two scenarios, namely:

1. the *perfect information* scenario in which future passenger demands are assumed to be known with certainty. Profit improvements estimated under this scenario provide an upper bound on the potential impact of our dynamic scheduling approach; and

2. the *historical average* scenario in which future passenger demands are estimated using a simple approach that averages historical demand data. Because airlines typically utilize more sophisticated, and hence more accurate, forecasting techniques, the impacts of dynamic scheduling estimated under this scenario provide a lower bound on the true impacts.

The quality of forecasts that are generated using average historical demand is shown in Figure 2-13. Unconstrained demand in Period 2 is depicted on the x-axis and the corresponding Period 2 demand forecasts derived using historical demand data are depicted on the y-axis. For perfect forecasts, all dots in these figures lie on the diagonal.

While Figure 2-13 provides a good visualization of forecast quality, we develop the following metrics to evaluate forecast quality. We begin by letting D_m^t be the true value of future demand in any market m and D_m^f be the corresponding forecast. We define *deviation* for market m as $D_m = D_m^f - D_m^t$; *absolute deviation* for market m as $|D_m|$; *relative deviation* for market m as $RD_m = \frac{(D_m^f - D_m^t)}{D_m^t}$; and *absolute rela-*

tive deviation for market m as $|RD_m|$. Positive deviation in a market indicates that forecasted demand is an overestimate of true passenger demand, while negative deviation indicates that true demand is underestimated. Absolute deviation measures the distance between estimated and true demand.

We partition markets into 8 groups based on their true market demand. The range of demands in each group are as follows: 0, (0,5], (5,10], (10,20], (20, 30], (30, 40], (40, 50], and above 50. In Figure 2-14, we show the average absolute deviation within each market group. Larger average absolute deviation values are observed as market size grows. In Figure 2-15, we show the average deviation within each market group. Notably, the average deviation for markets where $D_m^t = 0$ is always positive. Moreover, forecasts based on historical averages consistently overestimate demand for Day 7 (Figure 2-15). We depict the average absolute relative deviation within each market group in Figure 2-16; and the average relative deviation within each market group in Figure 2-17. When $D_m^t = 0$, relative deviation and absolute relative deviation are undefined and are therefore not reported.

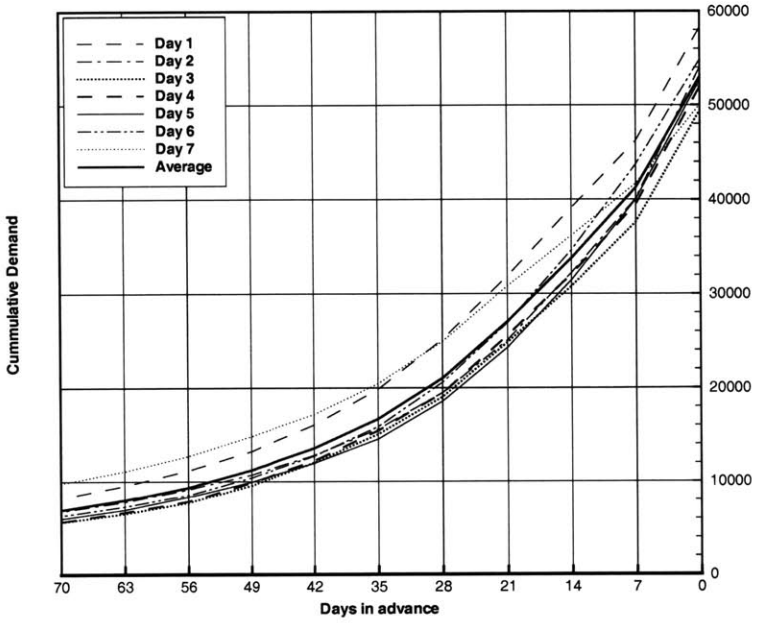


Figure 2-11: Cumulative demand

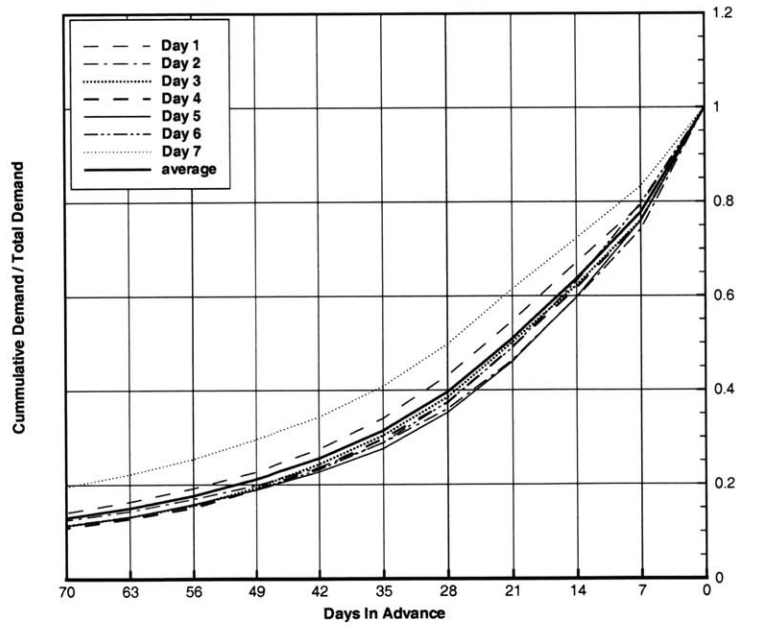


Figure 2-12: Cumulative demand as a fraction of total demand

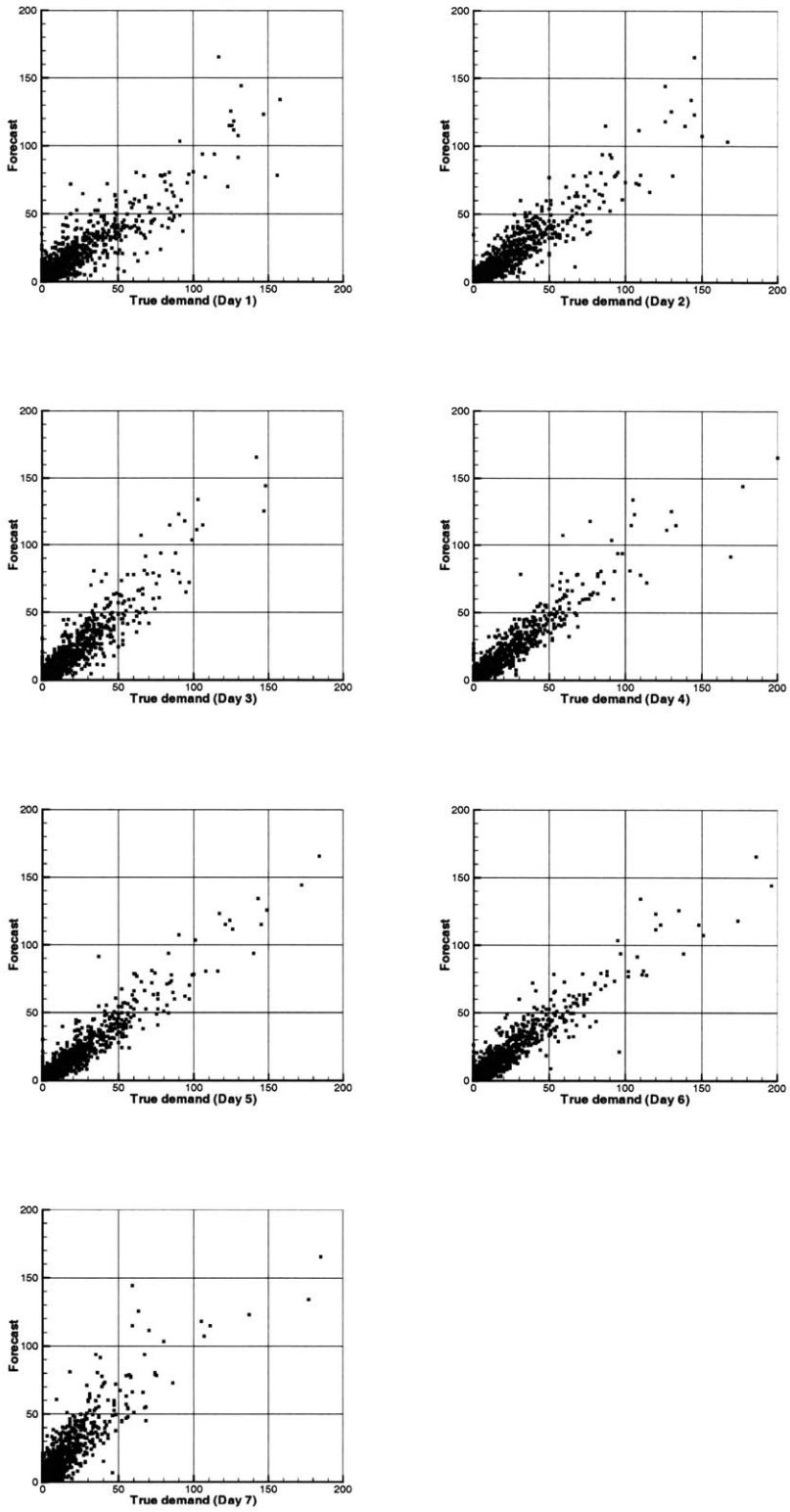


Figure 2-13: Forecast quality (scatter plot)

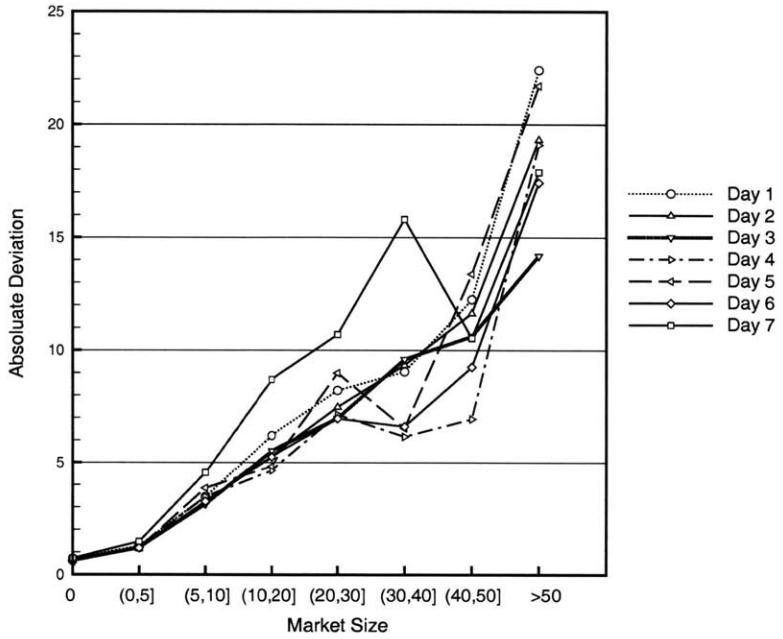


Figure 2-14: Average absolute deviation in each market group

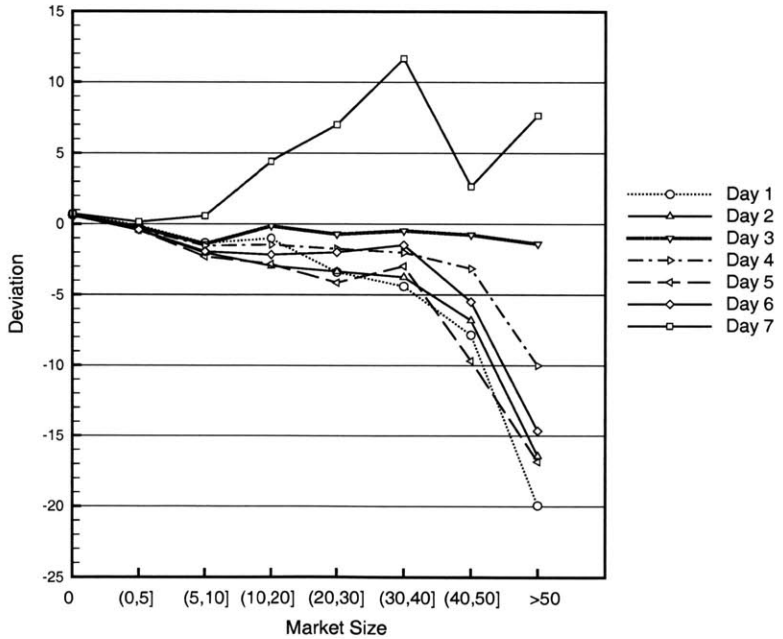


Figure 2-15: Average deviation in each market group

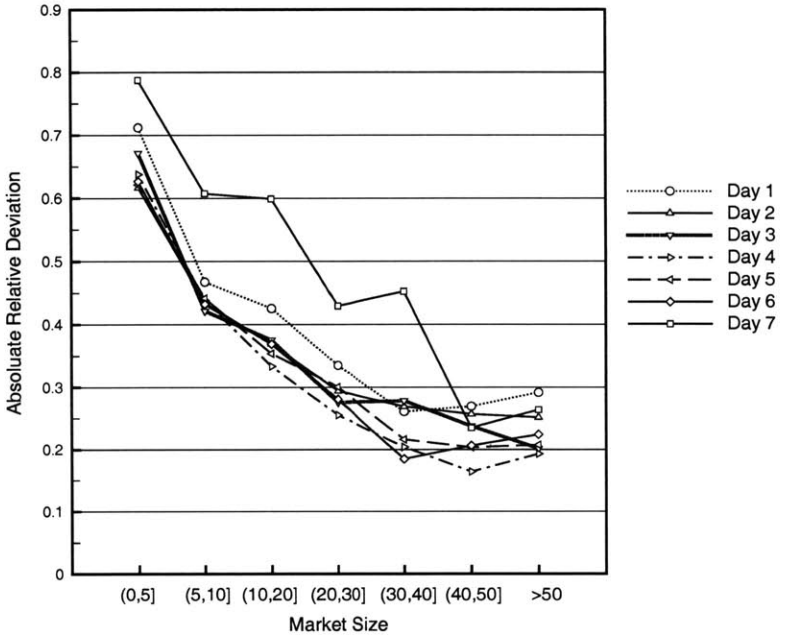


Figure 2-16: Average absolute relative deviation in each market group

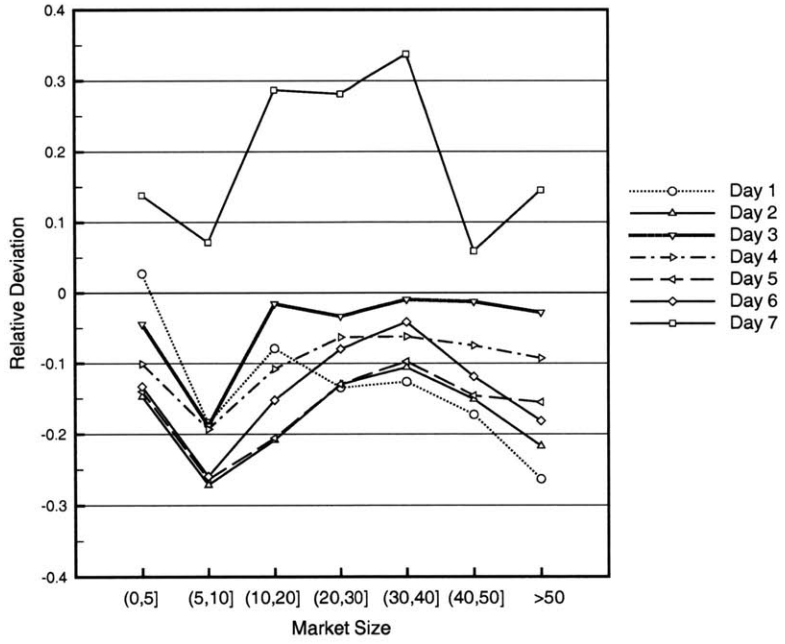


Figure 2-17: Average relative deviation in each market group

2.7.2 Results

In this section, we present the results of our experiments. We refer to the scenario with perfect information regarding future demands as *Forecast A*, and the scenario with forecast of future demands estimated using historical averages as *Forecast B*. Unless otherwise specified, when we refer to *profit increase*, *increase in revenue*, *increase in number of passengers* and similar metrics, we are measuring the change over the static case.

2.7.2.1 Profitability

Table 2.10 shows revenues, costs, and profits for the static case and the dynamic scheduling cases under Forecasts A and B, and reports the corresponding percentage changes relative to the static case. Under Forecast A, revenues increase 1-2% in each of the 7 days, while operating costs do not change significantly. The result is a 4-8% increase in profitability, or \$70-140k daily. The average profit increase is 5.26%, or \$99k daily (\$36 million annually). This result, achieved with perfect demand forecasts, constitutes an upper bound on the profitability gains achievable with dynamic scheduling. More modest improvements in revenue and profit are observed under Forecast B, specifically, a 2.64%, or \$50k increase in average daily profit (\$18 million annually) is observed. Arguably, the sophisticated forecasting tools employed by airlines result in demand forecasts that are of better quality than those used in Forecast B. This result achieved with crude demand estimates, then, represents a lower bound on the profitability potential of dynamic scheduling.

2.7.2.2 Comparison Between Re-Timing and Re-Fleeting

The two dynamic scheduling mechanisms, flight leg re-timing and flight leg re-fleeting, are examined and compared in this section. Specifically, we study the profit contributions of each, and compare the relative magnitude of their contributions.

Table 2.11 shows the results under Forecast A for dynamic scheduling (that is, re-fleeting and re-timing both), re-fleeting only, and re-timing only. Let $P^A(s)$ be

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average
				Static				
Revenue	9,058,867	8,018,606	6,870,656	7,363,064	7,946,805	8,333,631	7,176,026	7,823,951
Cost	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789
Profit	3,129,079	2,088,817	940,868	1,433,276	2,017,016	2,403,842	1,246,238	1,894,162
				Dynamic scheduling under Forecast A				
Revenue	9,217,867	8,136,941	6,934,930	7,446,343	8,037,549	8,457,232	7,263,078	7,927,706
	1.76%	1.48%	0.94%	1.13%	1.14%	1.48%	1.21%	1.33%
Cost	5,952,668	5,941,897	5,922,260	5,919,607	5,931,148	5,937,127	5,933,309	5,934,002
	0.39%	0.20%	-0.13%	-0.17%	0.02%	0.12%	0.06%	0.07%
Profit	3,265,199	2,195,045	1,012,670	1,526,736	2,106,401	2,520,105	1,329,769	1,993,704
	4.35%	5.09%	7.63%	6.52%	4.43%	4.84%	6.70%	5.26%
Profit incr.	136,120	106,227	71,803	93,461	89,385	116,263	83,531	99,541
				Dynamic scheduling under Forecast B				
Revenue	9,138,582	8,053,907	6,903,108	7,399,164	7,991,578	8,403,273	7,269,020	7,879,805
	0.88%	0.44%	0.47%	0.49%	0.56%	0.84%	1.30%	0.71%
Cost	5,929,631	5,923,936	5,916,015	5,936,996	5,934,338	5,935,877	5,972,766	5,935,651
	0.00%	-0.10%	-0.23%	0.12%	0.08%	0.10%	0.72%	0.10%
Profit	3,208,952	2,129,972	987,094	1,462,168	2,057,239	2,467,397	1,296,254	1,944,154
	2.55%	1.97%	4.91%	2.02%	1.99%	2.64%	4.01%	2.64%
Profit incr.	79,873	41,154	46,226	28,893	40,223	63,554	50,016	49,991

Table 2.10: Daily operating results under two forecast scenarios (in dollars)

the profit increase using mechanism s under Forecast A. The first observation is that P^A(re-timing + re-fleeting) $> P^A$(re-timing) + P^A(re-fleeting) for all days. The two mechanisms, then, are synergistic and achieve greater profit gains than the sum of those achieved by each mechanism individually.

Table 2.12 shows the results under Forecast B for dynamic scheduling, re-fleeting only, and re-timing only. Let $P^B(s)$ be the profit increase using mechanism s under Forecast B. P^B(re-timing +re-fleeting) $> P^B$(re-timing)+P^B(re-fleeting) is no longer true. In fact, in Days 2 and 4, P^B(re-timing + re-fleeting) $< P^B$(re-timing). However, the average profit improvement under dynamic scheduling is still greater than the sum of that under re-fleeting only and that under re-timing only. Moving from Forecast A to Forecast B, P^B(re-time) $\leq P^A$(re-time)[1] and P^B(re-fleet) $\leq P^A$(re-fleet) are observed. These results are as expected given the deterioration in forecast quality in Forecast B compared with Forecast A.

Table 2.13 shows the ratio of the profit increase under Forecast B to that under Forecast A, that is, $P^B(s)/P^A(s)$ for some s. We conclude that re-fleeting is more sensitive to forecast quality because P^B(re-fleeting)/P^A(re-fleeting) is significantly smaller than P^B(re-timing)/P^A(re-timing). When re-fleeting is applied alone under

[1] P^B(re-time) $> P^A$(re-time) by 0.05% for Day 2, which is because we set the optimality GAP in CPLEX to 0.1%.

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average
				Static				
Revenue	9,058,867	8,018,606	6,870,656	7,363,064	7,946,805	8,333,631	7,176,026	7,823,951
Cost	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789
Profit	3,129,079	2,088,817	940,868	1,433,276	2,017,016	2,403,842	1,246,238	1,894,162
				Dynamic scheduling				
Revenue	9,217,867	8,136,941	6,934,930	7,446,343	8,037,549	8,457,232	7,263,078	7,927,706
	1.76%	1.48%	0.94%	1.13%	1.14%	1.48%	1.21%	1.33%
Cost	5,952,668	5,941,897	5,922,260	5,919,607	5,931,148	5,937,127	5,933,309	5,934,002
	0.39%	0.20%	-0.13%	-0.17%	0.02%	0.12%	0.06%	0.07%
Profit	3,265,199	2,195,045	1,012,670	1,526,736	2,106,401	2,520,105	1,329,769	1,993,704
	4.35%	5.09%	7.63%	6.52%	4.43%	4.84%	6.70%	5.26%
Profit incr.	136,120	106,227	71,803	93,461	89,385	116,263	83,531	99,541
				Re-fleeting only				
Revenue	9,109,981	8,056,714	6,879,294	7,392,490	7,971,557	8,365,499	7,213,944	7,855,640
	0.56%	0.48%	0.13%	0.40%	0.31%	0.38%	0.53%	0.41%
Cost	5,940,955	5,934,028	5,925,018	5,932,752	5,931,862	5,927,903	5,941,608	5,933,447
	0.19%	0.07%	-0.08%	0.05%	0.03%	-0.03%	0.20%	0.06%
Profit	3,169,027	2,122,686	954,275	1,459,737	2,039,695	2,437,596	1,272,336	1,922,193
	1.28%	1.62%	1.43%	1.85%	1.12%	1.40%	2.09%	1.48%
Profit incr.	39,948	33,869	13,408	26,462	22,679	33,754	26,098	28,031
				Re-timing only				
Revenue	9,116,276	8,065,954	6,912,824	7,413,155	7,973,981	8,375,223	7,219,623	7,868,148
	0.63%	0.59%	0.61%	0.68%	0.34%	0.50%	0.61%	0.56%
Cost	5,929,731	5,929,673	5,929,673	5,929,673	5,929,789	5,929,673	5,929,615	5,929,689
	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Profit	3,186,545	2,136,281	983,151	1,483,483	2,044,192	2,445,551	1,290,008	1,938,459
	1.84%	2.27%	4.49%	3.50%	1.35%	1.74%	3.51%	2.34%
Profit incr.	57,466	47,464	42,284	50,207	27,176	41,709	43,770	44,297

Table 2.11: Comparison between re-fleeting and re-timing under Forecast A (in dollars)

Forecast B, there are two days (Days 2 and 3) when profits actually decline.

Examining Table 2.11 and Table 2.12, it is observed that P^A(re-timing) > P^A(re-fleeting) and P^B(re-timing) > P^B(re-fleeting) for all days. The profit increase when only re-timing is allowed is greater than that when only re-fleeting is allowed. This is because: 1) all flights can have the potential to be re-timed; however, not all flight legs can be re-fleeted; 2) under re-fleeting, the set of possible connecting itineraries remain unchanged. Re-timing can create new itineraries to increase the schedule's ability to serve a broader set of passenger demands, thus improving the ability to capture more high fare passengers and/or capture more passengers to fill empty seats.

In Table 2.14, we report the statistics about increases in revenue and number of passengers when re-timing and re-fleeting are applied alone under Forecast A. Results for Forecast A are presented because they are not complicated by the effect of forecast quality on re-timing and re-fleeting decisions, and thus clearly demonstrating the behavior of the two mechanisms. The first two rows show the increase in the numbers

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average
				Static				
Revenue	9,058,867	8,018,606	6,870,656	7,363,064	7,946,805	8,333,631	7,176,026	7,823,951
Cost	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789
Profit	3,129,079	2,088,817	940,868	1,433,276	2,017,016	2,403,842	1,246,238	1,894,162
				Dynamic scheduling				
Revenue	9,138,582	8,053,907	6,903,108	7,399,164	7,991,578	8,403,273	7,269,020	7,879,805
	0.88%	0.44%	0.47%	0.49%	0.56%	0.84%	1.30%	0.71%
Cost	5,929,631	5,923,936	5,916,015	5,936,996	5,934,338	5,935,877	5,972,766	5,935,651
	0.00%	-0.10%	-0.23%	0.12%	0.08%	0.10%	0.72%	0.10%
Profit	3,208,952	2,129,972	987,094	1,462,168	2,057,239	2,467,397	1,296,254	1,944,154
	2.55%	1.97%	4.91%	2.02%	1.99%	2.64%	4.01%	2.64%
Profit incr.	79,873	41,154	46,226	28,893	40,223	63,554	50,016	49,991
				Re-fleeting only				
Revenue	9,083,826	8,014,105	6,861,061	7,361,464	7,953,547	8,356,088	7,193,072	7,831,880
	0.28%	-0.06%	-0.14%	-0.02%	0.08%	0.27%	0.24%	0.10%
Cost	5,933,199	5,931,027	5,920,451	5,920,567	5,932,015	5,930,458	5,943,515	5,930,176
	0.06%	0.02%	-0.16%	-0.16%	0.04%	0.01%	0.23%	0.01%
Profit	3,150,627	2,083,078	940,610	1,440,897	2,021,532	2,425,629	1,249,557	1,901,704
	0.69%	-0.27%	-0.03%	0.53%	0.22%	0.91%	0.27%	0.40%
Profit incr.	21,548	(5,739)	(258)	7,622	4,516	21,787	3,319	7,542
				Re-timing only				
Revenue	9,113,048	8,067,139	6,909,084	7,396,036	7,972,271	8,359,889	7,214,381	7,861,693
	0.60%	0.61%	0.56%	0.45%	0.32%	0.32%	0.53%	0.48%
Cost	5,929,789	5,929,789	5,929,789	5,929,673	5,929,673	5,929,731	5,929,673	5,929,731
	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Profit	3,183,259	2,137,351	979,295	1,466,364	2,042,599	2,430,158	1,284,709	1,931,962
	1.73%	2.32%	4.08%	2.31%	1.27%	1.09%	3.09%	2.00%
Profit incr.	54,180	48,533	38,428	33,088	25,583	26,316	38,471	37,800

Table 2.12: Comparison between re-fleeting and re-timing under Forecast B (in dollars)

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average
P^A(re-fleeting)	39,948	33,869	13,408	26,462	22,679	33,754	26,098	28,031
P^B(re-fleeting)	21,548	-5,739	-258	7,622	4,516	21,787	3,319	7,542
P^B(re-fleeting)/P^A(re-fleeting)	53.9%	-16.9%	-1.9%	28.8%	19.9%	64.5%	12.7%	26.9%
P^A(re-timing)	57,466	47,464	42,284	50,207	27,176	41,709	43,770	44,297
P^B(re-timing)	54,180	48,533	38,428	33,088	25,583	26,316	38,471	37,800
P^B(re-timing)/P^A(re-timing)	94.3%	102.3%	90.9%	65.9%	94.1%	63.1%	87.9%	85.3%

Table 2.13: The ratio of profit increase under Forecast B to that under Forecast A when each mechanism is applied alone

of nonstop passengers and in nonstop revenues for each day compared to the static case. The next two rows show the increase in the numbers of connecting passengers and in connecting revenues in each day compared to the static case. We see that re-timing primarily captures connecting passengers.

The rest of the rows in this table analyze connecting passengers in detail. When re-timing is allowed, all connecting itineraries can be categorized into three groups:

- Itinerary Group A: these connecting itineraries are infeasible in the original schedule, but become feasible after re-timing;

- Itinerary Group B: these connecting itineraries are feasible in the original schedule, but become infeasible after re-timing;
- Itinerary Group C: these connecting itineraries are feasible before and after re-timing.

Rows 5 and 6 in this table report the numbers and associated revenues of connecting passengers traveling on Itinerary Group A after re-timing in each day. Rows 7 and 8 report the numbers and associated revenues of connecting passengers traveling on Itinerary Group B in the static case in each day. These passengers are not able to travel on the re-timed schedule and therefore are spilled or recaptured on alternative itineraries. Rows 9 and 10 report the increase in the numbers and associate revenues of connecting passengers traveling on Itinerary Group C. It is seen that because flight re-timing is able to create new itineraries to reach a larger set of passenger demands, it is able to change the revenue composition of connecting passengers and serve passengers with high revenue gains. On average, a larger amount of revenue from connecting passengers traveling on Itinerary Group A are captured than are lost through eliminating itineraries in Itinerary Group B and spilling passengers in Itinerary Group C. Markets with high revenue potential are served by reducing or eliminating service to markets with low revenue potential.

Another interesting result deserving attention is that $P^B(\text{re-time}) > P^A(\text{re-fleet})$ for all days except Day 6, implying that the profit improvement when re-timing alone is applied under Forecast B is larger than that when re-fleeting alone is applied under Forecast A.

In summary, from our experiments, we find that re-timing is less sensitive than re-fleeting to forecast quality; and re-timing contributes more than re-fleeting to the potential benefits of dynamic scheduling.

	Day 1		Day 2		Day 3		Day 4		Day 5		Day 6		Day 7		Average	
	Re-fleet only	Re-time only	Re-fleet only	Re-time only	Re-fleet only	Re-time only	Re-fleet only	Re-time only	Re-fleet only	Re-time only	Re-fleet only	Re-time only	Re-fleet only	Re-time only	Re-fleet only	Re-time only
Incr. in num. of nonstop pax	221	98	154	18	49	22	80	27	156	-12	76	22	32	6	110	26
Incr. in nonstop rev.	34,705	12,531	17,040	115	4,247	1,242	10,087	3,299	16,883	-766	7,639	3,879	4,857	-946	13,637	2,765
Incr. in num. of connecting pax	55	202	146	265	61	204	120	262	36	174	146	222	190	215	108	221
Incr. in connecting rev.	16,409	44,877	21,069	47,233	4,391	40,926	19,339	46,792	7,870	27,942	24,229	37,714	33,061	44,542	18,052	41,432
Num. of connecting pax on Itin. A	-	775	-	957	-	714	-	866	-	861	-	897	-	640	-	816
Associated rev.	-	155,813	-	174,496	-	127,481	-	159,286	-	148,771	-	165,141	-	121,638		150,375
Num. of connecting pax on Itin. B	-	334	-	358	-	403	-	383	-	376	-	391	-	244	-	355
Associated rev.	-	63,602	-	62,182	-	60,674	-	63,370	-	70,359	-	69,192	-	38,875		61,179
Change in num. of connecting pax on Itin. C	-	-238	-	-334	-	-107	-	-220	-	-311	-	-284	-	-181	-	-239
Associated rev.	-	-47,334	-	-65,081	-	-25,881	-	-49,124	-	-50,469	-	-58,236	-	-38,221	-	-47,764

Table 2.14: Passenger and revenue (in dollars) statistics under Forecast A when re-fleeting and re-timing are applied alone

2.7.2.3 Sensitivity to the Number of Re-Timed Flights

In our dynamic scheduling approach, schedule adjustments are made at each re-optimization point and finalized after the last re-optimization. In our case, the last (and the only) re-optimization point is 21-days prior to the departure day. Because schedule changes can raise many operational concerns (as detailed in Section 2.9.1), our goal is to determine a balance that minimizes the number of schedule changes while maximizing the benefits of changes. We examine the sensitivity of our results as we limit the number of allowed schedule changes, particularly the number of re-timed flights. We do not investigate the sensitivity of the solution quality to limits on the number of fleeting changes because re-fleeting within fleet families requires fewer operational changes.

Experiments are conducted under Forecast A when the number of re-timed flights is constrained. The number of re-fleeted flights is unrestricted. In Figure 2-18, we show the profit increase as a function of the total number of re-timed flights. In Figure 2-19, we show the profit achieved as a fraction of that achievable when all flight legs are allowed to be re-timed. No incremental profit increase is observed when the number of re-timed flights is allowed to exceed 300, therefore the maximum value on the x-axis of these figures is set to be 300. All curves are concave functions of the number of re-timed flights. The case when the number of re-timed flights is restricted to zero is equivalent to the re-fleeting only case. We observe that marginal profit improvements decrease as the number of re-timed flights increases; and profit improvement is nearly fully realized when the number of re-timed flights is 100. Hence, even re-timing a moderate number of the flight legs allows us to reap nearly all of the potential benefits.

Table 2.15 shows the profit increase when we limit the number of re-timed flights to 100 under Forecast B. The profit increases associated with limiting the number of re-timed flights are comparable to those when there is no limit. Interestingly, in Days 5, 6 and 7, larger profit improvements are achieved when limits are placed on the number of re-timed flights. It shows that when forecast is imperfect, having less

re-optimization flexibility does not necessarily lead to less profit increase. Had we conducted the experiments with the number of re-timed flights restricted between 0 and 100, we do not expect to see concave functions similar to those shown in Figure 2-18 and Figure 2-19.

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average
				Static case				
Revenue	9,058,867	8,018,606	6,870,656	7,363,064	7,946,805	8,333,631	7,176,026	7,823,951
Cost	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789	5,929,789
Profit	3,129,079	2,088,817	940,868	1,433,276	2,017,016	2,403,842	1,246,238	1,894,162
				Dynamic scheduling under Forecast B (No limit on num. of re-timed flights)				
Revenue	9,138,582	8,053,907	6,903,108	7,399,164	7,991,578	8,403,273	7,269,020	7,879,805
Cost	5,929,631	5,923,936	5,916,015	5,936,996	5,934,338	5,935,877	5,972,766	5,935,651
Profit	3,208,952	2,129,972	987,094	1,462,168	2,057,239	2,467,397	1,296,254	1,944,154
Profit incr.	2.55%	1.97%	4.91%	2.02%	1.99%	2.64%	4.01%	2.64%
				Dynamic scheduling under Forecast B (Num. of re-timed flights $\leq$ 100)				
Revenue	9,146,525	8,052,977	6,897,193	7,400,498	7,991,952	8,408,739	7,256,829	7,879,245
Cost	5,964,639	5,933,586	5,918,713	5,948,519	5,927,238	5,929,069	5,959,104	5,940,124
Profit	3,181,886	2,119,391	978,480	1,451,980	2,064,714	2,479,670	1,297,725	1,939,121
Profit incr.	1.69%	1.46%	4.00%	1.30%	2.36%	3.15%	4.13%	2.37%

Table 2.15: Profit increase when limiting the number of re-timed flights under Forecast B

2.7.2.4 Properties of New Connecting Itineraries

One of the benefits of flight re-timing is to create new connecting itineraries serving markets with greater than expected demand. In Figure 2-20, we show two types of itineraries that re-timing can create. In the figure, passengers cannot connect between the inbound flight and outbound flights a and d because connection times are either too short or too long. We refer to these as *infeasible connections*. When flight leg a is re-timed to b and d is re-timed to c, two new connections are created. We classify connecting itineraries enabled by schedule re-optimization as either *Type I connecting itineraries* or *Type II connecting itineraries*. Type I connecting itineraries are those that were infeasible prior to re-optimization because their connection times were less than the *minimum* allowed. Type II connecting itineraries are those whose connection times before re-optimization exceeded the *maximum* allowed. Passengers traveling on Type I (or Type II) connecting itineraries are called Type I (or Type II) passengers. It is more desirable to create Type I itineraries to serve Type I passengers, because Type II itineraries have much longer connection times and their salability is poor. In Table 2.16, we show the number of Type I and Type II passengers and the

average connection time for each type under Forecast A and Forecast B. Although no explicit incentive is given in our re-optimization model to create Type I itineraries and accommodate Type I passengers, Table 2.16 shows that about 80% of the passengers traveling on newly created itineraries are of Type I under both forecast scenarios, with an average connecting time of 30-34 minutes. The reason why the majority of passengers are Type I passengers is because when protecting connecting itineraries booked by previous passengers, inbound flights are more likely to be moved earlier, while outbound flights are more likely to be moved later. When an inbound flight is moved earlier, it creates Type I connecting itineraries; when an outbound flight is moved later, it also creates Type I connecting itineraries. The consequence is that more Type I itineraries are created and therefore, we serve more Type I passengers.

		Forecast A			Forecast B		
		Num. of pax	Pct.	Avg. conn. time	Num. of pax	Pct.	Avg. conn. time
Day 1	Type I	544	85.5%	32.4	514	74.6%	33.5
	Type II	92	14.5%	175.2	175	25.4%	144.7
Day 2	Type I	370	92.2%	31.0	782	76.5%	34.3
	Type II	31	7.8%	175.4	240	23.5%	150.9
Day 3	Type I	316	84.3%	30.5	570	73.9%	34.4
	Type II	59	15.7%	175.9	201	26.1%	164.3
Day 4	Type I	277	78.7%	31.0	581	83.4%	33.1
	Type II	75	21.3%	174.9	116	16.6%	174.2
Day 5	Type I	534	84.4%	31.9	651	84.1%	33.6
	Type II	99	15.6%	176.5	123	15.9%	175.6
Day 6	Type I	449	85.1%	31.3	394	76.5%	33.7
	Type II	78	14.9%	174.4	121	23.5%	170.5
Day 7	Type I	237	64.6%	31.4	674	91.9%	32.8
	Type II	130	35.4%	174.5	60	8.1%	174.4

Table 2.16: Statistics on Type I and Type II passengers ($MinCT = 25$ minutes and $MaxCT = 180$ minutes)

2.7.3 Quality of the Original Schedule

As stated in the beginning of Section 2.7, the original schedule used in our experiments is generated by an optimization model. In this section we investigate the quality of the original schedule and analyze whether the benefits of dynamic scheduling is resulted from a poorly designed original schedule, particularly, a poorly designed timetable.

We examine flight re-timing decisions under perfect information for the 7 days. Perfect information is used because re-timing decisions under perfect information indicate there may be better flight times for those flight legs. If a flight leg is re-

timed at least 5 out of 7 days, it is classified as frequently re-timed. If the re-timing decisions for a flight leg occurs in the same direction (that is, earlier or later than the original schedule) for all days, these re-timing decisions are called *consistent*. When a flight leg is frequently and consistently re-timed, it is likely that this flight is poorly scheduled in the original schedule. The benefits of employing dynamic scheduling techniques for this flight leg could be minimal if this flight leg had been optimally scheduled in the original schedule. In Table 2.17, we report the re-timing decisions for the frequently re-timed flight legs in our experiments. In the first row, we report that Flight leg 185 is postponed by 10 minutes on 6 out of the 7 days and by 15 minutes in one day. Flight leg 185 is clearly frequently and consistently re-timed, and hence, it might be poorly positioned in the original schedule. If, however, a flight leg is frequently, but not consistently re-timed (for example, Flight legs 50 and 216), we do not view it as poorly placed in the original schedule. Instead, variations in daily demand require it to be scheduled differently on a day-to-day basis.

Table 2.17 shows the 22 frequently re-timed flights, that is, re-timed at least 5 times. Of these 22 flights, 7 are moved inconsistently. In the worst case, then, the schedule for 15 flights, or less than 2% of the flight legs, in the original schedule is *not optimal*. Given that changes to one flight leg's schedule impacts the set of feasible schedules for a set of other flight legs, it is likely that alternative original schedules will similarly have some small percentage of flight legs that will be consistently rescheduled under dynamic scheduling approaches. We conclude, then, that our potential is very limited to improve the original schedule.

Flight #	Flight Detail					-15	-10	-5	0	+5	+10	+15	consistent?
185	SJC	1220	→	HUB	1410						6	1	
470	MCI	900	→	HUB	951	5	2						
610	YYC	725	→	HUB	937					1		6	
68	HUB	1655	→	SJC	1844				1			6	
171	HUB	1436	→	DFW	1856				1	1	4	1	
234	BUR	1525	→	HUB	1650	6			1				
263	HUB	1425	→	ABQ	1635				1		1	5	
455	STL	830	→	HUB	951	4	1		1	1			N
512	HUB	1439	→	MSP	1953				1		4	2	
600	HUB	755	→	BUR	918			6	1				
645	SAN	2000	→	HUB	2122				1	2	2	2	
50	HUB	925	→	ATL	1622		1	3	2	1			N
168	DTW	755	→	HUB	900				2	5			
195	LAS	650	→	HUB	804	4			2		1		N
216	SNA	1000	→	HUB	1119		1	3	2		1		N
220	SNA	1200	→	HUB	1320				2	2		3	
235	HUB	1812	→	ABQ	2022	1	1		2			3	N
326	LAS	1156	→	BOS	2012	3			2	2			N
417	IAH	600	→	HUB	644	5			2				
559	PDX	700	→	HUB	935	1			2			4	N
605	SNA	1715	→	HUB	1837	4		1	2				
709	EUG	555	→	HUB	825	5			2				

Table 2.17: Re-timing decisions for frequently re-timed flights

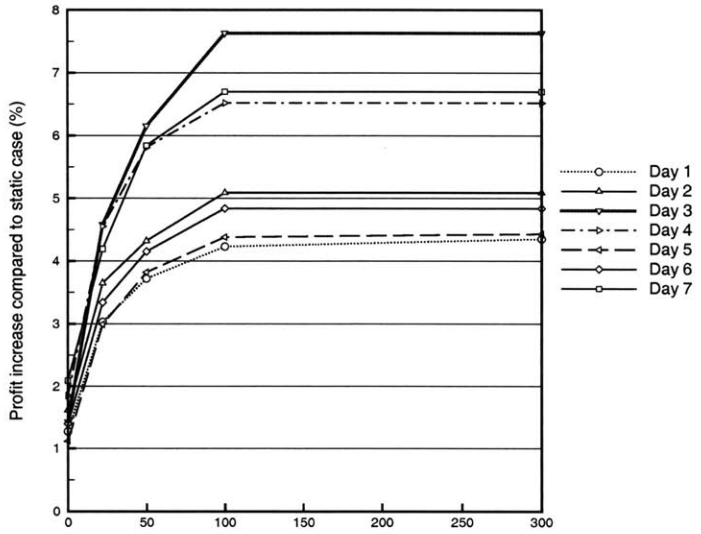


Figure 2-18: Profit increase as a function of total number of re-timed flights (Forecast A)

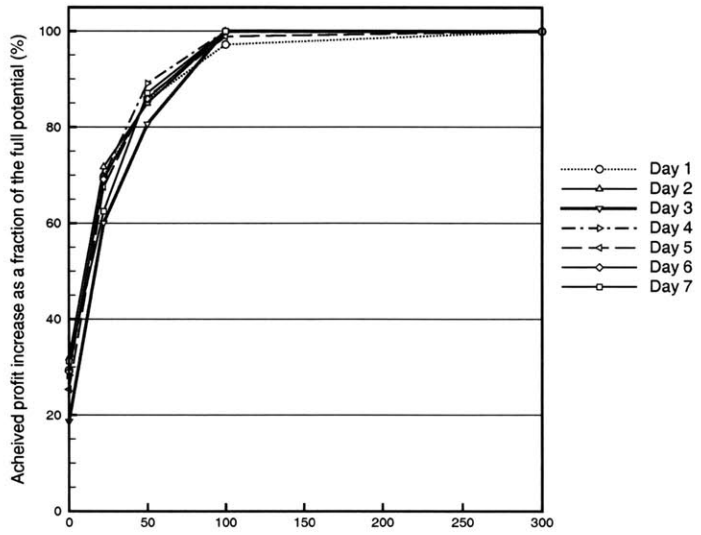


Figure 2-19: Achieved profit increase as a function of total number of re-timed flights (Forecast A)

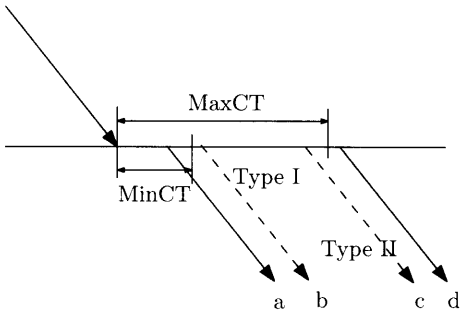


Figure 2-20: Two types of new connecting itineraries

2.8 Case Study 2: Weekly Schedules

In the previous section, a daily repeatable schedule is assumed in which the same flight schedule is repeated each day. The advantage of this assumption is reduced computational requirements and ease of implementation operationally. Given this daily schedule, dynamic scheduling can then be applied to adjust each day's schedule in response to passenger demand variations. The question we pose in this section is whether or not a weekly schedule, that is, one in which a different schedule is designed for each day of the week, is adequate to capture demand variations and eradicate the need for dynamic scheduling.

2.8.1 Schedule Generation

In Figure 2-21, we show the variations in daily flight load factors for a major U.S. airline in a 4-week period, depicting the variability in daily passenger demands. For each day, the mean load factor and several representative quantiles of the load factor histogram are shown. It is seen that the variations in load factors follow a weekly pattern. Clearly, higher revenues can be realized by flying a different schedule tailored to each day of the week, that is, a weekly flight schedule. A weekly schedule can be developed by expanding the daily time-lines of the time-space network to weekly time-lines. A weekly schedule design model, however, with its extended size, is much harder to solve than a daily schedule design model. In this thesis, we do not tackle the difficult problem of weekly schedule design, but rather generate a schedule that incorporates our knowledge about day-of-week demand variations and by solving the daily schedule design model (see Section 3.3) for each day of the week, using average demands for that day. We generate a Sunday schedule based on average Sunday demands, a Monday schedule based on average Monday demands, $\cdots$, and a Saturday schedule based on average Saturday demands. These seven daily schedules do not constitute a weekly schedule because each day's schedule is constructed independently and there is no guarantee that aircraft balance will be preserved from one day to the next. Notwithstanding this, the resulting daily schedules provide an approximation (albeit

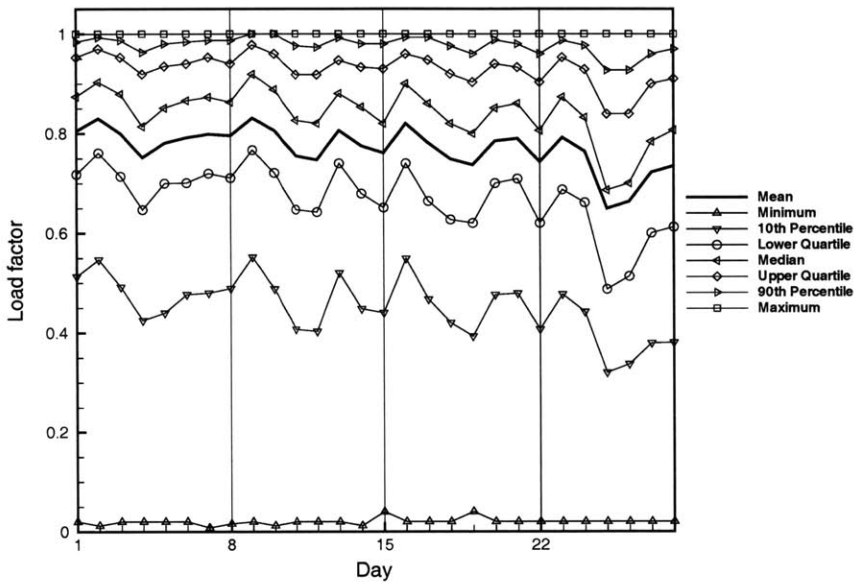


Figure 2-21: Daily mean load factor and quantiles of the load factor histogram for a major U.S. airline in a 4-week period. Days 1, 8, 15, and 22 correspond to Saturdays

an overly optimistic one) of the ability of a weekly schedule to capture day-of-week demand variations. If dynamic scheduling can still improve the economics of daily schedules constructed in this manner, it should have at least as great a benefit when applied to a weekly schedule.

In our analysis, we select a specific day of week, namely, Monday, and develop a daily schedule based on average Monday demand. Then, we evaluate the effects of applying dynamic scheduling to seven Mondays (indexed as $W1$, $W2$, $\cdots$, $W7$) in seven consecutive weeks. Although we do not perform tests for the remaining days of the week, we expect that the results would be similar.

2.8.2 Unconstrained Demand and Forecast Quality

We treat unconstrained demand and forecast quality in the same way as detailed in Section 2.7.1. In Figure 2-22, we depict the average cumulative demand curve and the

cumulative demand curves for the 7 experimental Mondays. We observe that there are variations in total unconstrained demand, despite the fact that these seven days are all the same day of the week. In Figure 2-23, we show cumulative demand curves as a fraction of total demand in each day. Not surprisingly, we note that the curves are more similar in Figure 2-23 than those for all days of one week in Figure 2-12.

Similar to our previous case study, two forecast scenarios are considered: 1) the perfect information scenario; and 2)the historical average scenario. In Figures 2-24 through 2-28, we show the quality of the forecasts generated from average historical demand. It is worth noting that the shapes of the curves in Figures 2-25 through 2-28 are more uniform than their counterparts in the previous case study (Figures 2-14 through 2-17).

Examining Figures 2-26 and 2-28, all curves are below zero for market group (0,5], (5,10], and (40,50] and above 50. It may seem odd, because there should be at least one curve above zero, however, this is because the historical average demands are averaged over twelve Mondays, but we only show seven of them in the figures, that is, the we only show the seven curves corresponding to the seven Mondays we perform experiments on.

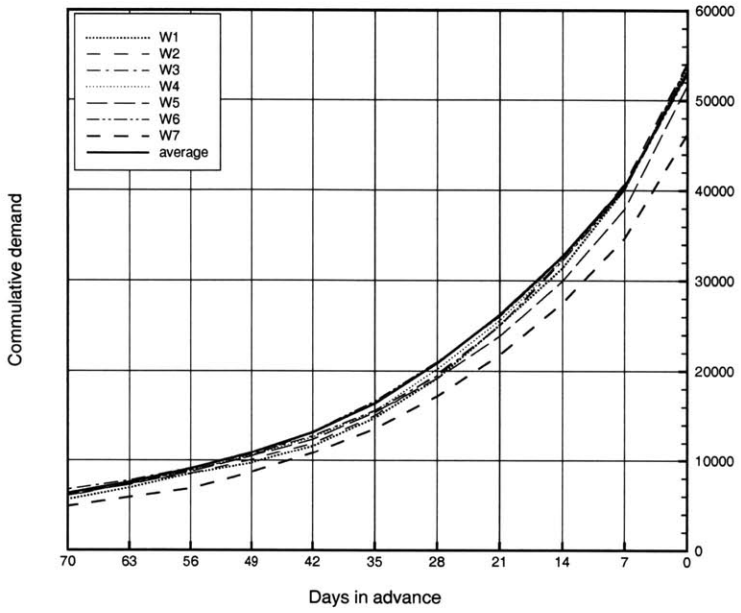


Figure 2-22: Cumulative demand curves

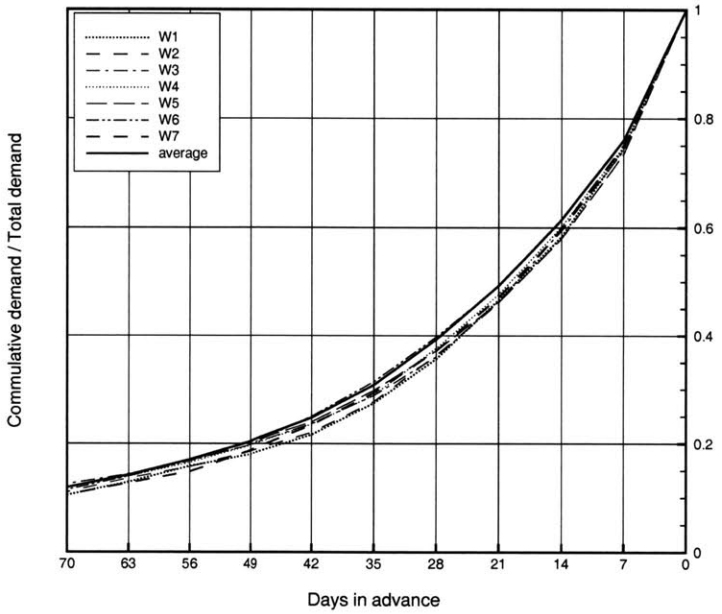


Figure 2-23: Cumulative demand curves as a fraction of total demand

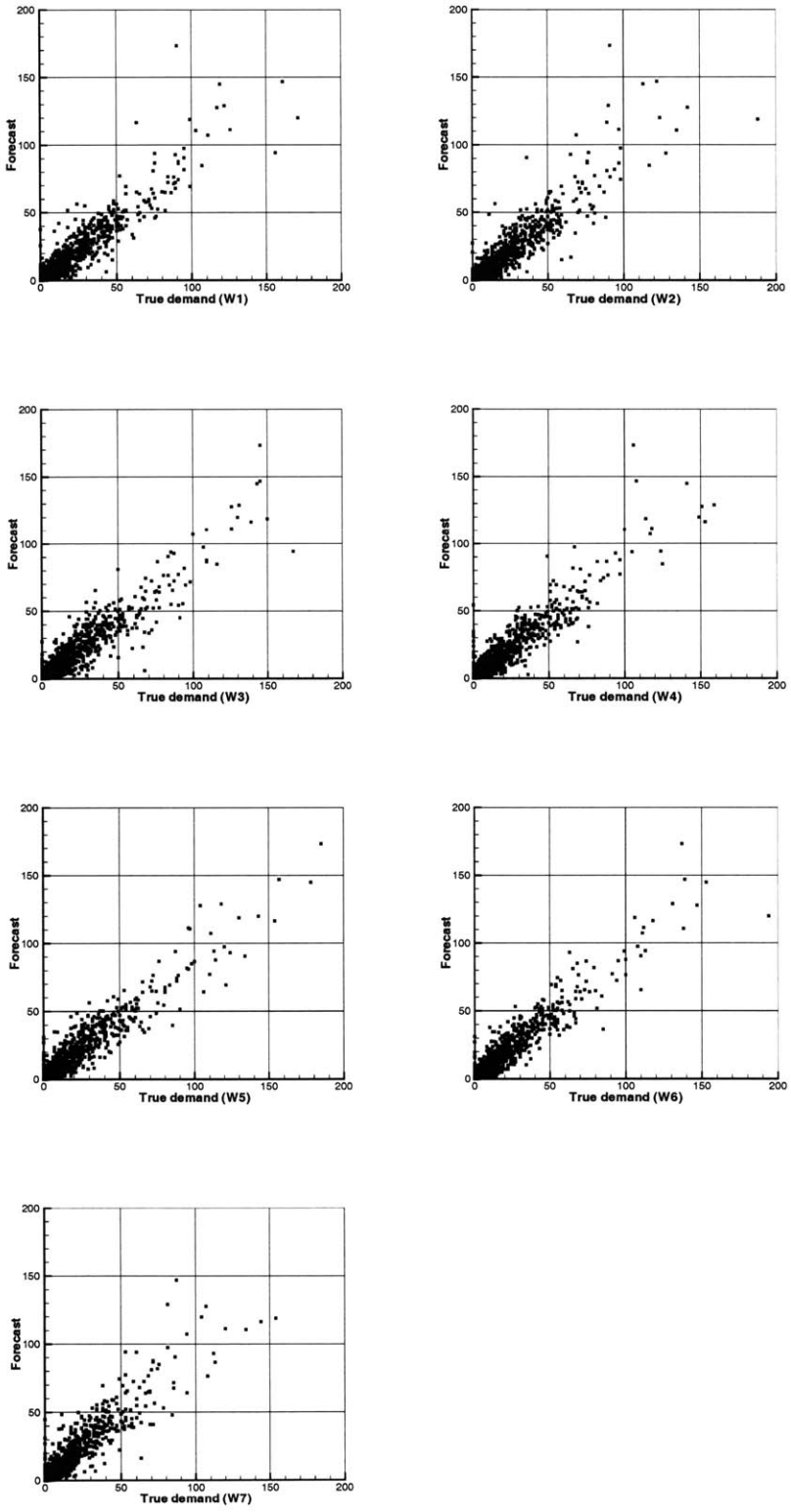


Figure 2-24: Forecast quality (scatter plot)

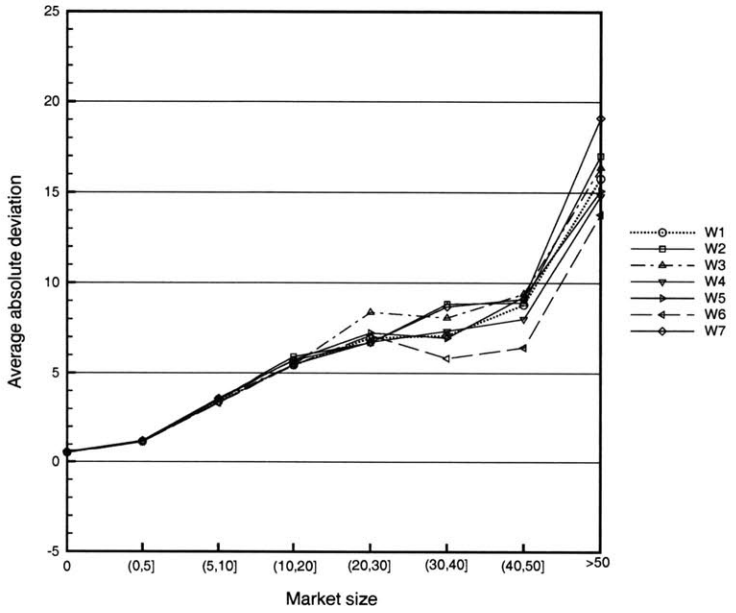


Figure 2-25: Average absolute deviation in each market group

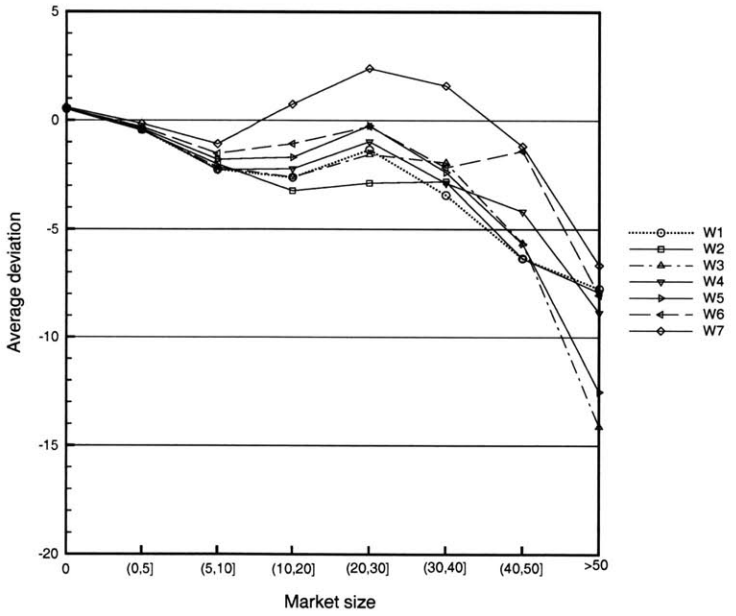


Figure 2-26: Average deviation in each market group

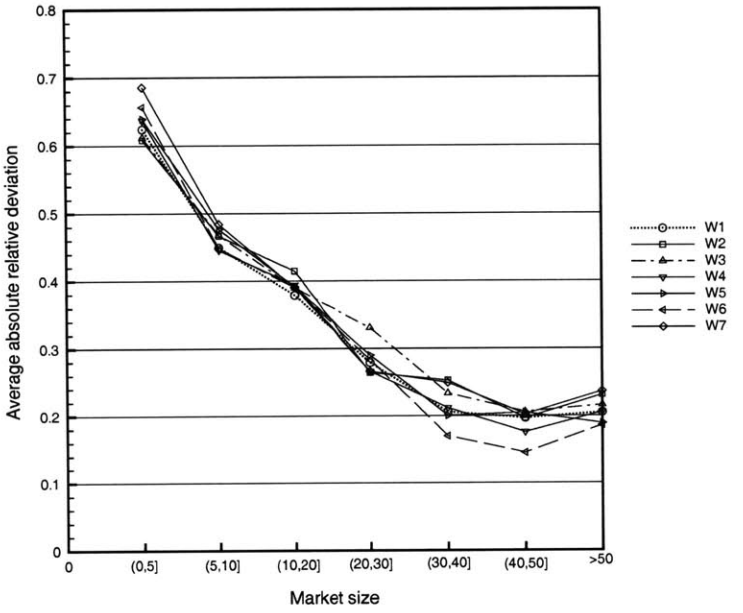


Figure 2-27: Average absolute relative deviation in each market group

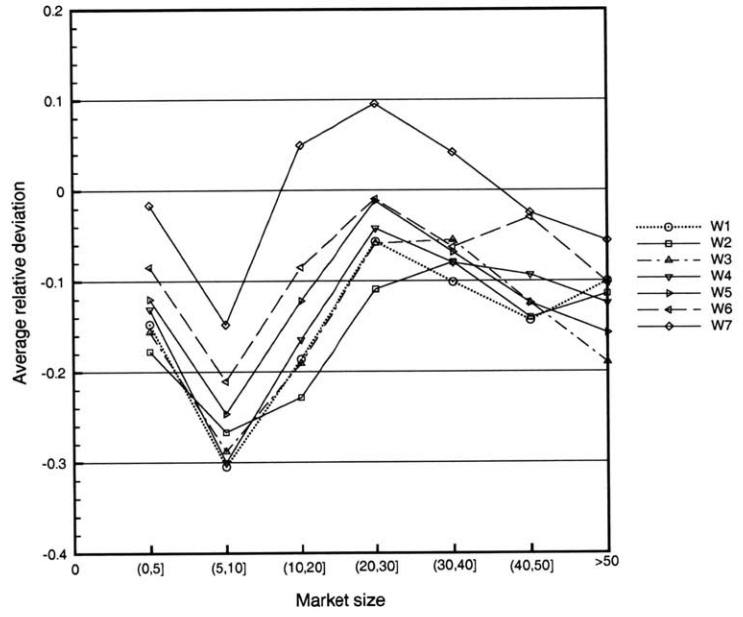


Figure 2-28: Average relative deviation in each market group

2.8.3 Results

In this section, we present the results of our experiments. Based on the results presented in Section 2.7, we limit the number of re-timed flights to 100 in all of the following experiments.

In Table 2.18, we show, for each day, the revenues, costs, and profits associated with the static schedule, and the dynamic scheduling cases under the two forecast scenarios. The benefit of dynamic scheduling remains significant. Under Forecast A, revenue goes up by 1.41%, costs increase by 0.28%, and profit improves by 4.97%, or \$92k daily (\$33 million annually) on average. Under Forecast B, revenue increases by 0.77%, costs increase by 0.29%, and profit improves by 2.28%, or \$42k daily (\$15 million annually). We believe that it is likely that the sophisticated forecasting engines used by airlines will result in demand forecasts of better quality than Forecast B, and hence, the results under Forecast B should represent a lower bound on the profitability potential of dynamic scheduling when applied to this particular day of week under weekly schedules.

	W1	W2	W3	W4	W5	W6	W7	Average
				Static				
Revenue	7,825,855	7,758,300	8,006,721	8,079,302	7,778,746	7,178,135	7,729,064	7,765,160
Cost	5,905,362	5,905,362	5,905,362	5,905,362	5,905,362	5,905,362	5,905,362	5,905,362
Profit	1,920,493	1,852,938	2,101,359	2,173,940	1,873,385	1,272,774	1,823,702	1,859,799
				Dynamic scheduling under Forecast A				
Revenue	7,939,360	7,863,389	8,122,944	8,203,797	7,878,834	7,272,242	7,839,553	7,874,303
	1.45%	1.35%	1.45%	1.54%	1.29%	1.31%	1.43%	1.41%
Cost	5,918,295	5,924,856	5,925,398	5,932,377	5,912,754	5,917,284	5,923,882	5,922,121
	0.22%	0.33%	0.34%	0.46%	0.13%	0.20%	0.31%	0.28%
Profit	2,021,065	1,938,532	2,197,547	2,271,420	1,966,081	1,354,958	1,915,672	1,952,182
	5.24%	4.62%	4.58%	4.48%	4.95%	6.46%	5.04%	4.97%
Profit incr.	100,572	85,594	96,188	97,480	92,697	82,185	91,970	92,384
				Dynamic scheduling under Forecast B				
Revenue	7,879,739	7,820,891	8,049,311	8,114,287	7,848,838	7,249,730	7,809,586	7,824,626
	0.69%	0.81%	0.53%	0.43%	0.90%	1.00%	1.04%	0.77%
Cost	5,915,842	5,918,301	5,915,032	5,915,032	5,922,767	5,920,097	5,949,482	5,922,364
	0.18%	0.22%	0.16%	0.16%	0.29%	0.25%	0.75%	0.29%
Profit	1,963,897	1,902,590	2,134,279	2,199,255	1,926,072	1,329,634	1,860,104	1,902,262
	2.26%	2.68%	1.57%	1.16%	2.81%	4.47%	2.00%	2.28%
Profit incr.	43,405	49,652	32,920	25,315	52,687	56,860	36,402	42,463

Table 2.18: Daily operating results under two forecast scenarios (in dollars)

2.8.4 Quality of the Original Schedule

We present the re-timing decisions of frequently re-timed flight legs in Table 2.19. 27 flight legs are frequently re-timed, among which 3 are not consistently re-timed. Compared to Table 2.17, we have slightly more flight legs frequently re-timed in this experiment, but the number still represents only about 2% of the total number of flight legs. For reasons stated in the previous section, we believe that our estimated profit improvements result principally from enhancements possible through dynamic scheduling rather than from an inadequate original schedule.

Flight #	Flight detail					-15	-10	-5	0	5	10	15	Consistent?
271	LAS	1210	→	HUB	1324	7							
276	HUB	1738	→	DEN	2025					2	2	3	
512	HUB	1439	→	MSP	1953						3	4	
167	LAS	2232	→	DTW	2927				1	1	3	2	
174	ONT	720	→	HUB	834				1	3	3		
185	SJC	1220	→	HUB	1410				1		3	3	
339	HUB	1346	→	ONT	1453				1		6		
606	HUB	1955	→	TUS	2037		4	2	1				
719	HUB	1258	→	FAT	1432	2		1	1	1	2		N
748	SBA	1220	→	HUB	1353	6			1				
50	HUB	925	→	ATL	1622		2	3	2				
111	PHL	1335	→	HUB	1525			5	2				
215	HUB	805	→	SNA	924	4	1		2				
216	SNA	1000	→	HUB	1119			4	2		1		N
231	HUB	1804	→	ORD	2327				2		3	2	
270	HUB	1002	→	LAS	1115	1		4	2				
293	LAS	1146	→	SMF	1317	2	3		2				
355	HUB	830	→	MCO	1543				2			5	
378	ONT	710	→	HUB	822	2	3		2				
455	STL	810	→	HUB	931		5		2				
479	PDX	1906	→	LAS	2113		5		2				
571	LAS	0	→	GEG	222	2	1	2	2				
604	HUB	1431	→	SNA	1546				2		2	3	
627	SLC	1850	→	HUB	1928		5		2				
637	HUB	1940	→	FAT	2115				2	3		2	
783	HUB	1625	→	BIL	1955				2			5	
814	YUM	710	→	HUB	810	2	1		2		1	1	N

Table 2.19: Re-timing decisions for frequently re-timed flights

2.9 Other Issues

In this section, we review two important issues with dynamic scheduling. One is the impact dynamic scheduling would have on aircraft maintenance routing, crew scheduling, and passenger itineraries; the other is the applicability of dynamic scheduling to other airlines.

2.9.1 Effects of Dynamic Scheduling on Aircraft Maintenance Routing, Crew Scheduling, and Passenger Itineraries

Changes to a published schedule impact several elements of the airline's operations, including aircraft maintenance routing, crew schedules, and passenger itineraries. The objective of aircraft maintenance routing is to develop a *maintenance feasible* route for each aircraft, that is, a sequence of flight legs visiting maintenance stations at regular intervals to ensure adequate opportunities for maintenance checks. Similarly, the goal of crew scheduling is to develop for each crew member, a sequence of flight legs to be operated by that crew. This sequence must satisfy numerous work rules defined by government agencies and collective bargaining agreements between labor and the airline. If maintenance routes and crew schedules are developed prior to the re-optimization point, changes in flight departure and arrival times can alter aircraft turn times and crew connection times, rendering the aircraft routes and crew schedules infeasible. Moreover, changes in aircraft fleeting, while having no impact on crew schedules, can require changes in the routes of the swapped aircraft to ensure that all affected aircraft have sufficient opportunities for maintenance. Although we do not conduct experiments to quantify the effects of dynamic scheduling on crews and maintenance routing, Berge and Hopperstad (1993) report that determining revised, feasible aircraft maintenance routing solutions after schedule changes is typically not difficult. If altering aircraft routing and crew scheduling is too onerous or expensive, the re-optimization point can be moved earlier to precede the release of the aircraft and crew plans. It is worth noting that airlines, who desire the ability to alter plans closer to the date of departure, are now moving toward delaying the release of aircraft routing and crew pairing plans.

Notification of schedule changes is necessary for nonstop passengers whose flights are moved earlier and for connecting passengers whose flights departing the origin city are moved earlier. This can be accomplished in several different ways, for example, automated emails, automated text messages, and automated phone calls over conventional telephone networks or using voice-over-internet protocol (VoIP).

2.9.2 Applicability of Dynamic Scheduling to Other Airlines

In this chapter, we conduct experiments for a major U.S. hub-and-spoke carrier and demonstrate that dynamic airline scheduling effectively improves profitability. In this section, we explore the question of whether or not dynamic scheduling can achieve similar improvements when applied to other airlines. We begin by identifying several conditions that are critical to the success of dynamic airline scheduling, and then show that these conditions are easily satisfied by major airlines in the U.S.. Hence, we conclude that that similar results are obtainable for other major airlines.

The underlying driver of the need for dynamic scheduling is the existence of imbalances in load factors across flight legs. If load factors are high on some flight legs and low on others, dynamic airline scheduling is an effective tool for making schedule adjustments that move excess capacity in some markets to markets where there is a shortage of capacity. If most of the flight legs have low load factors, there is no need to carry out dynamic scheduling. Similarly, if all flight legs are highly capacitated, dynamic airline scheduling provides little if any benefit. Recall that in Figure 1-1, we show the histogram of load factors for one major U.S. airline to illustrate imbalanced load factors over an airline's flight schedule. In Figure 2-29 we provide load factor histograms for two other major U.S. airlines, further demonstrating the load factor imbalance phenomenon. The bars in the figures correspond to the percentage of flight legs with associated load factors falling in the selected interval (y-axis on the left), and the lines correspond to the cumulative percentages (y-axis on the right). In the figure on the top, we observe that about 40% of the flight legs have load factors below 0.60, while at the same time about 20% of the flight legs have load factors of at least 0.90. In the figure on the bottom, the load factors on more than 20% of the flights are below 0.60, while the load factors of 30% of the flights are above 0.90. From these examples, we conclude that dynamic airline scheduling has the potential to make a difference in other airlines.

Next, we discuss factors that can affect the applicability of the two dynamic scheduling mechanisms. The prerequisite for performing flight leg re-fleeting is to

have several fleet families containing aircraft with different seating capacity, and at least two of these aircraft on the ground at the same time. This condition is easily satisfied by most airlines in the U.S. because they typically operate at least one fleet family at their hub airports.

Although flight leg re-timing can have the added benefit of creating more re-fleeting opportunities, its primary contribution is to adjust the number of seats to match demand. In order to apply flight leg re-timing effectively, it is necessary to have:

A de-peaked busy hub For maximal effectiveness, flight leg re-timing requires a relatively dense flight schedule with a de-peaked hub. A hub with a perfectly banked schedule will not have flight leg re-timing opportunities that lead to new connecting itineraries. Moreover, flight leg re-timing is ineffective in de-peaked hubs at which the number of arriving and departing flight legs is small. In our experiments, the number of departures and arrivals at hub airports is limited to five per 10-minute interval. Flint (2002) reports that when American Airlines de-peaked its hub in Chicago (ORD), the number of departures and arrivals at hubs were limited to one per minute, that is ten per 10-minute interval. In general, major U.S. airlines have 300 to 1000 flights inbound and outbound at their busiest hub. These hubs remain *busy* after de-peaking and provide sufficient opportunity for flight leg re-timing to create new connecting itineraries.

A large percentage of connecting passengers at the hub Having a large number of passengers connecting at the hub is vital to the success of flight leg re-timing. In Table 2.20, we report statistics related to connecting passengers on domestic itineraries at the 30 largest airports in the U.S.. For a given airport, we let x be the number of passengers with itineraries originating from or terminating at this airport and y be the number of connecting passengers at this airport. In the second column, we report the percentage of connecting passengers at the airport, that is, $y/(x+y)$; and in the third column, we report the

percentage of seats sold to connecting passengers for all flight legs to and from the airport, that is, $2y/(x + 2y)$. Connecting passengers account for a large percentage of passengers at many of these airports: 15 airports have more than 50% of all passengers connecting at the airport, and 27 airports have more than 50% of all sold seats on flight legs to and from the airport occupied by connecting passengers. In Table 2.21, we report the same information for each major U.S. airline at its hubs or major airports. All airlines except Alaska Airlines and Southwest Airlines have at least one or more airports at which a substantial percentage of the passengers are connecting.

Airport code	**Pct. of connecting passengers**	**Pct. of seats sold to connecting passengers**	**Rank by Size**
MEM	73.7%	84.9%	27
CLT	66.1%	79.6%	15
CVG	64.7%	78.6%	22
ORD	63.0%	77.3%	1
IAD	62.1%	76.6%	16
DEN	59.4%	74.6%	5
LAS	58.1%	73.5%	14
PHX	57.8%	73.3%	7
ATL	56.8%	72.5%	2
MCO	56.2%	71.9%	19
DFW	55.9%	71.7%	4
DTW	54.9%	70.9%	10
IAH	54.8%	70.8%	6
MSP	53.9%	70.0%	8
SLC	51.2%	67.8%	20
DCA	44.8%	61.9%	21
PHL	44.0%	61.1%	17
LAX	43.3%	60.4%	3
JFK	41.4%	58.6%	23
TPA	40.6%	57.7%	29
EWR	40.5%	57.7%	9
SEA	38.3%	55.4%	11
SNA	38.3%	55.4%	30
HNL	36.9%	53.9%	18
FLL	35.2%	52.0%	25
LGA	33.4%	50.0%	13
PDX	33.3%	50.0%	28
SAN	32.4%	48.9%	24
SFO	26.8%	42.3%	12
BOS	16.7%	28.6%	26

Source: U.S. D.O.T. DB1BMarket database, fourth quarter, 2005.

Table 2.20: Connecting passenger statistics on domestic itineraries at the 30 largest U.S. airport based on number of domestic passengers enplaned

2.10 Summary

Dynamic airline scheduling provides a way to manage capacity dynamically to match fluctuating passenger demand. In this research, we introduce a new mechanism for dynamic scheduling, namely, flight re-timing. We develop an optimization model that combines both flight re-fleeting and flight re-timing. We conduct experiments that demonstrate the significant benefit of dynamic scheduling and the synergy between flight re-fleeting and re-timing. The estimated profit increase is $36 million annually when perfect forecast is used in re-optimization. When imperfect forecast is used, the estimated profit increase remains significant at $18 million annually. Sensitivity analysis is conducted on the number of re-timed flights per day. It shows that the full benefit is achieved when a moderate number of flights (10% of the total number of flights) are re-timed.

The performance of flight re-timing and flight re-fleeting are compared when applied alone. It is shown that flight re-timing constitutes a larger portion of the total benefit than flight re-fleeting. We also find that flight re-timing is less sensitive to forecast quality. When forecasts are imperfect, flight re-timing realizes 85% of the benefit achieved when forecasts are perfect.

A side study is performed assuming a weekly schedule is in place to account for day-of-week demand variations. Dynamic scheduling continues to improve schedule profitability when a weekly schedule is applied. On average, the estimated profit increase is $32 million when forecasts are perfect and $15 million when forecasts are imperfect.

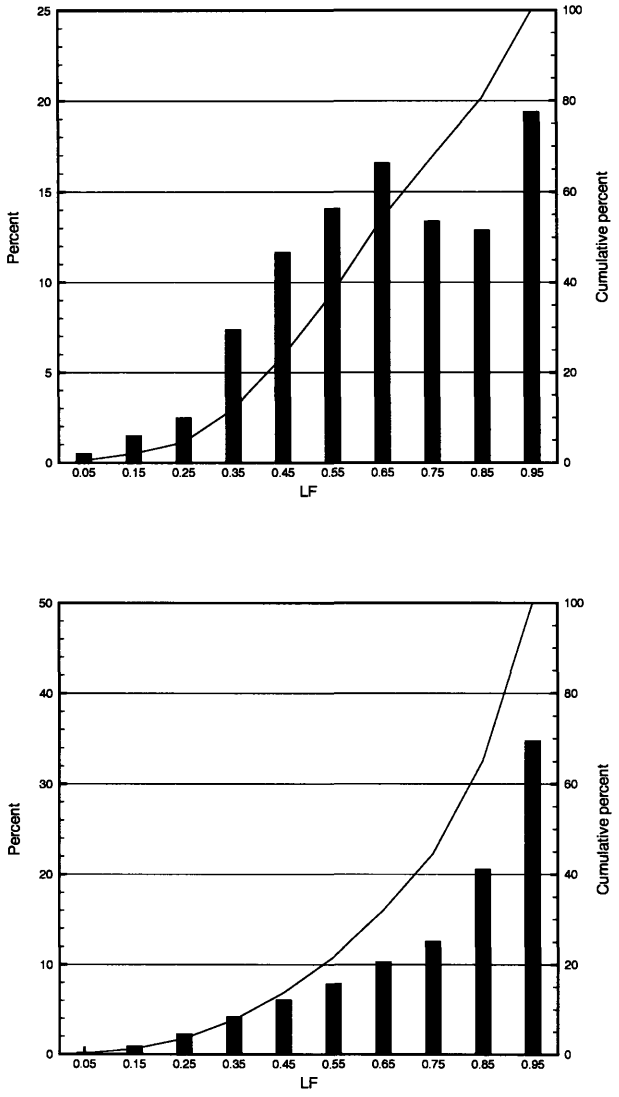


Figure 2-29: Histograms and cumulative percentages of flight load factors for two major US hub-and-spoke carriers. The figure on the top is based on 2004 data and the one on the bottom is based on 2003 data. Source: PODS Consortium, ICAT, MIT (2006)

Carrier	Airport code	Pct. of passengers changing Planes	Pct. of seats sold to connecting passengers
American	DFW	47.5%	64.4%
	LAX	12.7%	22.6%
	MIA	11.9%	21.3%
	ORD	33.0%	49.6%
	SJU	8.6%	15.9%
	STL	23.3%	37.8%
Alaska	ANC	16.4%	28.1%
	PDX	13.9%	24.4%
	SEA	19.7%	32.9%
Continental	CLE	21.7%	35.6%
	EWR	9.0%	16.5%
	IAH	40.3%	57.5%
Delta	ATL	52.3%	68.7%
	CVG	61.9%	76.5%
	SLC	46.6%	63.6%
AirTran	ATL	40.1%	57.2%
Northwest	DTW	44.0%	61.1%
	MEM	65.3%	79.0%
	MSP	44.8%	61.9%
United	DEN	41.2%	58.4%
	IAD	25.7%	40.9%
	LAX	21.5%	35.4%
	ORD	39.4%	56.5%
	SFO	22.2%	36.3%
US Airways (America West)	CLT	65.3%	79.0%
	PHL	30.7%	47.0%
	PHX	41.4%	58.5%
Southwest	BWI	16.6%	28.5%
	HOU	18.4%	31.1%
	LAS	12.8%	22.7%
	MDW	17.4%	29.6%
	PHX	15.7%	27.2%

Source: U.S. D.O.T. DB1BMarket database, fourth quarter, 2005.

Table 2.21: Connecting passenger statistics on domestic itineraries for major U.S. airlines at hub or major airports

Chapter 3

Robust Airline Schedule De-Peaking

3.1 Introduction

In the previous chapter, we demonstrate the benefits of dynamic airline scheduling, in which the set of possible dynamic scheduling decisions (re-timing and re-fleeting) are largely defined by the original schedule. Because some schedules provide limited opportunities to re-time or re-fleet, an important and challenging research question is how to design the flight schedule to maximize future dynamic scheduling capabilities. In this chapter, we address this question in the context of de-peaking an existing peaked hub-and-spoke schedule. We design *robust* de-peaking models to produce schedules that are more responsive, that is, more robust, to variations in demand when dynamic scheduling is applied.

Robust schedule planning has been increasingly applied in the airline industry. The objectives are typically either to lessen the impact of disruptions on the schedule; or facilitate repair of the schedule when disruptions occur. Examples of recent research to produce robust plans for the airline industry include: robust crew scheduling (Shebalov and Klabjan, 2004; Schaefer et al., 2005; Yen and Birge, 2006), robust fleet assignment (Rosenberger et al., 2004; Smith, 2004), robust aircraft routing (Ageeva, 2000), robust aircraft routing and flight schedule re-timing to reduce

passenger disruptions (Lan et al., 2006), robust passenger re-routing during disruption management (Karow, 2003), and a methodology to design a layered schedule that isolates the effects of disruptions (Kang and Clarke, 2003).

Research aimed at improving schedule robustness in light of demand variability is limited. Focusing on models and solution approaches for the dynamic re-fleeting problem, Sherali et al. (2005) propose to study the interaction between an original schedule and subsequent dynamic re-fleeting decisions. To our knowledge, the results of this study have not been reported to date. In a loosely related paper, Listes and Dekker (2005) argue that suitably distributed aircraft capacity is critical to the successful implementation of dynamic re-fleeting procedures, and propose a scenario aggregation-based approach to determine an optimal fleet composition that facilitates dynamic re-fleeting. They formulate the problem as a mixed-integer programming problem. To solve the linear programming relaxation, individual scenario problems are solved and aggregated to form a solution to the overall problem. Integer fleet composition solutions are then generated by applying rounding heuristics to the solution of the linear programming relaxation. Listes and Dekker conduct case studies to compare the effects of complete fleet re-assignment and re-fleeting only within families. They showed that by having a robust fleet composition, the load factors potentially increase up to 2.6% in the case of complete fleet re-assignment and up to 1.7% in the case of re-fleeting only within families.

In the robust de-peaking approach we propose, we begin by solving a *basic* de-peaking model in which a profit-maximizing flight schedule is obtained. Then, we solve our robust de-peaking model to generate a modified flight schedule in which the number of new itineraries that can be created through slight flight leg re-timings is maximized, while the resulting reduction in schedule profitability is limited to a prescribed amount. We present several alternate formulations of the robust de-peaking model, each with different mathematical and computational properties. Because the size of our robust de-peaking models grow substantially compared to a basic de-peaking model, approximate models are introduced aimed at improving computational tractability. We further explore the property of optimal solutions to identify

variables that can be excluded from the model without compromising solution quality. Our dynamic scheduling approach is applied to the solutions generated by the robust de-peaking model and the basic de-peaking model, and the results are reported.

We organize the remainder of this chapter as follows. In Section 3.2, we define the schedule de-peaking problem and briefly review related literature. In Section 3.3, we present a basic de-peaking model. We introduce the robust de-peaking model, the restricted robust de-peaking model and the associated solution algorithms in Sections 3.4, 3.5, and 3.6, respectively. Sections 3.7 reports our computational experiences and Section 3.8 provides an evaluation of our robust solutions. Finally, we provide concluding remarks in Section 3.9.

3.2 The Schedule De-Peaking Problem and Related Literature

Given an existing peaked schedule in a hub-and-spoke network, the schedule de-peaking problem is to determine when to schedule each of the flight legs in the peaked schedule and what fleet type to assign to each of these flight legs so as to maximize profits while satisfying flight cover constraints, aircraft balance constraints, aircraft count constraints, and de-peaking constraints (constraints that limit the number of departures and arrivals per unit of time at hubs). Singh (2006) develops an automated de-peaking tool in which three copies of each flight leg are considered; one at 15 minutes prior to the originally scheduled departure time in the peaked schedule, one at the scheduled departure time, and one at 15 minutes after the scheduled departure time. Incremental connection variables are introduced to evaluate changes in revenue. When a valid connection in the peaked schedule becomes invalid, it has negative revenue impact; when an invalid connection in the peaked schedule becomes valid, it has positive revenue impact. The model maximizes the revenue impact of incremental connection variables while satisfying cover constraints, balance constraints, count constraints, de-peaking constraints, and so on. A separate model is used repeatedly

to determine the revenue impact of incremental connection variables.

While not specifically addressing the problem of designing flight schedules for de-peaked hub operations, existing literature on airline schedule design nonetheless provides valuable insights to the modeling of the schedule design problem involving hub de-peaking. A sampling of schedule design research includes Simpson (1966), Simpson (1969), Chan (1972), Soumis et al. (1980), Nikulainen and Oy (1992), Dobson and Lederer (1993), Berge (1994), Marsten et al. (1996), Rexing et al. (2000), Yan and Tseng (2002), and Lohatepanont and Barnhart (2004).

Marsten et al. (1996) present a framework for *incremental* schedule design. Given a schedule, demands are estimated using a *schedule evaluation* model. Then, the fleet assignment problem is solved over the given schedule with the estimated demands. Potential combinations of flight leg additions and deletions to the current schedule are enumerated and evaluated to identify the best set of modifications to the flight schedule.

Rexing et al. (2000) recognize that adjusting scheduled flight departure times can result in a modified schedule providing improved flight connection opportunities and allowing a more cost-effective fleet assignment. Rexing et al. develop a model (referred to as *FAM with time windows*) and solution approach to assign aircraft types to flight legs and, simultaneously, schedule flight departure times. They report that the resulting schedule and fleet assignment translates to savings on the order of 30 to 50 million dollars annually.

Yan and Tseng (2002) develop a model and a solution algorithm to help carriers simultaneously solve for fleet assignment and appropriate timetable. The model is formulated as an integer multi-commodity network flow problem. They develop a solution algorithm based on Lagrangian relaxation and a sub-gradient method .

Lohatepanont and Barnhart (2004) build an integrated schedule design model based on the Itinerary-Based Fleet Assignment Model of Barnhart et al. (2002). A master flight list is generated containing all mandatory and optional flight legs, and demand correction terms are introduced to model demand and supply interactions. The integrated model is solved by an algorithm based on row and column generation,

and they estimate potential savings associated with the advanced approach of up to $350 million annually.

3.3 Basic De-Peaking Model

The terminologies and underlying networks used in this chapter are drawn from those described in the previous chapter. Readers should refer to Section 2.5.1, Section 2.5.2, and Section 2.5.3 of Chapter 2 for a detailed description.

The basic de-peaking model is built upon two models, namely: FAM with time windows (Rexing et al., 2000) and Itinerary-Based FAM (Barnhart et al., 2002). Input to the model includes the set of flight legs in the peaked schedule, each of which must be assigned to exactly one aircraft type. The departure time of the flight leg in the de-peaked schedule can differ from that of the peaked schedule, and these sets of decisions are modeled by creating a copy of each flight leg for every allowable departure time, and a set of constraints ensuring that exactly one copy is assigned one aircraft type. We introduce the following notation before detailing our basic de-peaking model.

Data and Parameters

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

$C(l)$: set of flight copies for flight leg $l \in L$.

$\langle l, k \rangle$: copy $k \in C(l)$ of flight leg $l \in L$.

Π : set of fleet types indexed by π.

S : set of cities.

G^{π} : set of ground arcs in fleet $\pi \in \Pi$'s network.

M : set of markets.

D_m : demand in market $m \in M$.

$fare_m$: average fare for demand in market $m \in M$.

$R(m)$: set of itineraries serving market $m \in M$.

$$\delta_{mr}^{lk} = \begin{cases} 1, & \text{if itinerary } r \in R(m) \text{ in market } m \in M \text{ traverses } \langle l, k \rangle; \\ 0, & \text{otherwise.} \end{cases}$$

$c_{lk\pi}$: cost to fly $\langle l, k \rangle$ with aircraft type $\pi \in \Pi$.

c_π : fixed cost of one aircraft of type $\pi \in \Pi$.

N^π : set of nodes in flight network of fleet type $\pi \in \Pi$.

n^π : number of aircraft available of fleet type $\pi \in \Pi$.

T : set of time intervals at the hub, indexed by t.

MAX^{at} : maximum number of aircraft arrivals at the hub in interval $t \in T$.

MAX^{dt} : maximum number of aircraft departures from the hub in interval $t \in T$.

$$\overline{\alpha}^i_{lk\pi} = \begin{cases} 1, & \text{if } \langle l,k \rangle \text{ in fleet } \pi\text{'s network originates at node } i \in N^\pi; \\ -1, & \text{if } \langle l,k \rangle \text{ in fleet } \pi\text{'s network terminates at node } i \in N^\pi; \\ 0, & \text{otherwise.} \end{cases}$$

$$\widehat{\alpha}^i_{g\pi} = \begin{cases} 1, & \text{if ground arc } g \in G^\pi \text{ originates at node } i \in N^\pi; \\ -1, & \text{if ground arc } g \in G^\pi \text{ terminates at node } i \in N^\pi; \\ 0, & \text{otherwise.} \end{cases}$$

$$\overline{\beta}_{lk\pi} = \begin{cases} 1, & \text{if } \langle l,k \rangle \text{ in fleet } \pi\text{'s network crosses the count line;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\widehat{\beta}_{g\pi} = \begin{cases} 1, & \text{if ground arc } g \in G^\pi \text{ crosses the count line;} \\ 0, & \text{otherwise.} \end{cases}$$

$$\gamma^{at}_{lk} = \begin{cases} 1, & \text{if } \langle l,k \rangle \text{ arrives at the hub during interval } t \in T; \\ 0, & \text{otherwise.} \end{cases}$$

$$\gamma^{dt}_{lk} = \begin{cases} 1, & \text{if } \langle l,k \rangle \text{ departs from the hub during interval } t \in T; \\ 0, & \text{otherwise.} \end{cases}$$

Decision Variables

x_{mr} : number of passengers assigned to itinerary $r \in R(m)$ in market $m \in M$.

$$f_{lk\pi} = \begin{cases} 1, & \text{fleet } \pi \in \Pi \text{ is used to fly flight copy } \langle l,k \rangle; \\ 0, & \text{otherwise.} \end{cases}$$

$y_{g\pi}$: number of aircraft of fleet type π traversing ground arc $g \in G^\pi$.

z_π : number of aircraft used of fleet type $\pi \in \Pi$.

The basic de-peaking model is presented as follows:

maximize

$$p^* = \sum_{m \in M} \sum_{r \in R(m)} x_{mr} fare_m - \sum_{l \in L} \sum_{k \in C(l)} \sum_{\pi \in \Pi} c_{lk\pi} f_{lk\pi} - \sum_{\pi \in \Pi} z_\pi c_\pi$$

subject to

$$\sum_{k \in C(l)} \sum_{\pi \in \Pi} f_{lk\pi} = 1, \forall l \in L \tag{3.1}$$

$$\sum_{l \in L} \sum_{k \in C(l)} f_{lk\pi} \overline{\alpha}^i_{lk\pi} + \sum_{g \in G^\pi} y_{g\pi} \widehat{\alpha}^i_{g\pi} = 0, \forall i \in N^\pi, \pi \in \Pi \tag{3.2}$$

$$\sum_{l \in L} \sum_{k \in C(l)} f_{lk\pi} \overline{\beta}_{lk\pi} + \sum_{g \in G^\pi} y_{g\pi} \widehat{\beta}_{g\pi} = z_\pi, \forall \pi \in \Pi \tag{3.3}$$

$$z_\pi \leq n^\pi, \forall \pi \in \Pi \tag{3.4}$$

$$\sum_{l \in L} \sum_{k \in C(l)} \gamma^{at}_{lk} \sum_{\pi \in \Pi} f_{lk\pi} \leq MAX^{at}, \forall t \in T \tag{3.5}$$

$$\sum_{l \in L} \sum_{k \in C(l)} \gamma^{dt}_{lk} \sum_{\pi \in \Pi} f_{lk\pi} \leq MAX^{dt}, \forall t \in T \tag{3.6}$$

$$\sum_{r \in R(m)} x_{mr} \leq D_m, \forall m \in M \tag{3.7}$$

$$\sum_{m \in M} \sum_{r \in R(m)} \delta^{lk}_{mr} x_{mr} \leq \sum_{\pi \in \Pi} f_{lk\pi} CAP_\pi, \forall l \in L, k \in C(l) \tag{3.8}$$

$$f_{lk\pi} \in \{0, 1\}, \forall l \in L, k \in C(l), \pi \in \Pi \tag{3.9}$$

$$y_{g\pi} \geq 0, \forall g \in G^\pi, \pi \in \Pi \tag{3.10}$$

$$x_{mr} \geq 0, \forall m \in M, r \in R(m) \tag{3.11}$$

$$z_\pi \geq 0, \forall \pi \in \Pi \tag{3.12}$$

The objective function is to maximize revenue less operating cost and fixed costs. Constraints (3.1) ensure that each flight leg is covered exactly once, while Constraints (3.2) enforce conservation of flow for each type of aircraft. Constraints (3.3) and (3.4) count the number of aircraft of each fleet used and limit that number to the number available. Constraints (3.5) and (3.6), the *de-peaking* constraints, limit the number of flight departures and arrivals per unit time at the hub. Constraints (3.7) are passenger

flow constraints restricting the number of passengers transported in each market to the value of that market's unconstrained demand. Constraints (3.8) are the capacity constraints, and Constraints (3.9) through (3.12) define the range of possible values for variables in this model.

3.4 Robust De-Peaking Model

Dynamic airline scheduling includes two elements, namely, flight leg re-fleeting and flight leg re-timing. While de-peaking the schedule does reschedule flight leg departure and arrival times, the set of flight legs in the network before and after de-peaking remains unchanged. Hence, the hub-and-spoke network structure is maintained, thereby allowing numerous aircraft swapping opportunities at the hub (Berge and Hopperstad, 1993). Given this, we focus our efforts on de-peaking to create flight re-timing opportunities that potentially can be executed during the dynamic scheduling stage.

In a hub-and-spoke network, an inbound flight leg f arriving at the hub at time t can connect to any outbound flight leg departing between $t+MinCT$ and $t+MaxCT$ to form a *feasible connecting itinerary*, where $MinCT$ represents the minimum connection time and $MaxCT$ denotes the maximum connection time at the hub. Any outbound flight departing earlier than $t + MinCT$ or later than $t + MaxCT$ forms an *infeasible connecting itinerary* with inbound flight leg f and contributes nothing toward revenue. The basic de-peaking model presented in Section 3.3 and virtually all schedule planning models in the literature maximize revenue less cost: the objective is to determine the optimal set of nonstop and feasible connecting itineraries to be included in the schedule. For infeasible connecting itineraries, these models are indifferent to the extent of their *infeasibility*, that is, the amount of time these infeasible connecting itineraries violate the connection time requirement. We define *slightly infeasible connecting itineraries* as those connecting itineraries whose connection times violate the minimum or maximum allowable connection times by a small margin, such as by 15 or fewer minutes. Slightly infeasible connecting itineraries have the property that they can be transformed into *feasible connecting itineraries* when

flights are re-timed by the dynamic scheduling process. Motivated by such an observation, our robust de-peaking model works to maximize both the number of feasible and slightly infeasible connecting itineraries in the de-peaked schedule. In doing so, we enhance the potential for dynamic scheduling procedures to increase capacity in markets experiencing greater than expected demand.

We begin by enumerating all *potentially connecting itineraries*, that is, all itineraries whose connection times are *feasible* if allowable flight leg re-timings can be applied. Note that this set includes the set of connecting itineraries that are feasible without any re-timings. The set of potentially connecting itineraries can thus be defined as all itineraries with connection times between T_1 and T_2, where $T_1 < MinCT$ and $T_2 \geq MaxCT$, and the precise values of T_1 and T_2 are a function of the maximum allowable amount of time that flight legs can be shifted. Let PC be the set of all potentially connecting itineraries. We assign a binary *connection variable* h_p with weight w_p to each potentially connecting itinerary $p \in PC$, where w_p measures the revenue that can be attained by p. Binary variable h_p takes value one if both flight leg copies forming potentially connecting itinerary p are selected in the solution; otherwise, $h_p = 0$. The objective, then, is to maximize the weighted sum of these connection variables values, that is,

$$\text{maximize} \sum_{p \in PC} w_p h_p.$$

We now introduce the following additional notation used in our robust de-peaking model.

(l_1, l_2) : ordered pair of inbound flight leg $l_1 \in L$ and outbound flight leg $l_2 \in L$ at the hub.

$m_{(l_1,l_2)}$: the market servable by the potentially connecting itinerary formed by flight leg l_1 and l_2.

$(\langle l_1, k_1 \rangle, \langle l_2, k_2 \rangle)$: a potentially connecting itinerary formed by $\langle l_1, k_1 \rangle$ and $\langle l_2, k_2 \rangle$, also referred to as p.

$CT(\langle l_1, k_1 \rangle, \langle l_2, k_2 \rangle)$: the connection time between $\langle l_1, k_1 \rangle$ and $\langle l_2, k_2 \rangle$.

C : set of flight leg pairs (l_1, l_2) such that $T_1 \leq CT(\langle l_1, k_1\rangle, \langle l_2, k_2\rangle) \leq T_2$ for some $k_1 \in C(l_1)$, $k_2 \in C(l_2)$.

$C(l_1, l_2)$: $\{(\langle l_1, k_1\rangle, \langle l_2, k_2\rangle) | T_1 \leq CT(\langle l_1, k_1\rangle, \langle l_2, k_2\rangle) \leq T_2\}$, where $(l_1, l_2) \in C$. By definition, we have $PC = \{p | p \in C(l_1, l_2), \forall (l_1, l_2) \in C\}$.

$S(p)$: binary set $\{\langle l_1, k_1^p\rangle, \langle l_2, k_2^p\rangle\}$, where $p \in C(l_1, l_2)$ and p corresponds to $(\langle l_1, k_1^p\rangle, \langle l_2, k_2^p\rangle)$.

w_p : revenue associated with connection variable $p \in C(l_1, l_2)$. Its value is computed as follows:

$$w_p = w_{m_{(l_1,l_2)}} = \begin{cases} log_e(D_{m_{(l_1,l_2)}} fare_{m_{(l_1,l_2)}}) + 1, & \text{if } D_{m_{(l_1,l_2)}} fare_{m_{(l_1,l_2)}} \geq 1; \\ D_{m(l_1,l_2)} fare_{m_{(l_1,l_2)}}, & \text{otherwise.} \end{cases}$$

$$\zeta_p^{lk} = \begin{cases} 1, & \text{if connection } p \text{ traverses } \langle l, k\rangle; \\ 0, & \text{otherwise.} \end{cases}$$

For a potentially connecting itinerary $p \in C(l_1, l_2)$, its weight w_p is derived from $D_{m_{(l_1,l_2)}} fare_{m_{(l_1,l_2)}}$, the potential revenue in market $m_{(l_1,l_2)}$. A natural logarithm is taken to decrease the difference between weights for h_p variables in order to create as many potentially connecting itineraries as possible in the schedule. If some potentially connecting itineraries have unusually large weights, the optimization model will then try to create those itineraries at the expense of eliminating many other potentially connecting itineraries.

Lemma 3.1 *By the definitions of h_p variables and $f_{lk\pi}$ variables, we have the following results:*

1. $\sum_{\pi \in \Pi} f_{lk\pi} \in \{0, 1\}, \forall l \in L, k \in C(l)$. *Exactly one $\sum_{\pi \in \Pi} f_{lk\pi}$ $(k \in C(l))$ takes value one in any feasible solution and the rest take value zero.*

2. $\sum_{p \in C(l_1,l_2)} h_p \in \{0, 1\}, \forall (l_1, l_2) \in C$.

Proof: Because $f_{lk\pi} \in \{0, 1\}$ $(\forall k \in C(l), l \in L, \pi \in \Pi)$ and $\sum_{k \in C(l)} \sum_{\pi \in \Pi} f_{lk\pi} = 1$ $(\forall l \in L)$, we have $\sum_{\pi \in \Pi} f_{lk\pi} \in \{0, 1\}$. We also know for $l \in L$, there exists exactly

one $k_l \in C(l)$, such that, $\sum_{\pi\in\Pi} f_{lk_l\pi} = 1$ and $\sum_{\pi\in\Pi} f_{lk\pi} = 0$ $(\forall k \in C(l), k \neq k_l)$. As a result, for all $p \in C(l_1, l_2)$, at most one of h_p takes value one and the rest take value zero. ■

3.4.1 Formulation 1

Now we develop the constraints for the robust de-peaking model. In Figure 3-1 we show an inbound flight leg l_1 and an outbound flight leg l_2 each with three copies. Three connection variables are illustrated, that is, $(\langle l_1, a\rangle, \langle l_2, d\rangle)$, $(\langle l_1, a\rangle, \langle l_2, e\rangle)$, $(\langle l_1, a\rangle, \langle l_2, f\rangle)$ and are indexed as h_1, h_2, and h_3. For each flight leg copy $\langle l, k\rangle$, if the term $\sum_{\pi\in\Pi} f_{lk\pi}$ equals one, this flight leg copy is selected in the solution; otherwise, it is not.

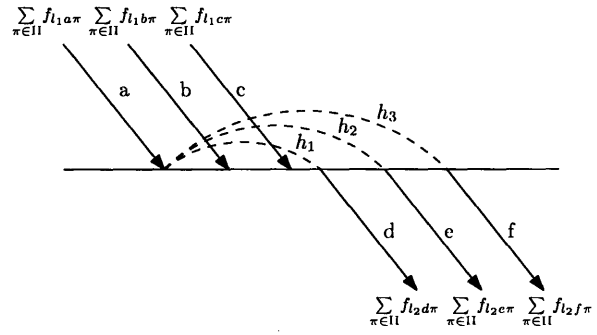


Figure 3-1: Illustration of connection variables

If h_1 takes value one in a feasible integer solution, $\sum_{\pi\in\Pi} f_{l_1a\pi}$ and $\sum_{\pi\in\Pi} f_{l_2d\pi}$ must both take value one. Moreover, if either $\sum_{\pi\in\Pi} f_{l_1a\pi}$ or $\sum_{\pi\in\Pi} f_{l_2d\pi}$ take value zero, h_1 can only take value zero. Such a relationship can be modeled by the following constraint:

$$2h_1 \leq \sum_{\pi\in\Pi} f_{l_1a\pi} + \sum_{\pi\in\Pi} f_{l_2d\pi}.$$

Following similar logic, we can formulate the constraints for the robust de-peaking model as follows:

$$\sum_{k\in C(l)} \sum_{\pi\in\Pi} f_{lk\pi} = 1, \forall l \in L \tag{3.13}$$

$$\sum_{l\in L} \sum_{k\in C(l)} f_{lk\pi}\overline{\alpha}^i_{lk\pi} + \sum_{g\in G^\pi} y_{g\pi}\widehat{\alpha}^i_{g\pi} = 0, \forall i \in N^\pi, \pi \in \Pi \tag{3.14}$$

$$\sum_{l\in L}\sum_{k\in C(l)} f_{lk\pi}\overline{\beta}_{lk\pi} + \sum_{g\in G^{\pi}} y_{g\pi}\widehat{\beta}_{g\pi} = z_{\pi}, \forall \pi \in \Pi \tag{3.15}$$

$$z_{\pi} \leq n^{\pi}, \forall \pi \in \Pi \tag{3.16}$$

$$\sum_{l\in L}\sum_{k\in C(l)} \gamma_{lk}^{at} \sum_{\pi\in\Pi} f_{lk\pi} \leq MAX^{at}, \forall t \in T \tag{3.17}$$

$$\sum_{l\in L}\sum_{k\in C(l)} \gamma_{lk}^{dt} \sum_{\pi\in\Pi} f_{lk\pi} \leq MAX^{dt}, \forall t \in T \tag{3.18}$$

$$\sum_{r\in R(m)} x_{mr} \leq D_m, \forall m \in M \tag{3.19}$$

$$\sum_{m\in M}\sum_{r\in R(m)} \delta_{mr}^{lk} x_{mr} \leq \sum_{\pi\in\Pi} f_{lk\pi} CAP_{\pi}, \forall l \in L, k \in C(l) \tag{3.20}$$

$$\sum_{m\in M}\sum_{r\in R(m)} x_{mr} fare_{mr} - \sum_{l\in L}\sum_{k\in C(l)}\sum_{\pi\in\Pi} c_{lk\pi} f_{lk\pi} - \sum_{\pi\in\Pi} z_{\pi} c_{\pi} \geq p^{*} \tag{3.21}$$

$$2h_p \leq \sum_{\langle l,k\rangle\in S(p)}\sum_{\pi\in\Pi} f_{lk\pi}, \forall p \in PC \tag{3.22}$$

$$f_{lk\pi} \in \{0,1\}, \forall l \in L, k \in C(l), \pi \in \Pi \tag{3.23}$$

$$y_{g\pi} \geq 0, \forall g \in G, \pi \in \Pi \tag{3.24}$$

$$x_{mr} \geq 0, \forall m \in M, r \in R(m) \tag{3.25}$$

$$z_{\pi} \geq 0, \forall \pi \in \Pi \tag{3.26}$$

$$h_p \in \{0,1\}, \forall p \in PC \tag{3.27}$$

Constraints (3.13) through (3.20) are the same as those used in the basic de-peaking model. Constraint (3.21) sets the lower bound on the profit attainable by the robust de-peaking solution to the profit achieved in the basic de-peaking solution. This constraint allows a trade-off between robustness and planned profitability. In this case, the model selects among the set of profit maximizing solutions the one that is most robust. Constraints (3.22) guarantee that if connection variable h_p equals one in the solution, the two flight leg copies that form p are selected in the solution. Constraints (3.23) through (3.27) specify the possible values of decision variables.

We refer to this formulation as Formulation 1. Next, we show a proposition on the optimal objective value of the Linear Programming (LP) relaxation of Formulation 1.

Proposition 3.1 *In the LP relaxation of Formulation 1, if variable $f_{lk\pi}$ takes value $\overline{f}_{lk\pi}$ $(\forall l \in L, k \in C(l), \pi \in \Pi)$, the optimal objective function of the LP relaxation is*

$$z_1 = \frac{1}{2} \sum_{p \in PC} w_p \left(\sum_{\langle l,k \rangle \in S(p)} \sum_{\pi \in \Pi} \overline{f}_{lk\pi} \right).$$

Proof: An upper bound on the objective value can be obtained by multiplying each constraint in Constraints (3.22) by its corresponding w_p value, and then summing over all $p \in PC$. This maximal value is attained when

$$h_p = \frac{1}{2} \sum_{\langle l,k \rangle \in S(p)} \sum_{\pi \in \Pi} \overline{f}_{lk\pi}, \forall p \in PC.$$

■

3.4.2 Formulation 2

Constraints (3.22) in Formulation 1 are used to model the relationship between h_p variables and $f_{lk\pi}$ variables. In Sections 3.4.2 through 3.4.4, we present alternative representations of this relationship. These various representations can be used to create different robust de-peaking formulations, all with the same set of optimal integer solutions but with different mathematical and computational properties.

Constraints (3.22) can be replaced by Constraints (3.28) to obtain Formulation 2.

$$h_p \leq \sum_{\pi \in \Pi} f_{lk\pi}, \forall \langle l,k \rangle \in S(p), \forall p \in PC. \tag{3.28}$$

Proposition 3.2 *Formulation 2 is equivalent to Formulation 1.*

Proof: Because h_p variables and $\sum_{\pi \in \Pi} f_{lk\pi}$ are binary (Lemma 3.1), we have the following results for any $p \in PC$:

$$\begin{aligned} & 2h_p \leq \sum_{\langle l,k \rangle \in S(p)} \sum_{\pi \in \Pi} f_{lk\pi} \\ \iff \quad & h_p \leq \min_{(l,k) \in S(p)} \left\{ \sum_{\pi \in \Pi} f_{lk\pi} \right\} \\ \iff \quad & h_p \leq \sum_{\pi \in \Pi} f_{lk\pi}, \forall \langle l,k \rangle \in S(p). \end{aligned}$$

Therefore, Constraints (3.22) and Constraints (3.28) are equivalent. Because all other constraints and the objective function are the same in both formulations, Formulation 2 is equivalent to Formulation 1. ■

An advantage of Constraints (3.28) is that the integrality constraints on h_p variables (Constraints 3.27) can be relaxed as is shown in the following proposition. For a mixed integer programming problem, even a relatively small reduction in the number of integer variables can significantly improve the tractability of the problem.

Proposition 3.3 *The integrality constraints on h_p variables can be relaxed in Formulation 2.*

Proof: In the objective function, all coefficients for h_p variables are positive and the formulation maximizes the objective function value. Therefore, in the optimal solution, h_p variables take values as large as possible. From Constraints (3.28), we know that the largest value h_p variables can take is:

$$h_p = \min_{\langle l,k \rangle \in S(p)} \left\{ \sum_{\pi \in \Pi} f_{lk\pi} \right\}.$$

By Lemma 3.1, $\sum_{\pi \in \Pi} f_{lk\pi} \in \{0,1\}$. Hence, $h_p \in \{0,1\}$. ■

The LP relaxation of a mixed integer programming (MIP) model relaxes all integrality constraints on variables. A key property of a formulation for a MIP problem is its *strength.* Consider two equivalent formulations of the same MIP model. One formulation is strictly stronger than the other if the feasible region of its LP relaxation

is strictly smaller than the other. A stronger formulation is more likely to produce objective function values closer to the value of the mixed integer program. In the following proposition, we show that Formulation 2 is strictly stronger than Formulation 1.

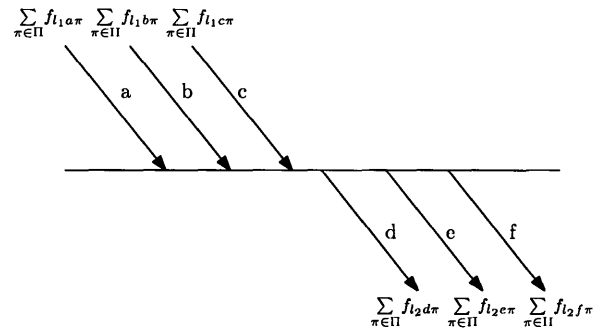
Proposition 3.4 *Formulation 2 is strictly stronger than Formulation 1.*

Proof: A feasible solution to the LP relaxation of Formulation 2 satisfies Constraints (3.28), hence the following constraints are satisfied:

$$\sum_{\langle l,k\rangle \in S(p)} h_p \leq \sum_{\langle l,k\rangle \in S(p)} \sum_{\pi \in \Pi} f_{lk\pi}, \forall p \in PC. \tag{3.29}$$

Because each connection variable is formed by two flight legs, we can substitute $\sum_{\langle l,k\rangle \in S(p)} h_p = 2h_p$ into Equation (3.29) to create Constraints (3.22). Because all other constraints are the same in both formulations, any feasible solution to the LP relaxation of Formulation 2 is a feasible solution to the LP relaxation of Formulation 1.

Next, we show that the feasible region of the LP relaxation of Formulation 2 is strictly smaller than that of the LP relaxation of Formulation 1. Take the example shown in Figure 3-2. Three copies of the inbound flight leg l_1 are created and in-



Connection variables	$\langle l_2, d\rangle$	$\langle l_2, e\rangle$	$\langle l_2, f\rangle$
$\langle l_1, a\rangle$	h_1	h_2	h_3
$\langle l_1, b\rangle$	h_4	h_5	h_6
$\langle l_1, c\rangle$	h_7	h_8	h_9

Figure 3-2: Example to illustrate formulation strength

dexed as a, b, and c, and three copies of the outbound flight leg l_2 are created and indexed as d, e, and f. These flight copies form nine connection variables, that is, $\mathbf{h} = \{h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9\}$. The flight leg copies corresponding to each connection variable are shown in the table. The following values are feasible in the LP relaxation of Formulation 1, but infeasible in the LP relaxation of Formulation 2:

$$\left\{\sum_{\pi\in\Pi} f_{l_1a\pi}, \sum_{\pi\in\Pi} f_{l_1b\pi}, \sum_{\pi\in\Pi} f_{l_1c\pi}, \sum_{\pi\in\Pi} f_{l_2d\pi}, \sum_{\pi\in\Pi} f_{l_2e\pi}, \sum_{\pi\in\Pi} f_{l_2f\pi}\right\} = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right\},$$

and

$$\mathbf{h} = \{\frac{1}{2}, \frac{3}{8}, \frac{3}{8}, \frac{3}{8}, \frac{1}{4}, \frac{1}{4}, \frac{3}{4}, \frac{1}{4}, \frac{1}{4}\}.$$

It is infeasible in the LP relaxation of Formulation 2 because Constraints (3.28) require that $h_2 \leq \{\sum_{\pi\in\Pi} f_{l_1a\pi}, \sum_{\pi\in\Pi} f_{l_2e\pi}\} \leq \min\{\frac{1}{2}, \frac{1}{4}\} = \frac{1}{4}$. ■

3.4.3 Formulation 3

Take the example shown in Figure 3-1, Formulation 3 generates three constraints for the three connection variables involving $\langle l_1, a\rangle$, specifically,

$$\begin{aligned} h_1 &\leq \sum_{\pi\in\Pi} f_{l_1a\pi}, \\ h_2 &\leq \sum_{\pi\in\Pi} f_{l_1a\pi}, \\ h_3 &\leq \sum_{\pi\in\Pi} f_{l_1a\pi}. \end{aligned}$$

Recognizing the fact that at most one of h_1, h_2, and h_3 can be non-zero in any feasible integer solution, the following constraint must hold:

$$h_1 + h_2 + h_3 \leq \sum_{\pi\in\Pi} f_{l_1a\pi}.$$

Hence, for all $p = (\langle l_1, k_1\rangle, \langle l_2, k_2\rangle) \in C(l_1, l_2)$ $((l_1, l_2) \in C)$, the following constraints must be satisfied:

$$\sum_{p\in\{p|\langle l_1,k_1\rangle\in S(p),p\in C(l_1,l_2)\}} h_p \leq \sum_{\pi\in\Pi} f_{l_1k_1\pi}, \forall k_1 \in C(l_1), \forall (l_1,l_2) \in C \tag{3.30}$$

and

$$\sum_{p\in\{p|\langle l_2,k_2\rangle\in S(p),p\in C(l_1,l_2)\}} h_p \leq \sum_{\pi\in\Pi} f_{l_2k_2\pi}, \forall k_2 \in C(l_2), \forall (l_1,l_2) \in C. \tag{3.31}$$

Constraints (3.30) and (3.31) can replace Constraints (3.22) to produce Formulation 3.

Proposition 3.5 *Formulation 3 is equivalent to Formulation 1.*

Proof: For a potentially connecting itinerary p' formed by $\langle l_1, k_1\rangle$ and $\langle l_2, k_2\rangle$, Constraints (3.30) and (3.31) in Formulation 3 correspond to:

$$h_{p'} + \sum_{p\in\{p|\langle l_1,k_1\rangle\in S(p),p\in C(l_1,l_2)\}\setminus\{p'\}} h_p \leq \sum_{\pi\in\Pi} f_{l_1k_1\pi},$$
$$h_{p'} + \sum_{p\in\{p|\langle l_2,k_2\rangle\in S(p),p\in C(l_1,l_2)\}\setminus\{p'\}} h_p \leq \sum_{\pi\in\Pi} f_{l_2k_2\pi}.$$

Summing the above equations, we get:

$$\begin{aligned} & 2h_{p'} + \sum_{p\in\{p|\langle l_1,k_1\rangle\in S(p),p\in C(l_1,l_2)\}\setminus\{p'\}} h_p \\ & + \sum_{p\in\{p|\langle l_2,k_2\rangle\in S(p),p\in C(l_1,l_2)\}\setminus\{p'\}} h_p \leq \sum_{\pi\in\Pi} f_{l_1k_1\pi} + \sum_{\pi\in\Pi} f_{l_2k_2\pi} \\ \Longrightarrow \quad & 2h_{p'} \leq \sum_{\pi\in\Pi} f_{l_1k_1\pi} + \sum_{\pi\in\Pi} f_{l_2k_2\pi}. \end{aligned}$$

This is true for all $p' \in PC$, hence, Constraints (3.30) and (3.31) in Formulation 3 imply Constraints (3.22) in Formulation 1.

When h_p variables and $f_{lk\pi}$ variables are binary, Constraints (3.22) are equivalent

to

$$h_p \leq \min_{\langle l,k\rangle \in S(p)} \left\{ \sum_{\pi \in \Pi} f_{lk\pi} \right\}.$$

Take any $(l_1, l_2) \in C$ with inbound flight leg $\langle l_1, k_1 \rangle$ forming potentially connecting itineraries with outbound flight legs $\langle l_2, k_2 \rangle$ ($k_2 \in \widehat{C}(l_2) \subseteq C(l_2)$). Let $\sum_{\pi \in \Pi} f_{l_2 k_{l_2} \pi} = 1$ ($k_{l_2} \in C(l_2)$) and $\sum_{\pi \in \Pi} f_{l_2 k_2 \pi} = 0$ ($k_2 \neq k_{l_2}, k_2 \in C(l_2)$). We have:

$$\begin{aligned}
&\sum_{p \in \{p | \langle l_1, k_1 \rangle \in S(p), p \in C(l_1, l_2)\}} h_p \\
\leq &\sum_{p \in \{p | \langle l_1, k_1 \rangle \in S(p), p \in C(l_1, l_2)\}} \min_{\langle l,k\rangle \in S(p)} \left\{ \sum_{\pi \in \Pi} f_{lk\pi} \right\} \\
= &\sum_{k_2 \in \widehat{C}(l_2)} \min \left\{ \sum_{\pi \in \Pi} f_{l_1 k_1 \pi}, \sum_{\pi \in \Pi} f_{l_2 k_2 \pi} \right\} \\
\leq &\min \left\{ \sum_{\pi \in \Pi} f_{l_1 k_1 \pi}, \sum_{\pi \in \Pi} f_{l_2 k_{l_2} \pi} \right\} \\
\leq &\sum_{\pi \in \Pi} f_{l_1 k_1 \pi}.
\end{aligned}$$

Therefore Constraints (3.30) are implied by Constraints (3.22). Similarly, we can prove that Constraints (3.31) are implied by Constraints (3.22). Hence, Constraints (3.30) and (3.31) are equivalent to Constraints (3.22). Because all other constraints and the objective function are the same in both formulations, Formulation 3 is equivalent to Formulation 1. ■

Proposition 3.6 *The integrality constraints on h_p variables can be relaxed in Formulation 3.*

Proof: By Lemma 3.1, among all copies of flight leg $l \in L$, let $f_{lk_l\pi}$ take value one; and the rest take value zero, that is, $f_{lk\pi} = 0$ ($k \in C(l), k \neq k_l$). For all connection variables $p \in C(l_1, l_2)$, the corresponding constraints in Constraints (3.30) and (3.31) are:

$$\sum_{p \in \{p | \langle l_1, k_{l_1} \rangle \in S(p), p \in C(l_1, l_2)\}} h_p \leq \sum_{\pi \in \Pi} f_{l_1 k_{l_1} \pi} = 1,$$

$$\sum_{p\in\{p|\langle l_2,k_{l_2}\rangle\in S(p),p\in C(l_1,l_2)\}} h_p \leq \sum_{\pi\in\Pi} f_{l_2k_{l_2}\pi} = 1,$$

$$\sum_{p\in\{p|\langle l_1,k_1\rangle\in S(p),p\in C(l_1,l_2)\}} h_p \leq \sum_{\pi\in\Pi} f_{l_1k_1\pi} = 0, \forall k_1 \in C(l_1) \setminus \{k_{l_1}\},$$

$$\sum_{p\in\{p|\langle l_2,k_2\rangle\in S(p),p\in C(l_1,l_2)\}} h_p \leq \sum_{\pi\in\Pi} f_{l_2k_2\pi} = 0, \forall k_2 \in C(l_2) \setminus \{k_{l_2}\}.$$

Therefore,

$$h_p \leq 1, \text{when } p = (\langle l_1, k_{l_1}\rangle, \langle l_2, k_{l_2}\rangle),$$
$$h_p = 0, \text{when } p \neq (\langle l_1, k_{l_1}\rangle, \langle l_2, k_{l_2}\rangle), p \in C(l_1, l_2).$$

Because the objective coefficients for h_p variables are positive and we are maximizing the objective function value, h_p variables take values as large as possible in the optimal solution. Hence in the optimal solution, we have:

$$h_p = 1, \text{when } p = (\langle l_1, k_{l_1}\rangle, \langle l_2, k_{l_2}\rangle),$$
$$h_p = 0, \text{when } p \neq (\langle l_1, k_{l_1}\rangle, \langle l_2, k_{l_2}\rangle), p \in C(l_1, l_2).$$

The above is true for all $p \in C(l_1, l_2), (l_1, l_2) \in C$, therefore all h_p variables take zero or one in the optimal solution even with the absence of the integrality constraints.■

Proposition 3.7 *Formulation 3 is strictly stronger than Formulation 2.*

Proof: For a potentially connecting itinerary p' formed by $\langle l_1, k_1\rangle$ and $\langle l_2, k_2\rangle$, Constraints (3.30) and (3.31) in Formulation 3 correspond to:

$$h_{p'} + \sum_{p\in\{p|\langle l_1,k_1\rangle\in S(p),p\in C(l_1,l_2)\}\setminus\{p'\}} h_p \leq \sum_{\pi\in\Pi} f_{l_1k_1\pi},$$

$$h_{p'} + \sum_{p\in\{p|\langle l_2,k_2\rangle\in S(p),p\in C(l_1,l_2)\}\setminus\{p'\}} h_p \leq \sum_{\pi\in\Pi} f_{l_2k_2\pi}.$$

Therefore, we have:

$$h_{p'} \leq \sum_{\pi \in \Pi} f_{l_1 k_1 \pi},$$
$$h_{p'} \leq \sum_{\pi \in \Pi} f_{l_2 k_2 \pi}.$$

The above is the same as Constraints (3.28). Because all other constraints are the same, any feasible solution to the LP relaxation of Formulation 3 is a feasible solution to the LP relaxation of Formulation 2.

Next we show that there exists some feasible solution to the LP relaxation of Formulation 2, which is infeasible to the LP relaxation of Formulation 3. Referring to the example shown in Figure 3-2, the following values are feasible in the LP relaxation of Formulation 2, but infeasible in the LP relaxation of Formulation 3:

$$\left\{ \sum_{\pi \in \Pi} f_{l_1 a \pi}, \sum_{\pi \in \Pi} f_{l_1 b \pi}, \sum_{\pi \in \Pi} f_{l_1 c \pi}, \sum_{\pi \in \Pi} f_{l_2 d \pi}, \sum_{\pi \in \Pi} f_{l_2 e \pi}, \sum_{\pi \in \Pi} f_{l_2 f \pi} \right\} = \left\{ \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{1}{4}, \frac{1}{4} \right\},$$

and

$$\mathbf{h} = \{ \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \}.$$

It is infeasible in the LP relaxation of Formulation 3, because Constraints (3.30) require that $h_1 + h_2 + h_3 \leq \sum_{\pi \in \Pi} f_{l_1 a \pi} = \frac{1}{2}$. ■

3.4.4 Formulation 4

In this section, we present a formulation that is weaker than Formulation 3, but exhibits better computational performance when exploring the branch-and-bound tree in our solution approach.

The connection variable h_p can take value zero or one, thus we can think of sending one unit of flow on potentially connecting itinerary p and have a pseudo-capacity (M_{lk}) for each flight copy $\langle l, k \rangle$. Take the example in Figure 3-3. Depicted are one inbound flight leg l_1, two outbound flight legs l_2 and l_3, each with three copies, and six connection variables h_1, h_2, $\cdots$ and h_6. The six connection variables all traverse

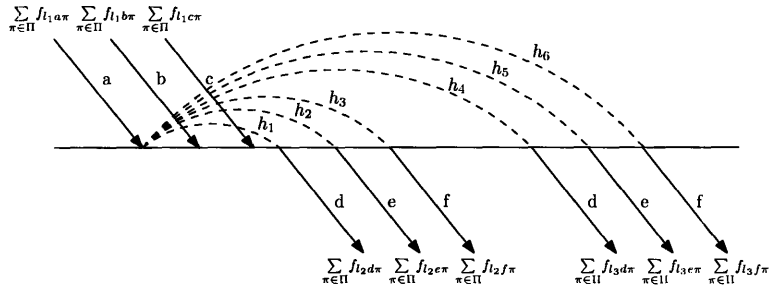


Figure 3-3: Illustration of Formulation 4

$\langle l_1, a \rangle$. If any of the six connection variables take value one in a feasible integer solution, $\sum_\pi f_{l_1 a \pi}$ must take value one. On the other hand, if $\sum_\pi f_{l_1 a \pi}$ takes value zero, none of the six connection variables can take value one. Such a relationship can be captured by the following constraint:

$$h_1 + h_2 + \cdots + h_6 \leq M_{l_1 a} \sum_{\pi \in \Pi} f_{l_1 a \pi}. \tag{3.32}$$

$M_{l_1 a}$ represents a sufficiently large number (the pseudo capacity of $\langle l_1, a \rangle$), which guarantees that if $\sum_{\pi \in \Pi} f_{l_1 a \pi} = 1$, Constraint (3.32) is not binding. Following similar logic, we replace Constraints (3.22) in Formulation 1 with Constraints (3.33) to model the relationship between h_p variables and $f_{lk\pi}$ variables in Formulation 4:

$$\sum_{p \in PC} \zeta_p^{lk} h_p \leq M_{lk} \sum_{\pi \in \Pi} f_{lk\pi}, \forall l \in L, k \in C(l). \tag{3.33}$$

Clearly smaller values of M_{lk} lead to a smaller feasible region and potentially tighter bounds on the optimal objective function value. We introduce a new binary parameter for each $(l_1, l_2) \in C$:

$$\rho_{(l_1,l_2)}^{lk} = \begin{cases} 1, & \text{if } \exists p \in C(l_1, l_2), \; s.t. \; \langle l, k \rangle \in S(p); \\ 0, & \text{otherwise.} \end{cases}$$

We know that:

$$\begin{aligned}
\sum_{p\in PC} \zeta_p^{lk} h_p &= \sum_{(l_1,l_2)\in C} \sum_{p\in C(l_1,l_2)} \zeta_p^{lk} h_p \\
&= \sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{lk} \sum_{p\in C(l_1,l_2)} \zeta_p^{lk} h_p \\
&\leq \sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{lk} \sum_{p\in C(l_1,l_2)} h_p \\
&\leq \sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{lk}. \qquad \text{(by Lemma 3.1)}
\end{aligned}$$

Hence, we set M_{lk} to $\sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{lk}$ to tighten Constraints (3.33) and include the following constraints in Formulation 4:

$$\sum_{p\in PC} \zeta_p^{lk} h_p \leq \sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{lk} \sum_{\pi\in\Pi} f_{lk\pi}, \forall l \in L, k \in C(l). \tag{3.34}$$

Proposition 3.8 *Formulation 4 is equivalent to Formulation 1.*

Proof: We first prove that Constraints (3.34) in Formulation 4 imply Constraints (3.22) in Formulation 1. The constraints corresponding to a particular $p' \in PC$ in Constraints (3.34) are:

$$\begin{aligned}
&\sum_{p\in PC} \zeta_p^{lk} h_p \leq \sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{lk} \sum_{\pi\in\Pi} f_{lk\pi}, \forall \langle l,k\rangle \in S(p') \\
\iff \quad & h_{p'} + \sum_{p\in PC\setminus\{p'\}} \zeta_p^{lk} h_p \leq \sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{lk} \sum_{\pi\in\Pi} f_{lk\pi}, \forall \langle l,k\rangle \in S(p').
\end{aligned}$$

If $\exists \langle l,k\rangle \in S(p')$, such that $\sum_{\pi\in\Pi} f_{lk\pi} = 0$, we have $h_{p'} = 0$. Therefore, $2h_{p'} \leq \sum_{\langle l,k\rangle\in S(p')} \sum_{\pi\in\Pi} f_{lk\pi}$, the constraint corresponding to p' in Constraints (3.22) holds. If $\sum_{\pi\in\Pi} f_{lk\pi} = 1$ $(\forall \langle l,k\rangle \in S(p'))$, regardless of the value of $h_{p'}$, the constraint corresponding to p' in Constraints (3.22) holds.

Now we prove that Constraints (3.22) in Formulation 1 imply Constraints (3.34) in Formulation 4. When h_p variables and $f_{lk\pi}$ variables are binary, Constraints (3.22)

are equivalent to:

$$h_p \leq \min_{(l,k)\in S(p)} \left\{ \sum_{\pi\in\Pi} f_{lk\pi} \right\}, \forall p \in PC.$$

For any inbound flight leg $\langle l_1, k_1 \rangle$, we have:

$$\begin{aligned}
\sum_{p\in PC} \zeta_p^{l_1k_1} h_p &\leq \sum_{p\in PC} \zeta_p^{l_1k_1} \min_{(l,k)\in S(p)} \left\{ \sum_{\pi\in\Pi} f_{lk\pi} \right\} \\
&= \sum_{(l_1,l_2)\in C} \sum_{p\in C(l_1,l_2)} \zeta_p^{l_1k_1} \min_{(l,k)\in S(p)} \left\{ \sum_{\pi\in\Pi} f_{lk\pi} \right\} \\
&= \sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{l_1k_1} \sum_{p\in C(l_1,l_2)} \zeta_p^{l_1k_1} \min_{(l,k)\in S(p)} \left\{ \sum_{\pi\in\Pi} f_{lk\pi} \right\} \\
&\leq \sum_{(l_1,l_2)\in C} \rho_{(l_1,l_2)}^{l_1k_1} \sum_{\pi\in\Pi} f_{l_1k_1\pi}. \qquad (3.35)
\end{aligned}$$

We prove the last inequality in the above equation as follows. For any inbound flight leg $\langle l_1, k_1 \rangle$, assume it forms potentially connecting itineraries with outbound flight legs $\langle l_2, k_2 \rangle$ ($k_2 \in \widehat{C}(l_2) \subseteq C(l_2)$). Let $\sum_{\pi\in\Pi} f_{l_2k_{l_2}\pi} = 1$ ($k_{l_2} \in C(l_2)$) and $\sum_{\pi\in\Pi} f_{l_2k_2\pi} = 0$ ($k_2 \neq k_{l_2}, k_2 \in C(l_2)$). We have:

$$\begin{aligned}
&\sum_{p\in C(l_1,l_2)} \zeta_p^{l_1k_1} \min_{(l,k)\in S(p)} \left\{ \sum_{\pi\in\Pi} f_{lk\pi} \right\} \\
=&\sum_{k_2\in\widehat{C}(l_2)} \min \left\{ \sum_{\pi\in\Pi} f_{l_1k_1\pi}, \sum_{\pi\in\Pi} f_{l_2k_2\pi} \right\} \\
\leq&\min \left\{ \sum_{\pi\in\Pi} f_{l_1k_1\pi}, \sum_{\pi\in\Pi} f_{l_2k_{l_2}\pi} \right\} \\
\leq&\sum_{\pi\in\Pi} f_{l_1k_1\pi}.
\end{aligned}$$

Following similar logic, Inequality (3.35) is true for any outbound flight leg. Therefore Constraints (3.34) in Formulation 4 are equivalent to Constraints (3.22) in Formulation 1. Because all other constraints and the objective function are the same in both formulations, Formulation 4 is equivalent to Formulation 1. ■

Proposition 3.9 *The integrality constraints on h_p variables can be relaxed in For-*

mulation 4.

Proof: In the objective function, the coefficients, or weights, for h_p variables are positive and we are maximizing the weighted sum of h_p variables. Therefore, in the optimal solution, h_p ($p \in PC$) variables take values as large as possible. For a particular $p' \in PC$, Constraints (3.34) are:

$$h_{p'} + \sum_{p \in PC} \zeta_p^{lk} h_p \leq \sum_{(l_1,l_2) \in C} \rho_{(l_1,l_2)}^{lk} \sum_{\pi \in \Pi} f_{lk\pi}, \forall \langle l,k \rangle \in S(p').$$

When $\sum_{\pi \in \Pi} f_{lk\pi} = 1$ ($\forall \langle l,k \rangle \in S(p')$), the above constraints are not binding (this is satisfied when we estimate M_{lk}). Therefore h'_p takes value one to maximize the objective function value. If $\exists \langle l,k \rangle \in S(p')$ such that $\sum_{\pi \in \Pi} f_{lk\pi}$ takes value zero in the feasible integer solution, $h_{p'}$ is forced to zero. Hence, integrality of h_p variables is guaranteed. ■

Proposition 3.10 *Formulation 3 is strictly stronger than Formulation 4.*

Proof: For an inbound flight leg $l_1 \in L$, by Constraints (3.30) in Formulation 3, we get:

$$\sum_{p \in \{p | \langle l_1,k_1 \rangle \in S(p), p \in C(l_1,l_2)\}} h_p \leq \sum_{\pi \in \Pi} f_{l_1k_1\pi}, \forall k_1 \in C(l_1), \forall (l_1,l_2) \in C$$

$$\Longrightarrow \quad \sum_{(l_1,l_2) \in C} \rho_{(l_1,l_2)}^{l_1k_1} \sum_{p \in \{p | \langle l_1,k_1 \rangle \in S(p), p \in C(l_1,l_2)\}} h_p \leq \sum_{(l_1,l_2) \in C} \rho_{(l_1,l_2)}^{l_1k_1} \sum_{\pi \in \Pi} f_{l_1k_1\pi}, \forall k_1 \in C(l_1)$$

$$\Longrightarrow \quad \sum_{p \in PC} \zeta_p^{l_1k_1} h_p \leq \sum_{(l_1,l_2) \in C} \rho_{(l_1,l_2)}^{l_1k_1} \sum_{\pi \in \Pi} f_{l_1k_1\pi}, \forall k_1 \in C(l_1).$$

Similarly, we can prove that for an outbound flight leg $l_2 \in L$, by Constraints (3.31) in Formulation 3, we get:

$$\sum_{p \in PC} \zeta_p^{l_2k_2} h_p \leq \sum_{(l_1,l_2) \in C} \rho_{(l_1,l_2)}^{l_2k_2} \sum_{\pi \in \Pi} f_{l_2k_2\pi}, \forall k_2 \in C(l_2).$$

Therefore,

$$\sum_{p \in PC} \zeta_p^{lk} h_p \leq \sum_{(l_1,l_2)} \rho_{(l_1,l_2)}^{lk} \sum_{\pi \in \Pi} f_{lk\pi}, \forall l \in L, k \in C(l).$$

The above are the same as Constraints (3.34). Hence, any feasible solution to the LP relaxation of Formulation 3 is feasible to the LP relaxation of Formulation 4 and Formulation 3 is at least as strong as Formulation 4.

Now we show that the feasible region of the LP relaxation of Formulation 3 is strictly smaller than that of the LP relaxation of Formulation 4. Take the example shown in Figure 3-3, and the following values are feasible in the LP relaxation of Formulation 4, but infeasible in the LP relaxation of Formulation 3:

$$\begin{pmatrix} \sum_{\pi \in \Pi} f_{l_1 a\pi} & \sum_{\pi \in \Pi} f_{l_1 b\pi} & \sum_{\pi \in \Pi} f_{l_1 c\pi} \\ \sum_{\pi \in \Pi} f_{l_2 d\pi} & \sum_{\pi \in \Pi} f_{l_2 e\pi} & \sum_{\pi \in \Pi} f_{l_2 f\pi} \\ \sum_{\pi \in \Pi} f_{l_3 d\pi} & \sum_{\pi \in \Pi} f_{l_3 e\pi} & \sum_{\pi \in \Pi} f_{l_3 f\pi} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

and

$$\{h_1, h_2, h_3, h_4, h_5, h_6\} = \{1, 0, 0, 0, 0, 0\}.$$

In the LP relaxation of Formulation 3, the following constraint is violated:

$$h_1 + h_2 + h_3 + h_4 + h_5 + h_6 \leq \sum_{\pi \in \Pi} f_{l_1 a\pi}.$$

■

In Table 3.3, we show the number of constraints needed to model the relationship between $f_{lk\pi}$ variables and h_p variables in each formulation. For $(l_1, l_2) \in C$, denote $L_{(l_1,l_2)}(l_1) = \{k | \langle l_1, k \rangle \in S(p), \exists p \in C(l_1, l_2)\}$ and $L_{(l_1,l_2)}(l_2) = \{k | \langle l_2, k \rangle \in S(p), \exists p \in C(l_1, l_2)\}$. Note that a is the number of flight copies created before and after the original flight leg. A sample value for the number of constraints needed in each formulation is provided based on the case study conducted in Section 3.8.

For a typical problem, Formulations 1 and 2 have the largest number of constraints

Formulation	Num. of constraints	Sample values
1	$\sum_{(l_1,l_2)\in C} \lvert C(l_1,l_2)\rvert$	879,505
2	$2\sum_{(l_1,l_2)\in C} \lvert C(l_1,l_2)\rvert$	1,759,010
3	$\sum_{(l_1,l_2)\in C} \left(\lvert L_{(l_1,l_2)}(l_1)\rvert + \lvert L_{(l_1,l_2)}(l_2)\rvert\right)$	317,785
4	$\lvert L\rvert(2a+1)$	5,824

Table 3.3: Number of constraints needed to model the relationship between $f_{lk\pi}$ variables and h_p variables in each formulation.

and do not provide a bound as tight as that in Formulation 3. They are not good candidates for implementation. Formulation 4 requires the least number of constraints, but provides a weaker bound than Formulation 3. We thus consider the question of whether or not we can achieve a bound similar to that provided by Formulation 3, but with fewer constraints than the number needed in Constraints (3.30) and (3.31) in Formulation 3. Because Constraints (3.30) require that:

$$\begin{aligned} & \sum_{p\in\{p|\langle l_1,k_1\rangle\in S(p),p\in C(l_1,l_2)\}} h_p \leq \sum_{\pi\in\Pi} f_{l_1k_1\pi}, \forall k_1 \in C(l_1), \forall (l_1,l_2)\in C \\ \Longrightarrow \quad & \sum_{k_1\in C(l_1)} \sum_{p\in\{p|\langle l_1,k_1\rangle\in S(p),p\in C(l_1,l_2)\}} h_p \leq \sum_{k_1\in C(l_1)} \sum_{\pi\in\Pi} f_{l_1k_1\pi}, \forall (l_1,l_2)\in C \\ \Longrightarrow \quad & \sum_{p\in C(l_1,l_2)\}} h_p \leq 1, \forall (l_1,l_2)\in C. \end{aligned} \tag{3.36}$$

The same results can be obtained from Constraints (3.31). We add Constraints (3.36) to Formulation 4 to tighten its bound. The number of constraints in Constraints (3.36) is 30,058 in our experiments, or 10% of the number of constraints in Constraints (3.30) and (3.31). As is demonstrated in Section 3.7, Constraints (3.36) tighten the bounds of Formulation 4 significantly, and also improve the bounds of Formulations 1 and 2 dramatically.

3.4.5 Additional Insights

In this section, we show two propositions related to the formulations presented in this chapter. They provide additional insights toward these formulations.

Lemma 3.2

$$\sum_{(l_1,l_2)\in C} \rho^{lk}_{(l_1,l_2)} \leq \sum_{p\in PC} \zeta^{lk}_p, \forall (l_1,l_2) \in C.$$

Proof: We have:

$$\begin{aligned} & \rho^{lk}_{(l_1,l_2)} \leq \sum_{p\in C(l_1,l_2)} \zeta^{lk}_p, \forall (l_1,l_2) \in C \\ \iff & \sum_{(l_1,l_2)\in C} \rho^{lk}_{(l_1,l_2)} \leq \sum_{(l_1,l_2)\in C} \sum_{p\in C(l_1,l_2)} \zeta^{lk}_p \\ \iff & \sum_{(l_1,l_2)\in C} \rho^{lk}_{(l_1,l_2)} \leq \sum_{p\in PC} \zeta^{lk}_p. \end{aligned}$$

■

Although it is hard to determine the relative strengths of Formulations 1 and 2, if the objective is to increase the number of potentially connecting itineraries in the solution, that is, if the weights for all connection variables are the same, we have the following result.

Proposition 3.11 *If $w_p = \overline{w} > 0$ $(\forall p \in PC)$, the LP relaxation of Formulation 4 is at least as strong as that of Formulation 1.*

Proof: Summing up all constraints in Constraints (3.34), we get:

$$\sum_{l\in L} \sum_{k\in C(l)} \sum_{p\in PC} \zeta^{lk}_p h_p \leq \sum_{l\in L} \sum_{k\in C(l)} \sum_{(l_1,l_2)\in C} \rho^{lk}_{(l_1,l_2)} \sum_{\pi\in\Pi} f_{lk\pi}. \tag{3.37}$$

We have:

$$\sum_{l\in L} \sum_{k\in C(l)} \sum_{p\in PC} \zeta^{lk}_p h_p = 2 \sum_{p\in PC} h_p.$$

We also have:

$$\sum_{l\in L}\sum_{k\in C(l)}\sum_{(l_1,l_2)\in C}\rho^{lk}_{(l_1,l_2)}\sum_{\pi\in\Pi}f_{lk\pi}\le\sum_{l\in L}\sum_{k\in C(l)}\sum_{p\in PC}\zeta^{lk}_{p}\sum_{\pi\in\Pi}f_{lk\pi}\qquad\text{by Lemma 3.2}$$

$$=\sum_{p\in PC}\left(\sum_{l\in L}\sum_{k\in C(l)}\zeta^{lk}_{p}\sum_{\pi\in\Pi}f_{lk\pi}\right)$$

$$=\sum_{p\in PC}\sum_{\langle l,k\rangle\in S(p)}\sum_{\pi\in\Pi}f_{lk\pi}.$$

Therefore, Equation (3.37) can be re-written as:

$$2\sum_{p\in PC}h_p\le\sum_{p\in PC}\sum_{\langle l,k\rangle\in S(p)}\sum_{\pi\in\Pi}f_{lk\pi}. \tag{3.38}$$

Assume that in the optimal solution to the LP relaxation of Formulation 1, variables $f_{lk\pi}$ take values $\overline{f}_{lk\pi}$, and that in the optimal solution to the LP relaxation of Formulation 4, variables $f_{lk\pi}$ take values $\widetilde{f}_{lk\pi}$ and variables h_p take values $\widetilde{h}_p$. Then the optimal value of the LP relaxation of Formulation 4 satisfies the following inequality:

$$z_2=\sum_{p\in PC}w_p\widetilde{h}_p=\sum_{p\in PC}\overline{w}\widetilde{h}_p=\overline{w}\sum_{p\in PC}\widetilde{h}_p$$

$$\le\frac{1}{2}\overline{w}\sum_{p\in PC}\sum_{\langle l,k\rangle\in S(p)}\sum_{\pi\in\Pi}\widetilde{f}_{lk\pi}\qquad\text{by Eq. (3.38)}$$

$$\le\frac{1}{2}\overline{w}\sum_{p\in PC}\sum_{\langle l,k\rangle\in S(p)}\sum_{\pi\in\Pi}\overline{f}_{lk\pi}$$

$$=z_1.$$

We now demonstrate why the last inequality holds. It is not difficult to show that $\widetilde{f}_{lk\pi}$ is a partial feasible solution to the LP relaxation of Formulation 1, that is, if variables $f_{lk\pi}$ are fixed to values $\widetilde{f}_{lk\pi}$, the LP relaxation of Formulation 1 is still feasible. By Proposition 3.1, the corresponding optimal objective value of the LP relaxation of Formulation 1 when $f_{lk\pi}$ variables take values $\widetilde{f}_{lk\pi}$ is:

$$\frac{1}{2}\sum_{p\in PC} w_p \sum_{(l,k)\in S(p)} \sum_{\pi\in\Pi} \widetilde{f}_{lk\pi} = \frac{1}{2}\overline{w} \sum_{p\in PC} \sum_{(l,k)\in S(p)} \sum_{\pi\in\Pi} \widetilde{f}_{lk\pi}$$
$$\leq \frac{1}{2}\overline{w} \sum_{p\in PC} \sum_{\langle l,k\rangle\in S(p)} \sum_{\pi\in\Pi} \overline{f}_{lk\pi}.$$

■

Proposition 3.12 *Assume that 2a flight copies are created and indexed as shown in Figure 3-4. Copy 0 corresponds to the flight time in the original peaked schedule. Let*

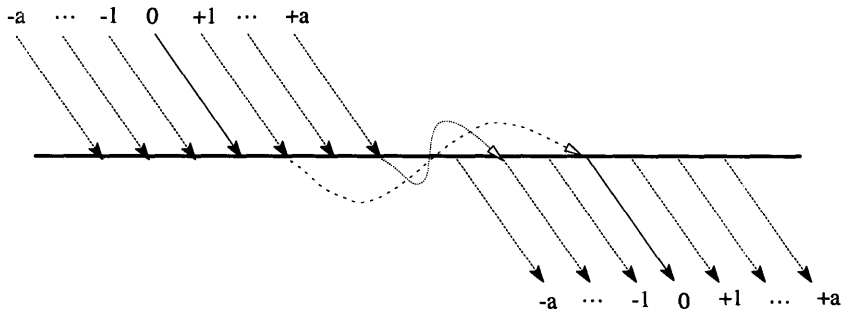


Figure 3-4: Flight copy indices

d be the time displacement between flight copies. Constraints (3.36) can always reject the optimal solution to the LP relaxation of Formulation 1, if:

1. *there exists some* $(l_1, l_2) \in C$, *such that* $T_1 \leq CT(\langle l_1, 0\rangle, \langle l_2, 0\rangle) \leq T_2$;
2. $4ad \leq T_2 - T_1$;
3. *at least one of the following sums are fractional in the optimal LP relaxation solution:* $\sum_{\pi\in\Pi} f_{l_1,k_1,\pi}$ $(k_1 \in \{-a, \cdots, 0, \cdots, +a\})$ *and* $\sum_{\pi\in\Pi} f_{l_2,k_2,\pi}$ $(k_2 \in \{-a, \cdots, 0, \cdots, +a\})$; *or* $m \geq 1$, *where* $m = argmax_b\{b | CT(\langle l_1, 0\rangle, \langle l_2, 0\rangle) - bd \geq T_1, b \in \overline{\mathbb{N}^-}\}$.

Proof: Without loss of generality, assume that $CT(\langle l_1, 0\rangle, \langle l_2, 0\rangle) \leq \frac{1}{2}(T_1 + T_2)$. Figure 3-5 shows the flight copies of l_1 and l_2 and each intersection in the grid corresponds to a possible connection variable these flight copies can form. If the connection

time for a connection variable is between T_1 and T_2, it corresponds to a potentially connecting itinerary and is called a valid connection variable. For example, intersection A corresponds to the connection variable formed by $\langle l_1, 0\rangle$ and $\langle l_2, 0\rangle$. By the definition of m, intersection B in the grid is the furthest point we can move if we leave A vertically and ensure that the corresponding connection variable is valid.

We have $CT(\langle l_1, -a\rangle, \langle l_2 + a\rangle) = CT(\langle l_1, 0\rangle, \langle l_2, 0\rangle) + 2ad \leq \frac{1}{2}(T_1+T_2)+\frac{1}{2}(T_2-T_1) \leq T_2$, therefore any intersection above the dashed line in the grid corresponds to a valid connection variable. Let $h_{(\langle l_1,k_1\rangle,\langle l_2,k_2\rangle)}$ denote the valid connection variable formed by $\langle l_1, k_1\rangle$ and $\langle l_2, k_2\rangle$.

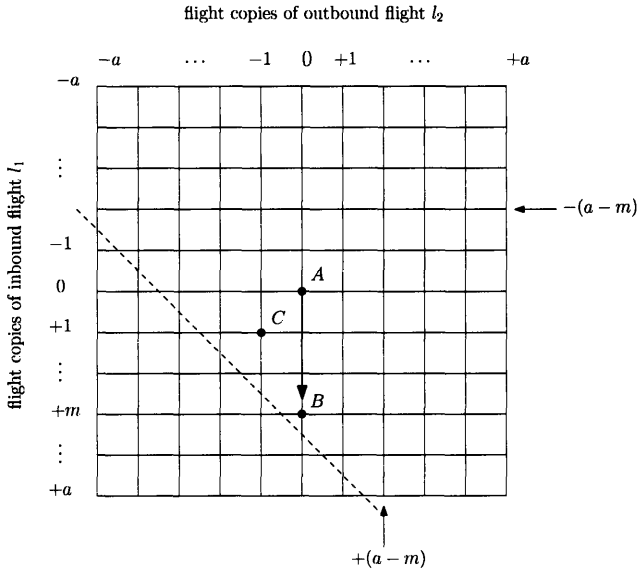


Figure 3-5: Connection variables formed by flight copies of l_1 and l_2

Let $f_{lk\pi} = \overline{f}_{lk\pi}$ in the optimal solution to the LP relaxation of Formulation 1. From Proposition 3.1, we can compute the values of connection variables in the optimal solution to the LP relaxation. Take any intersection above the dashed line, say intersection C, the value of this connection variable in the optimal LP solution is

$$h_{(\langle l_1,+1\rangle,\langle l_2,-1\rangle)} = \frac{1}{2}\sum_{\pi\in\Pi}\overline{f}_{l_1,+1,\pi} + \frac{1}{2}\sum_{\pi\in\Pi}\overline{f}_{l_2,-1,\pi}.$$

If we sum the optimal values of all valid connection variables formed by copies of l_1 and l_2, we get

$$\begin{aligned}
\sum_{p \in C(l_1,l_2)} h_p =& \frac{1}{2}\Bigg((m+1)\sum_{\pi\in\Pi} \overline{f}_{l_1,+a,\pi} + (m+2)\sum_{\pi\in\Pi} \overline{f}_{l_1,+(a-1),\pi} \\
& + \cdots + 2a\sum_{\pi\in\Pi} \overline{f}_{l_1,-(a-m-1),\pi} + (2a+1)\sum_{k=-(a-m)}^{-a}\sum_{\pi\in\Pi} \overline{f}_{l_1 k\pi} \Bigg) \\
& + \frac{1}{2}\Bigg((m+1)\sum_{\pi\in\Pi} \overline{f}_{l_2,-a,\pi} + (m+2)\sum_{\pi\in\Pi} \overline{f}_{l_2,-(a-1),\pi} \\
& + \cdots + 2a\sum_{\pi\in\Pi} \overline{f}_{l_2,+(a-m-1),\pi} + (2a+1)\sum_{k=+(a-m)}^{+a}\sum_{\pi\in\Pi} \overline{f}_{l_2 k\pi} \Bigg) \\
>& \frac{1}{2}\Bigg(\sum_{\pi\in\Pi} \overline{f}_{l_1,+a,\pi} + \sum_{\pi\in\Pi} \overline{f}_{l_1,+(a-1),\pi} \\
& + \cdots + \sum_{\pi\in\Pi} \overline{f}_{l_1,-(a-m-1),\pi} + \sum_{k=-(a-m)}^{-a}\sum_{\pi\in\Pi} \overline{f}_{l_1 k\pi} \Bigg) \\
& + \frac{1}{2}\Bigg(\sum_{\pi\in\Pi} \overline{f}_{l_2,-a,\pi} + \sum_{\pi\in\Pi} \overline{f}_{l_2,-(a-1),\pi} \\
& + \cdots + \sum_{\pi\in\Pi} \overline{f}_{l_2,+(a-m-1),\pi} + \sum_{k=+(a-m)}^{+a}\sum_{\pi\in\Pi} \overline{f}_{l_2 k\pi} \Bigg) \\
=& \frac{1}{2}\sum_{k=+a}^{-a}\sum_{\pi\in\Pi} \overline{f}_{l_1 k\pi} + \frac{1}{2}\sum_{k=-a}^{+a}\sum_{\pi\in\Pi} \overline{f}_{l_2 k\pi} \\
=& 1.
\end{aligned}$$

The strict inequality is due to Condition 3. Therefore adding Constraints (3.36) renders the current optimal LP relaxation solution infeasible. ■

Condition 1 is easily satisfied, simply requiring the flight copies corresponding to the original flight times form a potentially connecting itinerary. Because we set $a = 3$, $d = 10$ min, $T_1 = 10$ min, and $T_2 = 190$ min in our experiments, Condition 2 is satisfied. In the LP relaxation, it is fairly easy to have $\sum_{\pi\in\Pi} f_{l_1,k_1,\pi}$ ($k_1 \in \{-a, \cdots, 0, \cdots, +a\}$) and $\sum_{\pi\in\Pi} f_{l_2,k_2,\pi}$ ($k_2 \in \{-a, \cdots, 0, \cdots, +a\}$) take on fractional

values and/or $m \geq 1$.

3.5 Restricted Robust De-Peaking Model

The robust de-peaking model (Formulations 1 through 4) presented in Section 3.4 suffers from issues of computational tractability due to the large number of variables - for a typical problem in our experiments, there are approximately 350,000 x_{mr} variables, 870,000 h_p variables and another 90,000 $f_{lk\pi}$, $y_{g\pi}$ and z_π variables. Even the LP relaxations of these formulations cannot be solved in 20 hours on the workstation used in the research.

To address this tractability issue, we observe that the sole purpose for including variables x_{mr} in the model is to compute the revenue in Constraint (3.21), that is, the term $\sum_{m \in M} \sum_{r \in R(m)} x_{mr} fare_{mr}$. Instead of measuring the exact profit of the model and requiring it to be no less than p^*, we propose the following technique to approximate the profitability requirement. We can eliminate all x_{mr} variables by ensuring that the revenue associated with our robust de-peaking solution is at least as great as that in our basic de-peaking solution and the cost associated with our robust de-peaking solution is no more than that in our basic de-peaking model. To achieve this, we impose the following constraints to ensure that the revenue associated with our robust de-peaking solution is at least as great as that of the basic de-peaking solution:

1. We let Q denote the set of flight leg pairs forming connecting itineraries used in the basic de-peaking solution. Then, for each $(l_1, l_2) \in Q$, we let $Q(l_1, l_2) = \{(k_1, k_2) | MinCT \leq CT(\langle l_1, k_1 \rangle, \langle l_2, k_2 \rangle) \leq MaxCT\}$. We require any flight leg pair in Q to form a feasible connection in the robust de-peaking solution as follows:

$$\sum_{\pi \in \Pi} f_{l_1 k_1 \pi} + \sum_{\pi \in \Pi} f_{l_2 k_2 \pi} \leq 1, \forall (l_1, l_2) \in Q, (k_1, k_2) \notin Q(l_1, l_2).$$

2. We let PAX_l represent the number of passengers assigned to flight leg l in the

basic de-peaking solution. We require the capacity of the fleet type assigned to flight leg l in the robust solution to be no less than PAX_l, that is:

$$\sum_{k \in C(l)} \sum_{\pi \in \Pi} f_{lk\pi} CAP_\pi \geq PAX_l, \forall l \in L.$$

On the cost side, we limit the cost of our robust solution to c^*, the cost of the basic de-peaking solution. This treatment of revenue and cost guarantees that the profit of the robust de-peaking solution is no less than that of the basic de-peaking solution. With this guarantee, we can eliminate all x_{mr} variables from the robust de-peaking model, thereby significantly reducing the size of the model. Unlike the robust de-peaking model in which the objective is to select the most robust solution from the entire set of profit maximizing de-peaking solutions, the objective of this *restricted* model is to select the most robust solution from the *subset* of profit maximizing de-peaking solutions whose costs are no more than and revenues are no less than those in the basic de-peaking solution.

To summarize, we present our restricted versions of Formulations 1 through 4, first providing the set of constraints common to all formulations, and then presenting the additional constraints specific to each restricted formulation.

maximize $\sum_{p \in PC} w_p h_p$

subject to:

$$\sum_{k \in C(l)} \sum_{\pi \in \Pi} f_{lk\pi} = 1, \forall l \in L \tag{3.39}$$

$$\sum_{l \in L} \sum_{k \in C(l)} f_{lk\pi} \overline{\alpha}^i_{lk\pi} + \sum_{g \in G^\pi} y_{g\pi} \widehat{\alpha}^i_{g\pi} = 0, \forall i \in N^\pi, \pi \in \Pi \tag{3.40}$$

$$\sum_{l \in L} \sum_{k \in C(l)} f_{lk\pi} \overline{\beta}_{lk\pi} + \sum_{g \in G^\pi} y_{g\pi} \widehat{\beta}_{g\pi} = z_\pi, \forall \pi \in \Pi \tag{3.41}$$

$$z_\pi \leq n^\pi, \forall \pi \in \Pi \tag{3.42}$$

$$\sum_{l \in L} \sum_{k \in C(l)} \gamma_{lk}^{at} \sum_{\pi \in \Pi} f_{lk\pi} \leq MAX^{at}, \forall t \in T \tag{3.43}$$

$$\sum_{l \in L} \sum_{k \in C(l)} \gamma_{lk}^{dt} \sum_{\pi \in \Pi} f_{lk\pi} \leq MAX^{dt}, \forall t \in T \tag{3.44}$$

$$f_{lk\pi} \in \{0,1\}, \forall l \in L, k \in C(l), \pi \in \Pi \tag{3.45}$$

$$y_{g\pi} \geq 0, \forall g \in G^{\pi}, \pi \in \Pi \tag{3.46}$$

$$z_{\pi} \geq 0, \forall \pi \in \Pi \tag{3.47}$$

$$\sum_{l \in L} \sum_{k \in C(l)} \sum_{\pi \in \Pi} c_{lk\pi} f_{lk\pi} + \sum_{\pi \in \Pi} z_{\pi} c_{\pi} \leq c^{*} \tag{3.48}$$

$$\sum_{\pi \in \Pi} f_{l_1 k_1 \pi} + \sum_{\pi \in \Pi} f_{l_2 k_2 \pi} \leq 1, \forall (l_1, l_2) \in Q, (k_1, k_2) \notin Q(l_1, l_2) \tag{3.49}$$

$$\sum_{k \in C(l)} \sum_{\pi \in \Pi} f_{lk\pi} CAP_{\pi} \geq PAX_l, \forall l \in L \tag{3.50}$$

Constraints (3.39) through (3.47) are the same as those in the basic de-peaking model presented in Section 3.3. Constraint (3.48) limits the cost of the model to that of the basic de-peaking solution. Constraints (3.49) and (3.50), respectively, guarantee that connecting itineraries in the basic de-peaking solution remain feasible in the robust de-peaking solution, and ensure that the capacity of the aircraft assigned to $l \in L$ is no less than PAX_l. This ensures that the profit associated with the robust solution is not less than that of the basic de-peaking model. Additional constraints specific to each formulation are summarized as follows:

Formulation 1-R:

$$2h_p \leq \sum_{<l,k> \in S(p)} \sum_{\pi \in \Pi} f_{lk\pi}, \forall p \in PC \tag{3.51}$$

$$\sum_{p \in C(l_1, l_2)} h_p \leq 1, \forall (l_1, l_2) \in C \tag{3.52}$$

$$h_p \in \{0,1\}, \forall p \in PC \tag{3.53}$$

Formulation 2-R:

$$h_p \leq \sum_{\pi \in \Pi} f_{lk\pi}, \forall \langle l,k \rangle \in S(p), \forall p \in PC \tag{3.54}$$

$$\sum_{p \in C(l_1,l_2)} h_p \leq 1, \forall (l_1,l_2) \in C \tag{3.55}$$

$$0 \leq h_p \leq 1, \forall p \in PC \tag{3.56}$$

Formulation 3-R:

$$\sum_{p \in \{p | \langle l_1,k_1 \rangle \in S(p), p \in PC(l_1,l_2)\}} h_p \leq \sum_{\pi \in \Pi} f_{l_1 k_1 \pi}, \forall k_1 \in C(l_1), \forall (l_1,l_2) \in C \tag{3.57}$$

$$\sum_{p \in \{p | \langle l_2,k_2 \rangle \in S(p), p \in PC(l_1,l_2)\}} h_p \leq \sum_{\pi \in \Pi} f_{l_2 k_2 \pi}, \forall k_2 \in C(l_2), \forall (l_1,l_2) \in C \tag{3.58}$$

$$0 \leq h_p \leq 1, \forall p \in PC \tag{3.59}$$

Formulation 4-R:

$$\sum_{p \in PC} \zeta_p^{lk} h_p \leq \sum_{(l_1,l_2) \in C} \rho_{(l_1,l_2)}^{lk} \sum_{\pi \in \Pi} f_{lk\pi}, \forall l \in L, k \in C(l) \tag{3.60}$$

$$\sum_{p \in C(l_1,l_2)} h_p \leq 1, \forall (l_1,l_2) \in C \tag{3.61}$$

$$0 \leq h_p \leq 1, \forall p \in PC \tag{3.62}$$

3.6 Solution Approach

The solution approach to the robust de-peaking model is outlined in Figure 3-6. The basic de-peaking model is solved to provide c^* and Q, inputs to the robust de-peaking problem. While solving the restricted robust de-peaking model, for reasons stated in Section 3.7, Formulation 4-R is used to find good integer solutions, while the LP relaxation of Formulation 3-R is solved to obtain a bound (Z_L) on the optimal integer solution value and to gauge the optimality of the integer solutions obtained from Formulation 4-R.

To solve Formulation 4-R, an iterative approach is taken in which a *Restricted Master Problem* (RMP), containing only a subset of all possible h_p variables, is solved repeatedly until a near-optimal solution to the model is obtained. The rationale underlying the selection of the variables included in the RMP is as follows. Let $C_b = \{(l_1, l_2) \mid |C(l_1, l_2)| = b, (l_1, l_2) \in C\}$, where $0 \leq b \leq (2a+1)^2$, $b \in \mathbb{N}$, and a is the number of flight leg copies created before and after the original flight leg copy, respectively. By definition, $C = C_1 \bigcup C_2 \bigcup \cdots \bigcup C_{(2a+1)^2}$. Let X be a subset of C, and $P(X)$ be the set $\{p | p \in C(l_1, l_2), (l_1, l_2) \in X\}$. For any flight leg pair $(l_1, l_2) \in C_{(2a+1)^2}$, any pair of copies of flight leg l_1 and l_2 selected in an integer solution form a valid connection with the same contribution $w_p = w_{(l_1,l_2)}$. The removal of all columns corresponding to h_p, $p \in P(C_{(2a+1)^2})$ thus have the effect of reducing the objective function value for any solution by a constant amount. The result is that an optimal solution to the RMP is also optimal to the original model containing all decision variables.

Extending this approach, we can remove all columns corresponding to h_p with $p \in P(C_{(2a+1)^2-1})$, however, we can no longer guarantee that an optimal solution to RMP is optimal for the original problem because the solution to the RMP might select, from all $(2a+1)^2$ potential connections, the *only* pair of flight leg copies that does not create a potentially connecting itinerary. The risk of doing so is minimal but increases as the subscript b of C_b decreases. We let $\widetilde{C}_b = C_{(2a+1)^2} \bigcup C_{(2a+1)^2-1} \bigcup \cdots \bigcup C_b$ and eliminate all h_p variables with $p \in P(\widetilde{C}_b)$ in Formulation 4-R to create RMP. The branch-and-bound algorithm is used to find integer solutions to RMP. Once an integer solution is found, we denote the objective function value of RMP as Z_R and let its solution be denoted by $f'_{lk\pi}$. If connection $p \in P(\widetilde{C})$ is enabled in this current integer solution, we set h_p to 1; otherwise, we set h_p to 0, for each $p \in P(\widetilde{C})$. The objective function value of the current feasible solution for Formulation 4-R, denoted Z_I is then equal to $Z_R + \sum_{p \in P(\widetilde{C})} w_p h_p$. We compute the optimality gap as $gap = (Z_L - Z_I)/Z_I$. If the optimality gap is acceptable, that is, within our defined threshold value, the algorithm terminates; otherwise, we compute the value of $h_{(l_1,l_2)} = \sum_{p \in C(l_1,l_2)} h_p$, for each $(l_1, l_2) \in \widetilde{C}$. For each $(l_1, l_2) \in \widetilde{C}$ for which $h_{(l_1,l_2)} \neq 1$, all columns corresponding

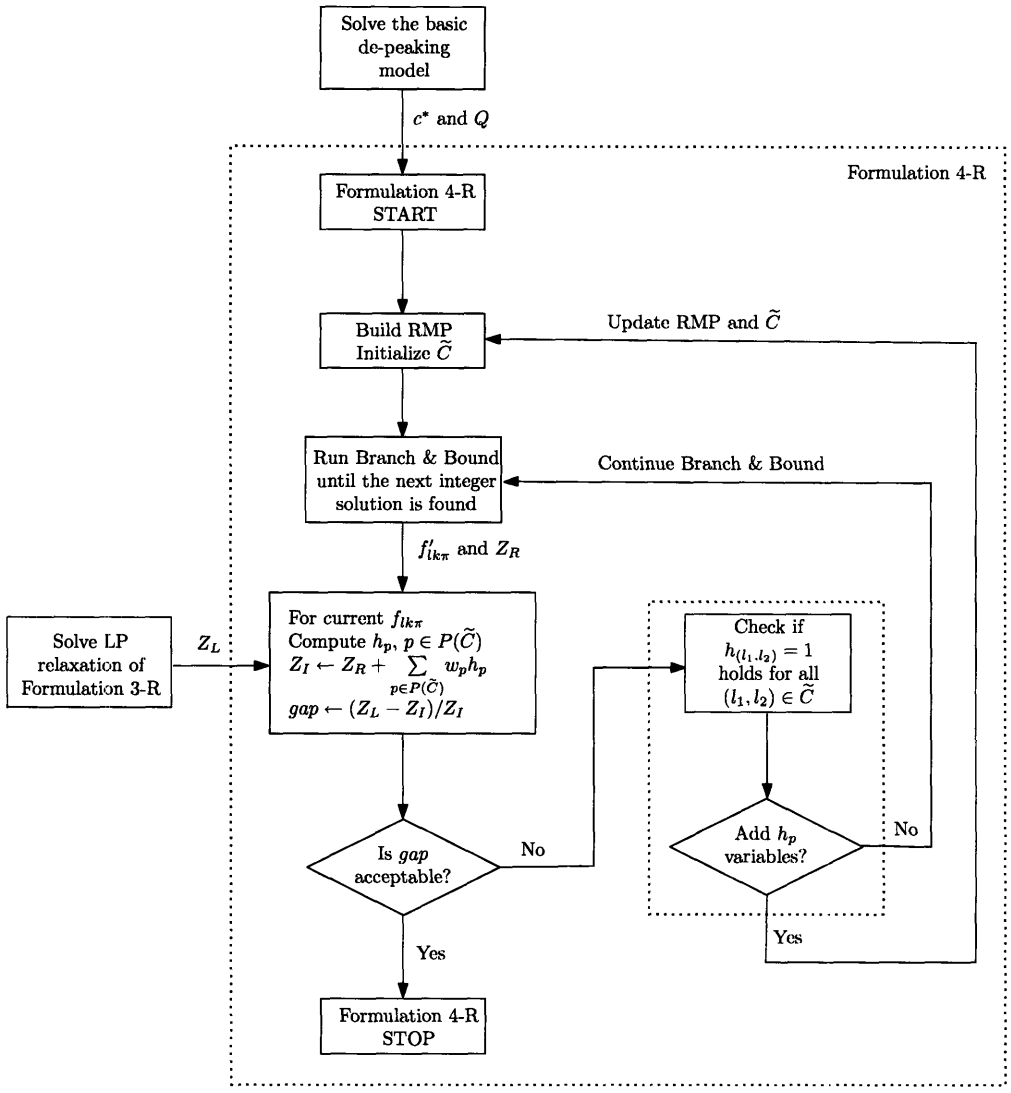


Figure 3-6: Solution algorithm for the robust de-peaking model

to h_p, $p \in C(l_1, l_2)$ are added to *RMP* and $\widetilde{C}_b$ is updated as follows: $\widetilde{C}_b \leftarrow \widetilde{C}_b \setminus (l_1, l_2)$. Next, *RMP* is re-solved. If $h_{(l_1,l_2)} = 1$ holds true for all $(l_1, l_2) \in \widetilde{C}$, we continue the branch-and-bound algorithm to identify another feasible solution for Formulation 4-R. Adding columns to the current *RMP* requires that we re-start the branch-and-bound algorithm from scratch (see ILOG, 2003, p. 523). This can be computationally time consuming, hence it can sometimes be beneficial to continue the current branch-and-bound search for additional integer solutions without interruption, even when the equality $h_{(l_1,l_2)} = 1$ does not hold for some $(l_1, l_2) \in \widetilde{C}$.

To improve tractability of the branch-and-bound algorithm, we replace branching based on variable dichotomy with branching based on Type I Special Ordered Sets for Constraints (3.39) and assign CAP_π as weights to each variable $f_{lk\pi}$ (see ILOG, 2003, p. 329). Suppose that at one node in the branch-and-bound tree, variables in the cover constraint for flight leg l_0, that is,

$$\sum_{k \in C(l_0)} \sum_{\pi \in \Pi} f_{l_0 k \pi} = 1,$$

take values $f^*_{l_0 k \pi}$ $(\forall \pi \in \Pi, k \in C(l_0))$ and some of these values are fractional. We compute $\overline{w} = \sum_{k \in C(l_0)} \sum_{\pi \in \Pi} f^*_{l_0 k \pi} CAP_\pi$ and partition the set of binary variables $f_{l_0 k \pi}$ $(\forall \pi \in \Pi, k \in C(l_0))$ into two groups, $G_1 = \{f_{l_0 k \pi} | CAP_\pi \leq \overline{w}, k \in C(l_0), \pi \in \Pi\}$ and $G_2 = \{f_{l_0 k \pi} | CAP_\pi > \overline{w}, k \in C(l_0), \pi \in \Pi\}$. On one branch of our branching strategy, we impose the following restriction: $\sum_{f_{l_0 k \pi} \in G_1} f_{l_0 k \pi} = 1$; and on the second branch, we require: $\sum_{f_{l_0 k \pi} \in G_2} f_{l_0 k \pi} = 1$. Hane et al. (1995) show that branching based on Type I special ordered sets is a more effective branching strategy than branching on individual variables (although Hane et al. apply a slightly different method to create G_1 and G_2).

3.7 Computational Experiences

The basic de-peaking model and the restricted robust de-peaking model are both implemented in C using ILOG CPLEX 9.0. Computational experiments are conducted on a workstation equipped with one Intel Pentium 4 2.8 GHz processor and 1 GB RAM.

In both the basic de-peaking model and the robust de-peaking model, seven copies of each flight leg l are created (that is, $a = 3$), placing one each at -30, -20, -10, 0, +10, +20, and +30 minutes offset from leg l's scheduled departure time in the peaked schedule. The number of departures and the number of arrivals are each limited to 5 per 10-minute interval, respectively. $MinCT = 25$ minutes and $MaxCT = 180$ minutes.

The basic de-peaking model is solved using the MIP solution routine from CPLEX callable library. In our experiments with the robust de-peaking model, $T_1 = MinCT -$ 20 minutes and $T_2 = MaxCT + 10$ minutes. We consider as potentially connecting itineraries those that violate the minimum connection time requirement by 20 minutes, while we consider a maximum 10-minute violation of the maximum connection time. Our rationale in doing so is that new connections with short durations (just slightly longer than $MinCT$) are more salable (although more likely to be disrupted) than connections with long durations (just shorter than $MaxCT$).

3.7.1 Comparison of the LP relaxations

In Table 3.4, we report, for Formulations 1-R through 4-R, the sizes of the models after CPLEX preprocessing (see ILOG, 2003, p. 322-324), and the optimal objective function values of the corresponding LP relaxations. In the first five rows in the table, we provide results when Constraints (3.36) (that is, Constraints (3.52), (3.61), and (3.55)) are excluded from Formulations 1-R through 3-R. In the second block, we report results when these constraints are included. Note that Constraints (3.36) are redundant in Formulation 3-R.

Our first observation is that Constraints (3.36) are effective in decreasing dramatically the values of Z_L and tightening the LP bounds of Formulations 1-R, 2-R and 3-R. Our second observation is that Z_L values increase as we move from Formulation 3-R to Formulation 2-R to Formulation 1-R. Formulation 3 provides the tightest bound regardless of whether Constraints (3.36) are included in Formulations 1, 2, and 3, or not. These results are consistent with the proofs presented in Section 3.4, where we compare the relative strengths of the LP relaxations of these Formulations. The number of columns for all formulations are similar, however, the number of rows differ widely. Formulation 3-R provides the tightest bound, and hence, we use it to generate bounds with which to measure the optimality gap of integer solutions.

	Formulation 1-R	Formulation 2-R	Formulation 3-R	Formulation 4-R
Num. of Rows	623,117	1,004,302	221,858	13,782
Num. of Cols.	654,480	509,454	514,834	514,748
Num. of NZ.	1,731,336	1,968,925	1,200,925	995,759
Z_L	694,690	333,865	129,142	180,331
$(Z_L - Z_I)/Z_I$	450.39%	164.52%	2.32%	42.87%
		With Constraints (3.36)		
Num. of Rows	652,690	1,032,284		36,408
Num. of Cols.	659,877	514,806		514,921
Num. of NZ.	2,359,001	2,451,719		1,469,236
Z_L	154,200.8	133,690.4		142,739
$(Z_L - Z_I)/Z_I$	22.17%	5.92%		13.09%

Table 3.4: Comparison of LP relaxations of the full problem across formulations ($Z_I = 126,217$)

3.7.2 Searching for Integer Solutions

In solving the *RMP* of Formulation 4-R, we must identify the value of b that allows near-optimal solutions to be generated within reasonable runtimes. In Table 3.5, we provide statistics of the values for h_p variables in the restricted robust de-peaking model when the values of $f_{lk\pi}$ variables are set to corresponding values in the basic de-peaking solution. In the first column, we report the values of each b with $|C_b| \geq 0$. In the second column, we provide the corresponding number of (l_1, l_2) pairs in C_b. In the third column, we record the value of $|P(C_b)|$, that is, the number of h_p variables for all (l_1, l_2) pairs in C_b. We provide in the fourth column, the number of h_p columns corresponding to $p \in P(C_b)$ as a fraction of the total number of h_p columns in the model. We show in the fifth column the number of (l_1, l_2) pairs that satisfy $h_{(l_1,l_2)} = 1$ in the basic de-peaking solution. Finally, we report in the last column, the fraction of leg pairs in C_b that satisfy $h_{(l_1,l_2)} = 1$, that is, $P_b = \left(\sum_{(l_1,l_2)\in C_b} h_{(l_1,l_2)}\right)/|C_b|$. $1 - P_b$ is a proxy that measures the likelihood of generating non-optimal solutions when variables h_p ($p \in P(C_b)$) are excluded from the model. The larger the value of P_b, the smaller the magnitude of the error associated with the exclusion. We observe that h_p variables corresponding to $p \in P(C_{49} \bigcup C_{48} \bigcup C_{47} \bigcup C_{46})$ account for more than 60% of the total number of h_p variables. In the basic de-peaking solution, nearly all leg pairs in the set $C_{49} \bigcup C_{48} \bigcup C_{47} \bigcup C_{46}$ satisfy $h_{(l_1,l_2)} = 1$. Hence, it is likely that $h_{(l_1,l_2)} = 1$ for most of the $(l_1, l_2) \in C_{49} \bigcup C_{48} \bigcup C_{47} \bigcup C_{46}$ in an optimal integer

solution to Formulation 4-R. We therefore set b equal to 46 in constructing our initial *RMP*. This eliminates 569,300 h_p variables or 65% of total h_p variables.

b	$\lvert C_b \rvert$	$\lvert P(C_b) \rvert$	$\frac{\lvert P(C_b) \rvert}{\sum_{b'=1}^{49} \lvert P(C_{b'}) \rvert}$	$\sum_{(l_1,l_2) \in C_b} h_{(l_1,l_2)}$	P_b
1	2,317	2,317	0.27%	56	2.4%
2	47	94	0.01%	12	25.5%
3	2,926	8,778	1.01%	223	7.6%
4	33	132	0.02%	8	24.2%
5	99	495	0.06%	30	30.3%
6	2,281	13,686	1.57%	378	16.6%
7	328	2,296	0.26%	149	45.4%
9	80	720	0.08%	21	26.3%
10	1,744	17,440	2.00%	415	23.8%
11	67	737	0.08%	25	37.3%
12	47	564	0.06%	7	14.9%
13	69	897	0.10%	35	50.7%
14	293	4,102	0.47%	165	56.3%
15	1,065	15,975	1.83%	292	27.4%
18	64	1,152	0.13%	32	50.0%
20	57	1,140	0.13%	33	57.9%
21	623	13,083	1.50%	298	47.8%
22	35	770	0.09%	23	65.7%
25	96	2,400	0.28%	58	60.4%
27	104	2,808	0.32%	59	56.7%
28	493	13,804	1.58%	312	63.3%
29	44	1,276	0.15%	24	54.5%
32	134	4,288	0.49%	86	64.2%
34	459	15,606	1.79%	300	65.4%
35	386	13,510	1.55%	267	69.2%
36	67	2,412	0.28%	46	68.7%
39	1,347	52,533	6.02%	1,034	76.8%
41	146	5,986	0.69%	125	85.6%
42	254	10,668	1.22%	236	92.9%
43	2,168	93,224	10.69%	1,848	85.2%
46	3,671	168,866	19.36%	3,378	92.0%
48	3,571	171,408	19.65%	3,471	97.2%
49	4,674	229,026	26.26%	4,674	100.0%
sum		872,193	100.00%	18,120	-

Table 3.5: Summary statistics of h_p variables in the restricted de-peaking model when $f_{lk\pi}$ variables are fixed to corresponding values in the basic de-peaking solution

Table 3.6 summarizes the computational performance of Formulations 1-R through 4-R on *RMP*. We can see that the number of rows, columns, and non-zeros of the *RMP* are significantly reduced when compared to those reported in Table 3.4 with Constraints (3.36). Not surprisingly, Formulation 1-R has a huge number of fractional integer variables at the root node of the branch-and-bound tree. The fourth row

	Fm. 1-R	Fm. 2-R	Fm. 3-R	Fm. 4-R
Num. of rows	224,871	344,551	98,363	24,431
Num. of columns	243,898	193,691	193,741	193,803
Num. of non-zeros	797,734	819,479	436,733	504,118
Num. of fractionl var. at root node	42,523	4,060	3,225	3,572
Num. of node searched until first integer solution	Not found	Not found	Not found	760
Total num. of nodes searched in 10 hrs.	$\sim 2,700$	~ 50	~ 100	$\sim 9,000$

Table 3.6: Branch-and-bound results for Formulations 1-R through 4-R on the initial Restricted Master Problem

reports the number of nodes searched in the branch-and-bound tree until the first integer solution is found. Formulations 1-R, 3-R, and 4-R fail to find integer solutions in 10 hours, the maximum allowable solution time. The last row reports the total number of nodes searched in 10 hours. Clearly, Formulation 4-R exhibits the best performance during branch-and-bound and, hence, it is used in our solution algorithm to search for integer solutions.

3.8 Case Study

In this section, using data obtained from a major U.S. airline, we demonstrate the potential impact of applying dynamic scheduling to the robust de-peaked schedule, and compare our results with those obtained when we apply dynamic scheduling to the basic de-peaked schedule. The airline providing us data operates a hub-and-spoke network with a banked schedule. There are approximately 1000 flight legs serving about 100 cities daily, and about 300 flight legs departing from and 300 flight legs arriving at the major hub each day. For ease of exposition, we let Schedule A denote the original banked schedule, Schedule B the de-peaked schedule obtained with the basic de-peaking model, and Schedule C the de-peaked schedule obtained with the robust de-peaking model.

3.8.1 Resulting Schedule Characteristics

In Table 3.7, we summarize the results of schedule de-peaking, simply comparing Schedules B and C to Schedule A, with percentage changes measured against Schedule A. Revenues, costs, profits, and average connection times are obtained by running our Passenger Mix Model (PMM, see Section 2.5.3) on the corresponding schedules. In both Schedules B and C total revenues are reduced as a result of de-peaking; however, total costs are also reduced. The result is that schedule profitability after de-peaking improves slightly. Additionally, de-peaking allowing one fewer aircraft to be used in Schedules B and C than in Schedule A. We also observe that the profit for Schedule C is greater than that in Schedule B, which results from the technique we employ in the robust de-peaking model to satisfy the profit requirement. In the next row, we report average connection times for each schedule. We observe that average passenger connecting time increases by about 9 minutes after de-peaking. This is similar to the 10.7-minute increase in average connection time reported by American Airlines after the de-peaking of its Chicago (ORD) hub (Flint, 2002). In the last two rows, we report the objective values for Formulation 3-R for Schedules B (the basic de-peaking solution) and C (the robust de-peaking solution). The basic de-peaking solution

has an objective function value of 117,174, with an optimality gap of 10.10%, while the robust de-peaking solution has an objective function value of 126,218, with an optimality gap of 2.28%. Our solution approach improves upon the basic de-peaking solution and finds an integer solution with a satisfactory optimality gap.

	Schedule A	Schedule B	Schedule C
Revenue	8,170,245	8,146,066	8,165,746
	-	-0.30%	-0.06%
Cost	6,001,400	5,929,789	5,929,789
	-	-1.19%	-1.19%
Profit	2,168,845	2,216,277	2,235,957
	-	2.19%	3.09%
No. of aircraft used	171	170	170
Avg. conn. Time	72.7	81.6	81.1
Z_I	-	117,174	126,218
$(Z_L - Z_I)/Z_L$	-	10.21%	2.32%

Table 3.7: Resulting schedule characteristics ($Z_L = 129,142$)

In Figure 3-7, we depict the departure and arrival operations at the hub under Schedule A, clearly demonstrating peaked departures and arrivals. In Figure 3-8, we show the departure and arrival operations at the hub under Schedule B, this time exhibiting a smoothing of arrival and departure activities. We do not provide the departure and arrival operations at the hub under Schedule C because they are similar to those under Schedule B.

In Figure 3-9, we present connection time distributions. The dotted line corresponds to planned connection times for passenger itineraries in Schedule A based on actual booking records from the major U.S. airline. The average connecting time is 68.3 minutes. The spike between 25 minutes and 95 minutes corresponds to passenger itineraries with the outbound flight leg in the same bank as the inbound flight leg. The tail between 95 minutes to 180 minutes corresponds to passenger itineraries in which the outbound flight leg is in the next bank after that of the inbound flight leg. The other three lines in Figure 3-9 correspond to connection times for Schedules A, B, and C when a PMM is used to assign passengers to itineraries. Although some of the underlying assumptions of PMM are quite strong, the shape of the distribution obtained from PMM for Schedule A closely matches that obtained from actual passenger data. The distributions of connection times for Schedules B and C are much

smoother than those for Schedule A due to schedule de-peaking.

3.8.2 Comparison of Dynamic Scheduling Results

For Schedules B and C, we conduct the same dynamic scheduling experiments detailed in Chapter 2 for operations spanning one week. A one-time re-optimization point is set 21-days prior to departure; the number of re-timed flight legs is limited to 100 per day; and the number of re-fleeted flight legs is unconstrained. Forecast A assumes perfect forecasts are used in re-optimizing the schedule, and Forecast B uses historical demand averages as future demand forecasts. In Table 3.8, we let *Profit incr. (DS)* denote the profit increase achieved when dynamic scheduling (DS) is applied to each schedule, and the *Percent value* represents the percent improvement in profit. The average daily profit improvements for Schedule B is \$98,857, or 5.22% for Forecast A and \$44,959, or 2.37% for Forecast B. The average daily profit improvement for Schedule C is \$78,625, or 4.09% for Forecast A and \$35,228, or 1.83% for Forecast B. Dynamic scheduling helps to increase schedule profitability in both Schedules B and C.

In the next block in this table, we compare Schedules B and C and report changes in revenue, cost and profit that result solely from robust planning (RP). Although the average profit improvement over the static case in Schedule C is less than that in Schedule B, the overall profitability of Schedule C is greater than that of Schedule B. We observe average daily schedule profitability improvements of \$7,089, or 0.36% under Forecast A and \$17,592, or 0.91% under Forecast B. When perfect forecasts (Forecast A) are available, dynamic scheduling produces re-optimized schedules of sufficient quality that incorporating robustness has little impact. When perfect forecasts are not available (Forecast B), however, designing robustness into the original schedules (Schedule C) provides added flexibility and leads to larger profit improvements. Hence, by incorporating robustness considerations into schedule design, we produce a more profitable schedule. The increase of profit in the static case is attributable to improved network connectivity due to our robust planning approach. For example, \$32,929 more connecting revenue is achieved in Schedule C than in Schedule B. Ta-

ble 3.9 shows the dynamic scheduling results of Schedule B and Schedule C for each individual day of the experimental week.

	Schedule B			Schedule C		
	Static	Dynamic scheduling		Static	Dynamic scheduling	
		Forecast A	Forecast B		Forecast A	Forecast B
Nonstop rev.	3,998,124	4,030,383	4,010,964	3,992,517	4,029,400	4,013,353
Conn. Rev.	3,825,827	3,896,564	3,868,281	3,858,756	3,908,750	3,877,992
Revenue	7,823,951	7,926,947	7,879,245	7,851,273	7,938,150	7,891,345
Cost	5,929,789	5,933,928	5,940,124	5,929,789	5,938,041	5,934,632
Profit	1,894,162	1,993,020	1,939,121	1,921,484	2,000,109	1,956,713
Profit incr. (DS)		98,857	44,959		78,625	35,228
Pct. profit incr. (DS)		5.22%	2.37%		4.09%	1.83%
Incr. in nonstop rev. (RP)				-5,607	-983	2,388
Incr. in conn. rev. (RP)				32,929	12,186	9,712
Incr. in rev. (RP)				27,322	11,203	12,100
Incr. in cost (RP)				0	4,113	-5,492
Incr. in profit (RP)				27,322	7,089	17,592
Pct. profit incr. (RP)				1.44%	0.36%	0.91%

Table 3.8: Comparisons between Schedule B and Schedule C when averaged over a week's operation (in dollars). The number of re-timed flights is limited to 100 and the number of re-fleeted flights is unconstrained

3.8.3 Quality of the Robust Schedule

In this section, we examine if the benefits of applying dynamic scheduling to the robust Schedule C result from a poorly designed original schedule. We study the quality of Schedule C by analyzing the re-timing decisions made by the re-optimization model. To prevent re-timing decisions from being affected by noise due to inexact forecasts, we conduct our analysis using perfect information (Forecast A). If a flight leg is *frequently and consistently re-timed*, it might indicate that this flight leg is poorly positioned in the schedule. We classify a flight leg as being *frequently re-timed* if it is re-timed at least 5 out of 7 days. Similarly, we classify a flight leg as being *consistently re-timed* if the re-timing decisions for this flight leg occur in the same direction (that is, earlier or later than the original schedule) for each day it is re-timed. We summarize the re-timing decisions for Schedule C in Table 3.10. There are 29 frequently re-timed flight legs, among which 21 (or 2% of the total number of flight legs) are consistently re-timed. In Section 2.7, we similarly report that about 2% of the flight legs are frequently and consistently re-timed from their Schedule B departure and arrival times. After comparing the sets of flight legs in Tables 2.17 and

3.10, however, we find that the majority of the flights in the two lists differ. Because only 4 flight legs, or 0.4% of all flight legs, are consistently and frequently re-timed in *both* Schedules B and C (Table 3.11), we conclude that our ability to improve upon Schedule C through a priori flight leg re-timings is very limited.

3.9 Summary

In this chapter, we develop a robust de-peaking model to allow maximal flexibility in making future schedule changes dynamically. To the best of our knowledge, it is the first approach of this type. To evaluate our robust de-peaking model, we first develop a basic de-peaking model to provide a basis for comparison, and to calibrate schedule profitability when there are no uncertainties, and no need to embed robustness into the schedule. We then design the robust de-peaking model to maximize the weighted sum of potentially connecting itineraries, while achieving the same profitability as the solution to the basic model. Four formulations of this model are presented and studied. We develop a restricted robust de-peaking model to reduce problem size by an approximate treatment of the profitability requirement. We solve the restricted robust de-peaking model with a decomposition based approach involving a variable reduction technique and a new form of column generation. In our computational experiments, we demonstrate significant computational improvements as a result of our new approach. We show through experiments using data from a major U.S. airline, that that the schedule generated by the robust de-peaking model can potentially improve profit by 0.4-0.9% in addition to the improvement achieved by dynamic scheduling. In these same experiments, we also demonstrate, as expected, that robust schedules with built-in flexibility achieve greater improvements in profitability when forecasts are imperfect.

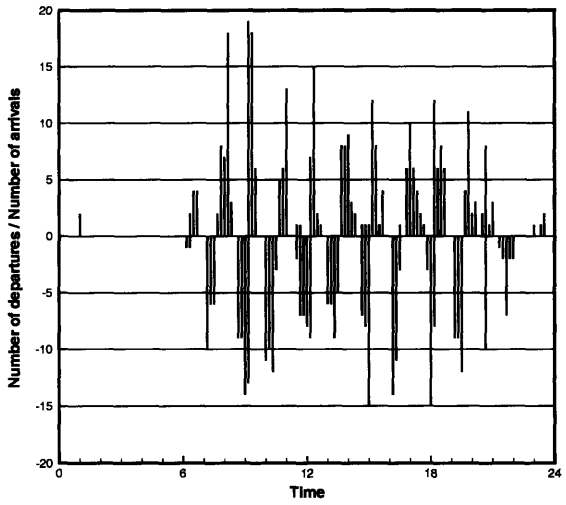


Figure 3-7: Number of departures (negative values) and arrivals (positive values) at hub under the banked schedule

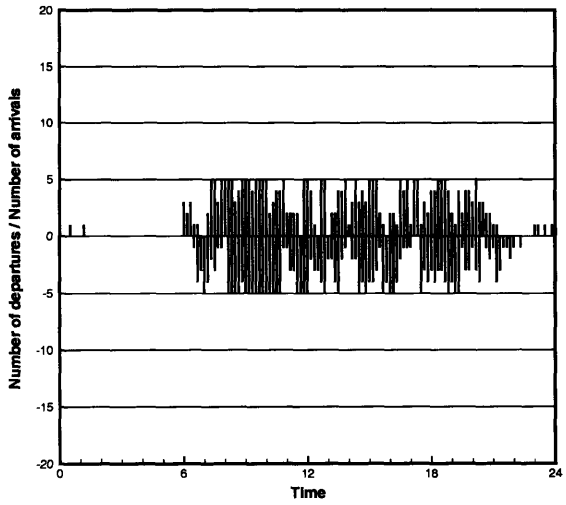


Figure 3-8: Number of departures (negative values) and arrivals (positive values) at hub under the de-peaked schedule obtained with the basic de-peaking model

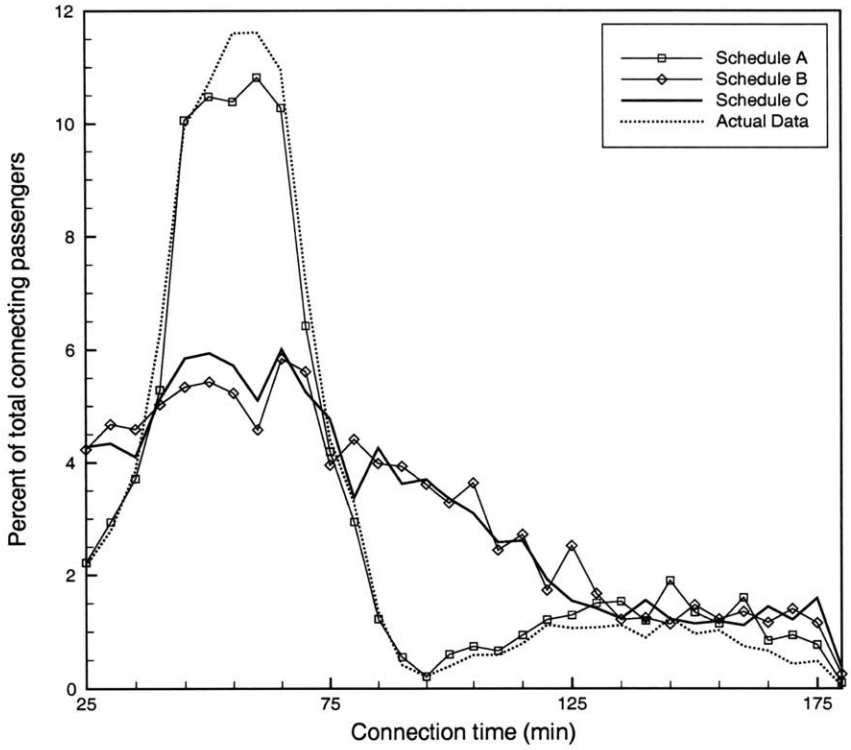


Figure 3-9: Distributions of passenger connection times based on actual passenger data under Schedule A and results from PMM under Schedules A, B, and C

	Schedule B			Schedule C		
	Static	Dynamic scheduling		Static	Dynamic scheduling	
		Forecast A	Forecast B		Forecast A	Forecast B
			Day 1			
Revenue	9,058,867	9,213,541	9,146,525	9,104,822	9,236,500	9,179,631
Cost	5,929,789	5,952,147	5,964,639	5,929,789	5,944,574	5,939,501
Profit	3,129,079	3,261,394	3,181,886	3,175,033	3,291,926	3,240,129
Profit incr. (DS)		132,315	52,807		116,893	65,096
		4.23%	1.69%		3.68%	2.05%
Profit incr. (RP)				45,954	30,532	58,243
				1.47%	0.94%	1.83%
			Day 2			
Revenue	8,018,606	8,136,941	8,052,977	8,031,747	8,131,864	8,067,222
Cost	5,929,789	5,941,897	5,933,586	5,929,789	5,948,798	5,934,447
Profit	2,088,817	2,195,045	2,119,391	2,101,958	2,183,065	2,132,775
Profit incr. (DS)		106,227	30,574		81,107	30,817
		5.09%	1.46%		3.86%	1.47%
Profit incr. (RP)				13,141	-11,979	13,383
				0.63%	-0.55%	0.63%
			Day 3			
Revenue	6,870,656	6,934,930	6,897,193	6,902,123	6,941,957	6,914,351
Cost	5,929,789	5,922,260	5,918,713	5,929,789	5,923,931	5,924,523
Profit	940,868	1,012,670	978,480	972,334	1,018,026	989,828
Profit incr. (DS)		71,803	37,612		45,692	17,494
		7.63%	4.00%		4.70%	1.80%
Profit incr. (RP)				31,467	5,356	11,348
				3.34%	0.53%	1.16%
			Day 4			
Revenue	7,363,064	7,446,343	7,400,498	7,404,744	7,469,634	7,421,413
Cost	5,929,789	5,919,607	5,948,519	5,929,789	5,928,067	5,926,856
Profit	1,433,276	1,526,736	1,451,980	1,474,955	1,541,567	1,494,557
Profit incr. (DS)		93,461	18,704		66,612	19,602
		6.52%	1.30%		4.52%	1.33%
Profit incr. (RP)				41,680	14,831	42,577
				2.91%	0.97%	2.93%
			Day 5			
Revenue	7,946,805	8,036,567	7,991,952	7,965,054	8,053,850	7,996,434
Cost	5,929,789	5,931,148	5,927,238	5,929,789	5,940,358	5,932,868
Profit	2,017,016	2,105,419	2,064,714	2,035,266	2,113,493	2,063,566
Profit incr. (DS)		88,403	47,698		78,227	28,301
		4.38%	2.36%		3.84%	1.39%
Profit incr. (RP)				18,250	8,074	-1,148
				0.90%	0.38%	-0.06%
			Day 6			
Revenue	8,333,631	8,457,232	8,408,739	8,344,732	8,455,586	8,418,011
Cost	5,929,789	5,937,127	5,929,069	5,929,789	5,939,380	5,932,256
Profit	2,403,842	2,520,105	2,479,670	2,414,943	2,516,206	2,485,756
Profit incr. (DS)		116,263	75,828		101,263	70,813
		4.84%	3.15%		4.19%	2.93%
Profit incr. (RP)				11,101	-3,899	6,085
				0.46%	-0.15%	0.25%
			Day 7			
Revenue	7,176,026	7,263,078	7,256,829	7,205,689	7,277,659	7,242,353
Cost	5,929,789	5,933,309	5,959,104	5,929,789	5,941,180	5,951,974
Profit	1,246,238	1,329,769	1,297,725	1,275,901	1,336,479	1,290,379
Profit incr. (DS)		83,531	51,487		60,578	14,478
		6.70%	4.13%		4.75%	1.13%
Profit incr. (RP)				29,663	6,709	-7,346
				2.38%	0.50%	-0.57%

Table 3.9: Comparisons between Schedule B and Schedule C for each individual day in a week's operation (in dollars). The number of re-timed flights is limited to 100 and the number of re-fleeted flights is unconstrained

Flight #	Flight detail					-15	-10	-5	0	+5	+10	+15	Consistent?
109	LAS	1900	→	LAX	2000						1	5	
645	SAN	2000	→	HUB	2122					3	1	2	
719	HUB	1348	→	FAT	1522	1		3				2	N
720	FAT	1545	→	HUB	1725			4				2	N
771	HUB	1730	→	SLC	2008						1	5	
50	HUB	925	→	ATL	1622		1	3	1	1			N
62	GEG	1430	→	HUB	1707	3	1		1			1	N
78	HUB	1743	→	LAS	1849		5		1				
238	ORD	1905	→	HUB	2036	2	3		1				
798	FLG	1907	→	HUB	2000	3	1	1	1				
194	LAS	1557	→	ORD	2133				2	2		2	
216	SNA	1010	→	HUB	1129	4			2				
220	SNA	1200	→	HUB	1320				2	1	1	2	
254	LAS	2023	→	SFO	2150				2		1	3	
263	HUB	1435	→	ABQ	1645				2	2		2	
278	TPA	820	→	HUB	944	3		1	2				
286	HUB	1507	→	BOS	2305				2			4	
378	ONT	610	→	HUB	722	4			2				
459	SMF	625	→	HUB	811	2	2		2				
563	DFW	1900	→	LAS	1946	2		2	2				
600	HUB	745	→	BUR	908			4	2				
642	HUB	1357	→	LGB	1515			1	2			3	N
682	HUB	1017	→	MRY	1205			3	2			1	N
746	DSM	745	→	HUB	852	4			2				
774	HUB	1012	→	SBP	1149	1			2	1	1	1	N
779	HUB	724	→	FAT	905		2	2	2				
780	FAT	925	→	HUB	1107		2	2	2				
782	SAN	1410	→	HUB	1530	2		1	2		1		N
822	CLD	1250	→	HUB	1422	1	3		2				

Table 3.10: Re-timing decisions for frequently re-timed flights in Schedule C

Schedule	Flight #	Flight detail					-15	-10	-5	0	+5	+10	+15
B	220	SNA	1200	→	HUB	1320				2	2		3
C	220	SNA	1200	→	HUB	1320				2	1	1	2
B	263	HUB	1425	→	ABQ	1635				1		1	5
C	263	HUB	1435	→	ABQ	1645				2	2		2
B	600	HUB	755	→	BUR	918			6	1			
C	600	HUB	745	→	BUR	908			4	2			
B	645	SAN	2000	→	HUB	2122				1	2	2	2
C	645	SAN	2000	→	HUB	2122					3	1	2

Table 3.11: Flights frequently and consistently re-timed in both Schedule B and Schedule C

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 4

Future Research Directions

Thirteen years ago, the pioneering work by Berge and Hopperstad (1993) brought us to the new era of dynamic scheduling. In subsequent years, researchers have further studied this topic, and some airlines have deployed dynamic scheduling approaches. Thirteen years later, with this thesis, we enrich the dynamic scheduling literature by introducing a new mechanism, that of flight re-timing. In view of the recent trend among airlines to de-peak hub operations, we combine flight re-fleeting and re-timing into an integrated schedule re-optimization approach and conduct experiments to evaluate its potential impact. In addition, we construct flight schedules with embedded flexibility to facilitate subsequent applications of dynamic scheduling in the booking period.

Even with these advances, there are still many topics left to be explored. We suggest the following future research directions to enhance knowledge in this area:

- **Evaluate the impact of dynamic scheduling on aircraft maintenance routing and crew planning.** An aircraft maintenance routing plan specifies the sequence of flight legs each aircraft operates and the resulting opportunities for routine maintenance. Similarly, crew pairing plans provide the sequence of flight legs each crew operates while satisfying the myriad crew work rules. If aircraft maintenance routing and crew pairings are planned before the last re-optimization point, the final re-optimized schedule might produce infeasibilities

in the maintenance or crew plans. An interesting future exercise is to evaluate the extent of the impact of dynamic scheduling to maintenance and crew plans; or to assess the potential to restore these plans when infeasibilities result. Typically, we expect that a feasible aircraft maintenance routing is not hard to find for a re-optimized schedule in a hub-and-spoke network. Crew duties, however, have to obey numerous work rules and can be hard to repair. Should it be too costly to repair disrupted crew duties, it might be desirable to develop re-optimization models incorporating additional constraints to ensure crew pairing feasibility.

- **Evaluate the case of multiple re-optimization points during the booking process.** For reasons of simplicity, we employ a one-time re-optimization point 21 days prior to departure. An interesting and challenging research question is to determine the optimal times and frequency with which to re-optimize the schedule.

- **Integrate dynamic scheduling with flexible passenger booking.** Talluri (2001) recognizes the fact that a large number of passengers are indifferent to the route they take to get from their origins to their destinations, as long as the different routes offer similar levels of service, that is, the departure and arrival times are approximately the same, and costs are the same. Talluri defines such a collection of routes as a *Route Set* and proposes a flexible booking approach in which passengers book route sets instead of specific routes, with the exact routing immediately determined if the booking request is accepted by the airline. Gallego and Phillips (2004) extend this flexible booking approach by delaying the determination of the exact routing until a later time. By delaying the exact route assignment of passengers, airlines can wait to see how future demands materialize and route previously booked passengers on flight legs with lower utilizations, thereby saving seats on congested flight legs for late-booking, high-fare passengers. This approach, which takes advantage of passenger routing flexibility and provides airlines with another mechanism for

handling mismatches between capacities and demands, can be integrated with dynamic airline scheduling to create an approach that is highly responsive to demand variations.

- **Evaluate and enhance the Passenger Mix Model.** In this research, we adopt a simplified treatment of passenger recapture in the passenger mix model. Specifically, we assume perfect recapture between selected subset of itineraries, and these subsets are specified empirically, not using more sophisticated consumer choice models. An interesting future research question is to evaluate the implications of such a simplification.

- **Define additional metrics to measure embedded schedule flexibility.** In this research, we use the weighted sum of *potentially connecting itineraries* as the metric defining schedule robustness, or flexibility to re-time flight legs and create additional itineraries where needed. An important research topic is to specify other metrics, especially those that integrate re-timing and re-fleeting impacts, and measure their effectiveness in creating robust schedules from which realized profits can be maximized.

- **Build a feedback loop between robust planning models and dynamic scheduling models.** The original schedule largely defines the set of feasible dynamic scheduling decisions; but dynamic scheduling decisions provide valuable information about the quality of the original schedule. If profits improve by frequently and consistently re-timing or re-fleeting particular flight legs, it might be possible to modify the original schedule to reduce the need to re-time and re-fleet these flight legs. A challenging research question is to assess the potential for improving the original schedule, using a feedback loop and simulations with our dynamic and robust scheduling approaches.

THIS PAGE INTENTIONALLY LEFT BLANK

Bibliography

Ageeva, Y. *Approaches to Incorporating Robustness into Airline Scheduling*. Master's thesis, Massachusetts Institute of Technology, 2000.

Bandet, P. O. *Armada-Hub: An Adaption of AF Fleet Assignment Model to Take into Account the Hub Structure in CDG*. In *Proceedings of the 34th AGIFORS Annual Symposium*. 1994.

Barnhart, C., Kniker, T. S., and Lohatepanont, M. *Itinerary-Based Airline Fleet Assignment*. *Transportation Science*, 36(2):pp. 199–217, 2002.

Berge, M. *Timetable Optimization: Formulation, Solution Approaches, and Computational Issues*. In *Proceedings of the 34th AGIFORS Annual Symposium*. 1994.

Berge, M. E. and Hopperstad, C. A. *Demand Driven Dispatch: A Method for Dynamic Aircraft Capacity Assignment, Models and Algorithms*. *Operations Research*, 41(1):pp. 153–168, 1993. Special Issue on Stochastic and Dynamic Models in Transportation.

Bish, E. K., Suwandechochai, R., and Bish, D. R. *Strategies for Managing the Flexible Capacity in the Airline Industry*. *Naval Research Logistics*, 51:pp. 654–684, 2004.

Chan, Y. *Route Network Improvement in Air Transportation Schedule Planning*. Technical Report R72-3, M.I.T. Flight Transportation Laboratory, 1972.

Desaulniers, G., Desrosiers, J., Rioux, B., Solomon, M., and Soumis, F. *Daily Aircraft Routing and Scheduling*. *Management Science*, 43:pp. 841–854, 1997.

Dobson, G. and Lederer, P. J. *Airline Scheduling and Routing in a Hub-and-Spoke System. Transportation Science*, 27:pp. 281–297, 1993.

Donoghue, J. A. *Hub Machine. Air Transport World*, 39(6):p. 5, 2002.

Dorinson, D. M. *The Evolution of Airline Distribution Channels and Their Effects on Revenue Management Performance.* Master's thesis, Massachusetts Institute of Technology, 2004.

Dror, M., Trudeau, P., and Ladany, S. P. *Network Models for Seat Allocation on Flights. Transportation Research*, 22B:pp. 239–250, 1988.

Etschmaier, M. M. and Mathaisel, D. F. X. *Aircraft Scheduling - The State of the Art.* In *Proceedings of the 24th AGIFORS Annual Symposium*, pp. 181–225. 1984.

Farkas, A. *The Influence of Network Effects and Yield Management on Airline Fleet Assignment Decisions.* Ph.D. thesis, Massachusetts Institute of Technology, 1995.

Feldman, J. M. *IT Systems Start to Converge. Air Transport World*, 37(9):pp. 78–81, 2000.

Flint, P. *No Peaking. Air Transport World*, 39(11):pp. 22–27, 2002.

Flottau, J. *Lufthansa Reorganizes Flight Operations, 'Depeaks' Frankfurt. Aviation Daily*, p. 3, 2003.

Frank, M., Mederer, M., Hatzack, W., and Friedemann, M. *Dynamic Capacity in Airline Revenue Management.* In *Proceedings of the 45th AGIFORS Annual Symposium.* 2005.

Gallego, G. and Phillips, R. *Revenue Management of Flexible Products. Manufacturing & Service Operations Management*, 6(4):pp. 1–17, 2004.

Gillie, J. *More use Net to Book Flights. theNewsTribune.com*, 2006. URL `http://www.thenewstribune.com/business/story/5425105p-4900433c.html`.

Glover, F., Glover, R., Lorenzo, J., and McMillan, C. *The Passenger-Mix Problem in the Scheduled Airlines. Interfaces*, 12(3):pp. 73–79, 1982.

Goodstein, J. *Re-Fleeting Applications at United Airlines.* In *Proceedings of the 37th AGIFORS Annual Symposium*, pp. 75–88. 1997.

Hane, C., Barnhart, C., Johnson, E., Marsten, R., Nemhauser, G., and G., S. *The Fleet Assignment Problem: Solving a Large-Scale Integer Program. Mathematical Programming*, 70:pp. 211–232, 1995.

Hirschman, D. *How Hub 'De-Peaking' Would Work. The Atlanta (GA) Journal-Constitution*, 2004.

ILOG. *ILOG CPLEX 9.0 User's Manual*, 2003.

Jarrah, A. I. *An Efficient Airline Re-Fleeting Model for the Incremental Modification of Planned Fleet Assignment. Transportation Science*, 34(4):pp. 349–363, 2000.

Jarrah, A. I., Yu, G., Krishnamurthy, N., and Rakshit, A. *A Decision Support Framework for Airline Cancellations and Delays. Transportation Science*, 27:pp. 266–280, 1993.

Kang, L. S. and Clarke, J.-P. *Degradable Airline Scheduling.* Technical report, Massachusetts Institute of Technology, 2003.

Karow, M. *Virtual Hubs: An Airline Schedule Recovery Concept and Model.* Master's thesis, Massachusetts Insitute of Technology, 2003.

Klabjan, D., Johnson, E. L., Nemhauser, G. L., Gelman, E., and Ramaswamy, S. *Airline Crew Scheduling with Time Windows and Plane-Count Constraints. Transportation Science*, 36(3):pp. 337–348, 2002.

Kniker, T. *Itinerary-Based Fleet Assignment.* Ph.D. thesis, Massachusetts Institute of Technology, 1998.

Lan, S., Clarke, J.-P., and Barnhart, C. *Planning for Robust Airline Operations: Optimizing Aircraft Routings and Flight Departure Times to Minimize Passenger Disruptions. Transportation Science*, 40(1):pp. 15–28, 2006.

Levin, A. *Scheduling and Fleet Routing Models for Transportation Systems. Transportation Science*, 5:pp. 232–255, 1971.

Listes, O. and Dekker, R. *A Scenario Aggregation-Based Approach for Determining a Robust Airline Fleet Composition for Dynamic Capacity Allocation. Transportation Science*, 39(3):pp. 367–382, 2005.

Lohatepanont, M. and Barnhart, C. *Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment. Transportation Science*, 38(1):pp. 19–32, 2004.

Marsten, R. E., Subramanian, r., and Gibbons, L. *Junior Analyst Extraordinare (JANE). AGIFORS Proceedings*, 1996.

McDonald, M. *Endangered Species?. Air Transport World*, 39(6):pp. 34–38, 2002.

Nikulainen, M. and Oy, K. *A Simple Mathematical Method to Define Demand for Schedule Planning. AGIFORS Proceedings*, 1992.

Ott, J. *'De-Peaking' American Hubs Provides Network Benefit. Aviation Week & Space Technology*, 2003.

Peterson, R. M. *The Penultimate Hub Airplane. Internal Memo, Boeing Commercial Airplane Group, Seattle, WA*, 1986.

Phillips, B. L., Boyd, D. W., and Grossman, T. A. *An Algorithm for Calculating Consistent Itinerary Flows. Transportation Science*, 25:pp. 225–239, 1991.

Rexing, B., Barnhart, C., and Kniker, T. *Airline Fleet Assignment with Time Windows. Transportation Science*, 34(1):pp. 1–20, 2000.

Rosenberger, J. M., Johnson, E. L., and Nemhauser, G. L. *A Robust Fleet-Assignment Model with Hub Isolation and Short Cycles. Transportation Science*, 38(3):pp. 357–368, 2004.

Rosenberger, J. M., Schaefer, A. J., Goldsman, D., and Johnson, E. L. *A Stochastic Model of Airline Operations. Transportation Science*, 36:pp. 357–377, 2002.

Schaefer, A. J., Johnson, E. L., Kleywegt, A. J., and Nemhauser, G. L. *Airline crew scheduling under uncertainty. Transportation Science*, 39(3):pp. 340–348, 2005.

Shebalov, S. and Klabjan, D. *Robust Airline Crew Pairing: Move-up Crews*. Technical report, University of Illinois at Urbana-Champaign, 2004.

Sherali, H. D., Bish, E. K., and Zhu, X. *Polyhedral Analysis and Algorithms for a Demand-Driven Refleeting Model for Aircraft Assignment. Transportation Science*, 39(3):p. 349, 2005.

Simpson, R. W. *Computerized Schedule Construction for an Airline Transportation System*. Technical report, M.I.T. Flight Transportation Laboratory, 1966.

Simpson, R. W. *Scheduling and Routing Models for Airline Systems*. Technical report, M.I.T. Flight Transportation Laboratory, 1969.

Singh, M. *Hub-Depeaking Techniques at United Airlines*. In *AGIFORS Schedluling & Strategic Planning Conference*. 2006.

SL. *American To Save $ 4.5 Million By Switching Terminals at DFW. Aviation Daily*, 350(17):p. 3, 2002.

Smith, B. *Robust Airline Fleet Assignment*. Ph.D. thesis, Georgia Institute of Technology, 2004.

Soumis, F., Ferland, J., and Rousseau, J.-M. *MAPUM: A Model for Assigning Passengers to a Flight Schedule. Transportation Research, Part A*, 15A:pp. 155–161, 1981.

Soumis, F., Ferland, J. A., and Rousseau, J.-M. *A Model for Large Scale Aircraft Routing and Scheduling Problems. Transportation Research*, 14B:pp. 191–201, 1980.

Talluri, K. *Airline Revenue Management with Passenger Routing Control: A New Model with Solution Approaches. International Journal of Services Technology and Management*, 2(1/2):pp. 102–115, 2001.

Talluri, K. T. *Swapping Applications in a Daily Airline Fleet Assignment. Transportation Science*, 30(3):pp. 237–248, 1996.

Thengvall, B. G., Bar, J. F., and Yu, G. *Balancing User Preferences for Aircraft Schedule Recovery During Irregular Operations. IIE Transactions*, 32:pp. 181–193, 2000.

UAL. *UAL Corporation Reports Fourth Quarter And Full Year 2005 Results*. 2006. URL `http://www.united.com/press/detail/0,6862,53614-1,00.html`.

Wells, A. T. and Wensveen, J. G. *Air Transportation: A Management Perspective*. Brooks/Cole, 5th edition, 2004. ISBN 0-534-39384-5.

Yan, S. and Tseng, C.-H. *A Passenger Demand Model For Airline Flight Scheduling and Fleet Routing. Computers and Operations Research*, 29(11):pp. 1559–1581, 2002.

Yen, J. W. and Birge, J. R. *A Stochastic Programming Approach to the Airline Crew Scheduling Problem. Transportation Science*, 40(1):pp. 3–14, 2006.

Yu, G. and Luo, S. *On the Airline Schedule Perturbation Problem Caused by the Ground Delay Program. Transportation Science*, 31:pp. 298–311, 1997.