Planning for Robust Airline Operations: Optimizing Aircraft Routings and Flight Departure Times to Achieve Minimum Passenger Disruptions

by

Shan Lan

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

June 2003

Signature of Author ....
Department of Civil and Environmental Engineering
9 May 2003

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

Certified by ....
John-Paul Clarke
Associate Professor, Aeronautics and Astronautics Engineering
Thesis Supervisor

Accepted by ....
Oral Buyukozturk
Chairman, Department Committee on Graduate Studies

Planning for Robust Airline Operations: Optimizing Aircraft Routings and Flight Departure Times to Achieve Minimum Passenger Disruptions

by

Shan Lan

Submitted to the Department of Civil and Environmental Engineering
on 9 May 2003, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Transportation

Abstract

Airlines typically construct schedule plans based on the assumption that every flight leg departs and arrives as planned. Because this optimistic scenario rarely occurs, these plans are frequently disrupted and airlines often incur significant costs in addition to the originally planned costs. Flight delays and schedule disruptions also cause passenger delays and disruptions, and disrupted passengers experience very long delays and contribute to a significant amont of the total passenger delay. A more robust plan can alleviate flight and passenger delays and disruptions and their effects in the operation, and eventually reduce the operation costs. In this dissertation, we first define various robustness criteria in the context of airline schedule planning. Then we present two new approaches for robust airline schedule planning to achieve minimum passenger disruptions: Robust Aircraft Maintenance Routing, and Flight Schedule Retiming.

Because each airplane usually flies a sequence of flights, delay of one flight might propagate along the aircraft route to downstream flights and cause further delays and disruptions. We propose a new approach to reduce delay propagations by intelligently routing aircraft. We formulate this problem as a mixed integer programming problem with stochastically generated inputs. An algorithmic solution approach is presented. Computational results obtained by using data from a major U.S. airline show that this approach could reduce delay propagations significantly, thus improving on-time performance and reducing passenger disruptions.

Passengers miss their connections if there is not enough time for them to connect. These passengers experience very long delays. We develop a new approach to minimize the number of passenger misconnections by re-timing the departure times of flights. within a small time window. We generate copies for each flight arc in the flight network and let the model pick the set of flight copies that minimizes the number of disrupted passengers. We show various ways to formulate the problem and study the properties of these models. An algorithmic solution approach is presented. Computational results obtained by using data from a major U.S. airline show that this approach could significantly reduce the number of passenger misconnections.

Thesis Supervisor: Cynthia Barnhart
Title: Professor, Civil and Environmental Engineering

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

Contents

List of Figures

List of Tables

Acknowledgments

I would like to express my greatest gratitude to my advisors, Professor Cynthia Barnhart and Professor John-Paul Clarke, for their advise, supervision, and support. Their insightful ideas, knowledge, and experience in the field contributed in large measures to the success of this dissertation. They have been understanding mentors and trusted friends. I feel highly honored for working with them and value this experience very much.

Deep appreciation goes to Professor Georgia Perakis for serving on my committee, her encouragements, and insightful questions and suggestions. I would also like to thank Professor Ismail Chabini, Professor Moshe Ben-Akiva, Professor Nigel Wilson and other professors at MIT, particularly those at the Center for Transportation and Logistics, for their help, support and guidance.

I am grateful to the Zakhartchenko Fellowship and the Center for Transportation and Logistics at MIT for providing me financial support.

Many thanks are due to my colleagues at MIT: Stephane Bratu, Amy Mainville Cohn, Amr Farahat, Laura Kang, Manoj Lohatepanont, Su Shen, Peng Sun and many others in Professor Clarke's student group. Their contributions to my research are greatly appreciated. Many thanks are also to my officemates at room 5-012 and fellow students at CTL: John Bottom, Xiaojing Dong, Jenny farvar, Song Gao, Hai Jiang, Hong Jin, Scott Ramming, Tomer Toledo, Joan Walker, Vinay Yadappanavar, Maya Geryes Abou Zeid and many others, for their friendship. Staffs at the Center for Transportation and Logistics and the Department of Civil Engineering, particularly, Maria Marangiello, Cynthia Stewart, Sydney Miller, and Ginny Siggia have been very helpful and supportive. I thank them very much.

Finally, I would like to express my deepest gratitude to my family, my parents for their love and support for all these years, my wife, Jiang Zhu, for her understanding, patience and support, and my daughter, Vicky Lan, for the happy time she brings to me.

Chapter 1

Introduction

1.1 The Airline Schedule Planning Process

Airline schedule planning is the process of generating a schedule that has the most revenue potential and resolves a host of related issues involving fleet assignment, aircraft maintenance routing and crew scheduling. Airline schedule planning has been extensively studied in the past decade and numerous models and algorithmic approaches have been developed. Barnhart and Talluri (1997) [12], Lohatepanont (2001) [56] and Cohn and Barnhart (2003) [30] have presented structural overviews of this planning process and detailed literature reviews. Here, we give a brief introduction.

The airline schedule planning problem is huge, complex and impractical to solve directly. This problem is therefore often divided into sequential subproblems and each one is solved sequentially. Airline schedule planning consists of four stages: schedule generation, fleet assignment, maintenance routing and crew scheduling (See Figure 1-1).

Schedule Generation

The schedule generation problem determines markets, frequencies and specific times to fly. The schedule affects every operational decision and has the biggest impact on an airline's profitability. The schedule planning step typically begins 12 months before the schedule goes into operation and lasts approximately 9 months (Lohatepanont (2001)

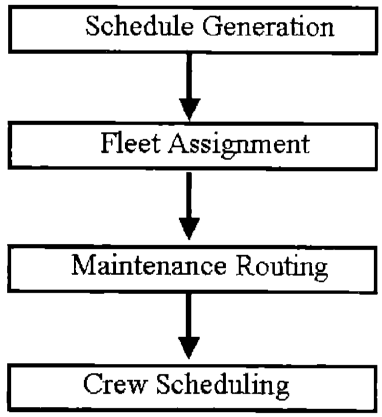

Figure 1-1: Airline Schedule Planning Process

[56]). The schedule generation step begins with route development, in which the airline decides which markets it wants to serve based on forecasted demand information. In this process, airlines try to keep serving their profit-making markets and look for new opportunities to expand. Then airlines need to determine the appropriate service frequency in these markets. The objective here is to match the frequency to the forecasted demand. Finally, airlines need to determine at what time these flights should be offered.

In the literature, optimization models are limited in the area of schedule generation. This is mainly because that schedule generation involves strategic decisions of an airline and its competitors, which is hard to be captured in a mathematical model. In addition, there are an infinite number of feasible schedules, and it is very difficult to determine the cost of each schedule, because it depends on the solutions to the subproblems. Despite these difficulties, Teodorovic and Krcmar-Nozic (1989) [76] present a methodology that determines optimal flight frequencies on a network maximizing total profit and market share and minimizing the total schedule delay of all passengers on the network. Berge (1994) [16] presents a sub-timetable optimization approach wherc a small part of the network is optimized and augmented to the master timetable. However, most of the research in this area focus on improving an existing schedule. Rexing et al. (2000) [65] and Klabjan et al. (1999) [50] present models that allow the flight times to vary within a given time window. Lohatepanont and Barnhart (2001) [57] present an integrated model for schedule design and fleet assignment that considers an extended schedule with optional flight legs and determines which flights to keep and which to discard. Using the same idea of moving flight departure times in a small time window, we develop airline operation models to minimize the number of passengers who miss their connections in the operation. The details of these models are reported in Chapter 4.

Fleet Assignment

The fleet assignment problem assigns a specific aircraft type to each flight in the schedule and tries to match the seat capacity of aircraft to the demand for each flight. The goal of this problem is to minimize operating expenses and lost revenue caused by insufficient capacity. This assignment has to satisfy some constraints such as balance of aircraft in the network and restrictions on the number of available aircraft of each type.

Much research has been done in this area. Daskin and Panayotopoulos (1989) [33], Abara (1989) [1], Hane et al. (1995) [42], Gu et al. (1994) [41], Clarke et al. (1996) [26], Rexing et al. (2000) [65], Barnhart, Kniker and Lohatepanont (2001) [10], Barnhart, Farahat, and Lohatepanout (2001) [7] and many others study this problem from various aspects. Hane et al. (1995) [42] present a multi-commodity flow model with side constraints to solve the fleet assignment problem. They develop various ways to reduce the problem size and improve the solution time, which make it possible to solve realistic problems with thousands of flights in a reasonable time. Gu et al. (1994) [41] study the complexity and behavior of the model presented in the paper by Hane et al. (1995) [42]. Rexing et al. (2000) [65] present a model to achieve minimum fleet assignment costs by allowing flight departure times to move in a small time window. Barnhart , Kniker and

Lohatepanont (2001) [10] present an itinerary-based model determining fleet assignment costs more accurately than the model by Hane et al. (1995) [42]. Barnhart, Farahat, and Lohatepanout (2001) [7] present a composite-variable based fleet assignment model that can model fleet assignment costs as accurately as the itinerary-based model but is easier to solve. The fleet assignment models have been widely applied in practice, and significant savings have been achieved. For example, a 100 million dollars per year savings in operating costs at Delta Airlines has been reported by Subramanian et al. (1994) [73].

Maintenance Routing

The maintenance routing problem constructs a set of routes, one for each aircraft, to ensure that all aircraft are maintained at the right place and right time. US Federal Aviation Association (FAA) regulations require that airlines perform periodic maintenance checks on their aircraft after a certain number of hours of flying. These requirements are strictly enforced and an airplane will be grounded if it does not meet them. The objective in this step is to find maintenance feasible routes for airplanes, given a fleeted schedule and the number of available aircraft of each fleet type. We discuss this problem in detail in Section 3.1.

Crew Scheduling

The crew scheduling problem assigns cockpit and cabin crews to flights to achieve minimum cost. Crew scheduling is often done in two phases. First, the *crew pairing* problem is solved to construct sequences of flights (crew pairings) in such a way that each flight is included in exactly one pairing. Second, the *crew assignment* problem is solved to assign specific crews to the pairings to create schedules. These schedules must satisfy some requirements. For instance, pilots are qualified to fly only certain types of aircraft; flight crews cannot be away from their base or stay on duty for a time period longer than the respective limits. The focus of crew scheduling research has historically been the crew pairing problem, in part because of tractability issues associated with the crew assignment problem (Barnhart and Talluri (1997) [12]). Because the crew pairing structure is very complex and the cost function is nonlinear, almost all of the existing models formulate the crew pairing problem as a set partitioning problem with one binary decision variable for each pairing. This formulation eliminates the need to formulate explicitly the complex pairing structure and allows linearization of the cost function.

Because the number of variables is huge, sometimes exceeding hundreds of millions, most crew pairing research has focused on solution techniques for solving such large-scale integer programs. Some recent work includes Anbil et al. (1992) [5] , Barnhart et al. (1994) [8], Barnhart and Shenoi (1998) [11], Beasley and Cao (1996) [13], Chu et al. (1997) [25], Desaulniers et al. (1997) [34] , Hoffman and Padberg (1993) [43], Klabjan et al. (1999) [49], Klabjan and Schwan (1999) [51], Vance et al. (1997) [79].

Once pairings are generated, they are combined with rest periods, vacations, training time, etc. to create extended work schedules that can be performed by an individual. The objective of the crew assignment problem is to find a minimum cost assignment of employees to these work schedules. There are two approaches for crew assignment: rostering and bidline generation. Rostering is a common practice in Europe by which schedules are constructed for specific individuals. Bidline generation is a common practice in North America. In this case, the cost-minimizing subset of schedules is selected without individual specific action. Employees reveal their relative preferences for these schedules through a bidding process. The airline then assigns schedules to employees based on individual priority rankings.

Many of the research results on airline schedule planning have been applied in the airline industry and have improved airlines' performances. This notwithstanding, almost all optimization models in this area have assumed that flights, crews, and passengers will operate as planned. Thus, airlines typically construct plans that maximize revenue or minimize cost based on the assumption that every flight leg departs and arrives as planned. Because this optimistic scenario rarely occurs, these plans are frequently disrupted and airlines often incur significant costs in addition to the originally planned cost. Currently, the "optimal" planned schedules generated by US airlines' schedule planning systems are far from optimal in operations. It is estimated that the financial impact of irregularities on the daily operations of a single major U.S. domestic carrier may exceed \$440 million per annum in lost revenue, crew overtime pay, and passenger hospitality costs (Clarke and Smith (1999) [28]). The cost of delays and disruptions is not only significant, but also rapidly increasing. The Air Transport Association estimates that delays cost consumers and airlines about \$5.2 billion in 1999 and \$6.5 billion in 2000. (Air Transport Association website (2003) [3]).

1.2 Delay, Cancellation and Disruption

There are many reasons that can cause flight delays and cancellations, for instance, severe weather conditions, unexpected aircraft and personnel failures, and congestion at the airport and in the airspace. In the year 2000, about 30% of the flights were delayed, and about 3.5% of the flights were cancelled . Since the schedule planning system used by airlines that creates "optimal" schedules does not attempt to manage possible delays and cancellations, the delays and cancellations often cause disruptions to airline schedules, sometimes with significant negative effects.

Flight delays and cancellations not only lead to aircraft and crew schedule disruptions but also cause passengers to be disrupted from their original itinerary. Passengers are disrupted if their planned itineraries become infeasible because one of the flights in the planned itinerary is cancelled or there is insufficient time for them to connect between two flights. In 2000, it is estimated that about 4% of passengers were disrupted, among which about half of them are connecting passengers (Bratu and Barnhart (2002) [21]). The impacts of passenger disruptions are tremendous (Bratu and Barnhart (2002) [21]). First, disrupted passengers incur very long delays: in one case study, the average delay for disrupted passengers is estimated to be about 419 minutes, while the average delay for non-disrupted passengers is 14 minutes. Second, passenger disruptions cause huge direct revenue losses. Associated revenue losses include delay cost for passengers, airline revenue loss due to passengers being served by other airlines, and overnight passenger costs. Third, there are some other significant potential losses, such as loss of goodwill.

In recent years (prior to September 11, 2001), flight delays and cancellations increased significantly in the U.S.. In 2000, 30% of the flights were delayed, a 100% increase compared to 1995, and about 140,000 flights were cancelled, a 500% increase compared to 1995 (Bratu and Barnhart (2002) [21]). As staggering as these numbers are, it is estimated that flight delays and cancellations might increase dramatically in the future: air traffic in US is expected to double in the next 10-15 years, and each 1% increase in air traffic will bring about a 5% increase in delays (MIT Global Industry Program [61] and Schaefer et al. (2001) [70]). This will lead to much more frequent and serious schedule disruptions and tremendous revenue loss unless airline schedule planning and operations are significantly improved. This has motivated our research in robust airline schedule planning.

1.3 Contributions and Outline

1.3.1 Contributions

Our contributions in this dissertation are summarized as follows:

1. We provide alternative definitions of robustness in the context of airline schedule planning. These definitions can be applied to the overall airline schedule planning process, not just to our two approaches.

2. We propose a new approach to generate aircraft maintenance routes minimizing delay propagation. We formulate the problem as a mixed integer program, and develop an algorithmic approach to solve it. We investigate the value of our robust plan over the plan generated by conventional approaches using data from a major U.S. airline. The results show that our approach could reduce delay propagation significantly, improve on-time performance and reduce number of passengers missing their connections.

3. We propose a new approach to reduce passenger misconnections. We develop optimization models to minimize the expected total number of passenger misconnections by moving flight departure times within a small time window. We analyze the model properties and develop an algorithmic approach. The computational results obtained by using data from a major U.S. airline show that this approach, which has desirable computational properties, can reduce the number of disrupted passengers significantly.

4. We present models that integrate our models with other steps in the airline schedule planning process therefore gaining more robustness in schedule plans.

1.3.2 Thesis Outline

In Chapter 2, we first survey some general robust planning methodology and some methods proposed by researchers to generate robust airline schedule plans. We then provide alternative definitions of robustness in the context of airline schedule planning. Finally we present a modeling framework for robust airline schedule planning. In Chapter 3, we present a robust aircraft maintenance routing model and its associated solution approach. By routing aircraft in different ways, we can reduce the delay propagating through out the network. We also present and analyze proof-of-concept results using data from a major U.S. airline. In Chapter 4, we present the idea of rescheduling each flight within a small time window to achieve minimum passenger disruptions. We show various ways to model this problem and analyze the properties of the models. We also present and analyze proof-of-concept results using data from a major U.S. airline. In Chapter 5, we discuss possible extensions of our models to obtain more robust airline schedules. Finally, in Chapter 6, we summarize our contributions and describe directions for future research.

Chapter 2

Robust Airline Schedule Planning

2.1 Introduction

There are two ways to deal with schedule disruptions. One is to re-optimize the schedule after the schedule disruptions occur, which is done in the operations stage. Another approach to managing schedule disruptions is to build robustness into the planned schedules, which must be done in the planning stage.

2.1.1 Schedule Recovery

When disruptions occur, the airlines must reschedule flight operations. An airline's recovery policy determines which flight departures to postpone and cancel and how to reroute the aircraft, pilots and passengers. Although very few airlines use automated recovery policies, lots of research has been done in this area. Airlines typically recover from disruptions in stages (Rosenberger et al. (2001) [66]). The first stage is to recover aircraft by rerouting aircraft and delaying and/or cancelling flight legs. The second stage is to recover crew by reassigning pilots and cabin crew and calling upon reserve pilots and cabin crew. The third stage is to recover passengers. Related literature includes Cao and Kanafani (1997) [23], Jarrah et al. (1993) [44], Lettovsky (1997) [54], Luo and

Yu (1997) [58], Mathaisel (1996) [60], Teodorovic and Guberinic (1984) [75], Teodorovic and Stojkovic (1995) [77], Thengvall et al. (2000) [78], Yan and Tu (1997) [82], Yan and Yang (1996) [83] and Yu et al. (2003) [85]. Interested readers are referred to Clarke and Smith (2000) [29] and Rosenberger et al. (2001) [66] for a detailed review.

2.1.2 Robust Planning

Building robust into schedule is a proactive way to deal with the schedule disruptions. A more robust plan can alleviate the effects of disruptions on the operations, and hence reduce operations costs. Hence, building robust schedule plans is as important as re-optimizing disrupted schedule. Although the literature for schedule recovery is prolific, research on robust airline schedule planning is very limited (See Section 2.2), partly because of some challenges of modeling this problem as described next.

In this dissertation we will focus on building robustness into the schedule plans. We hope to build schedule plans that are "robust" by considering explicitly possible delays and cancellations in our planning model. These robust schedule plans should be less sensitive to delays and cancellations or more repairable than those generated by conventional approaches, and therefore reduce the impact of schedule disruptions during operations. Robust plans might not be optimal for "planned" operations, but should minimize overall realized costs.

Modeling Robust Airline Schedule Planning: Challenges

The robust airline schedule planning problem is very challenging. First, robustness is difficult to define with definition of robustness often being problem specific. Depending on the particular problem instance, a robust plan might be a plan that yields the minimum cost for the worst case, or minimum expected cost, or a plan that minimizes costs and satisfies a certain level of service requirements, or a plan that achieves other objectives. There hasn't been a systematic way in the literature to define robustness in the context of airline schedule planning.

Second, the flows of aircraft, crew, passengers, and other resources interact over the hub-and-spoke networks operated by most major airlines in the U.S.. Hub-and-spoke networks present major operational challenges. When severe weather conditions reduce a hub airport capacity, departing aircraft queues start to form and delays grow exponentially and propagate throughout the operation, affecting aircraft, crew and passengers.

Third, optimization models capturing stochasticity are often computationally intractable for airline problems due to their large size. In fact, deterministic optimization models for airline schedule planning are challenging to solve. The additional complexity and problem size when stochasticity is considered leads to issues of tractability.

Last, it is difficult to balance robustness and costs when modeling these problems. Conventional models for airline schedule planning minimize planned costs, while airlines' ultimate goal is to minimize operation costs, which can be viewed as the sum of the planned costs and the costs of delays and disruptions. Adding robustness to schedules will reduce costs for delays and disruptions and hopefully lead to reductions in operation costs. Schedule plans with added robustness might result in higher planned costs, but it is hoped that the benefit of robust plans should exceed the increased planned costs. The value of robustness is, however, hard to quantify and thus it is difficult for airlines to determine how much they should pay to achieve certain levels of robustness. Based on the above analysis, we can conclude that creative ideas are needed to develop robust airline schedule planning methods.

Robust Aircraft Maintenance Routing

Given these inherent difficulties, we adopt the strategy of finding the most robust solution without significantly added cost. Our first model focuses on the maintenance routing problem for the following reason. In most optimization models for the aircraft maintenance routing problem, the objective is to maximize *through revenue*, the potential revenue obtained by offering passengers the opportunity to stay on the same aircraft rather than make a connection at an airport. In practice, this additional revenue is very difficult to determine accurately and the financial impact is relatively small (Cordeau et al. (2000) [32], Klabjan et al. (1999) [50]). The aircraft maintenance routing problem can thus be cast as a feasibility problem. This gives us the opportunity to achieve robustness through optimizing some appropriate objective. The challenge is what robustness can achieve for the maintenance routing problem.

Because each airplane flies a sequence of flights, arrival delay of one flight may cause departure delay for the next flight flown by the same aircraft if there is not enough slack between these two flights. This phenomenon is called *delay propagation*. Delay propagation often causes delays for downstream flights, and delays and disruptions for crews and passengers on these flights. Especially at hubs where airplanes, crews and passengers flows are interrelated, schedules are very sensitive to delays. Thus, preventing flight delay propagation might help to reduce delays and disruptions for downstream flights, passengers and crews. This motivates us to look for methods that can reduce delay propagation. Because delays propagate along aircraft routes, delay propagation can be reduced if aircraft routes are selected intelligently. In Chapter 3, we present a method to select aircraft routes that satisfy maintenance requirements and minimize the expected total propagated delay. We recognize this is only a step towards building robust schedules, but it is a step that could significantly reduce delay propagation, improve on-time performance and reduce the number of passengers missing their connections.

Flight Schedule Retiming

As described in Section 1.2, every year large numbers of passengers are disrupted either because their flights are cancelled or because they do not have enough time to connect to the next flight in their itinerary. This causes significant losses to airlines and passengers and motivates us to look for methods that can reduce passenger disruptions. In fact, if "slack" time is inserted into the schedule appropriately, many passengers might be saved from missing their connections. To do this, we associate with all flights small time window and solve an optimization problem to determine the flight departure times to minimize the number of missed connections.

2.2 Literature for Robust Planning

Although robust airline schedule planning is a relatively new concept, "robust planning" has been studied by many researchers and applied in various fields such as robot design, manufacturing, supply chain management and logistics, telecommunications, economics, ecology, water and environmental management, and portfolio management in finance (For detailed reviews, readers are referred to [72], [18], [53], [86], [45], [80]). The methodologies used include stochastic programming (Birge and Louveaux (1997) [19]), scenario planning (Kouvelis and Yu (1997) [53], Mulvey et al. (1995) [62]), and fuzzy optimization (Zimmermann (1991) [86] and Sakawa (1993) [69]).

2.2.1 General Robust Planning Methodology

Stochastic Programs are mathematical programs for which some of the data incorporated in the objective or constraints is uncertain. Uncertainty is usually characterized by probability distributions of the parameters. The goal is to find a feasible solution for all (or almost all) the possible data instances, and to maximize or minimize the expectation of some function of the decisions and the random variables. The most widely applied and studied stochastic programming models are called *two-stage models.* The decision maker takes some action in the first stage, after which a random event occurs affecting the outcome of the first-stage decision. A recourse decision can then be made in the second stage that minimizes the expected costs of the consequences of that decision. One natural generalization of the two-stage model is the *multi-stage models.* The decision maker makes one decision now, waits for some uncertainty to be realized, and then makes another decision based on what has happened. The objective is to minimize the expected costs of all decisions taken ([72]).

Stochastic optimization sometimes fail to meet decision making needs in an environment characterized by significant uncertainty. First of all, the stochastic programming models are often computationally intractable when applied to practical problems, duc to their large size. Another shortcoming of the stochastic programming approach is that it requires the decision makers to determine the probability distributions of random variables. This is not a trivial exercise in many cases, particularly when the decision environments have multiple interdependent uncertain factors.

Mulvey et al. (1995) [62] adopt a scenario based approach to model robustness under two dimensions. The first dimension is *solution robust*: a solution is robust with respect to optimality if it remains close to optimal for any input data scenario to the model. The second one is *model robust*: a solution is robust with respect to feasibility if it remains almost feasible for any scenario realization. There are two distinct components in their model: a structural component that is fixed and free of any noise in its input data, and a control component that is subjected to noisy input data. This model allows the introduction of higher moments of distribution of random variables and uses a feasibility penalty function to penalize violations of the control constraints under some of the scenarios.

Kouvelis and Yu (1997) [53] describe and formulate robustness for problems where associated probabilities of realizations of uncertainties are unknown. They use *min-max regret* rules to minimize the worst-case costs among all possible realizations of uncertainties. They use a scenario-based approach to represent the input data uncertainty. A specific input data scenario represents a potential realization, which occurs with some positive but perhaps unknown probability, of the important parameters.

To deal with imprecision quantitatively, concepts and techniques of probability theory are often employed. Bellman and Zadeh (1970) [14] believe that this assumes imprecision can be equated with randomness, which they think is a questionable assumption. To differentiate between *randomness* and *fuzziness*, they develop the concepts of fuzzy goals, fuzzy constraints, and fuzzy decisions. Since then, much research has been done in this area, and the most successful application is in fuzzy control. Applying these concepts to optimization has attracted researchers, but the methodology is still far from mature for real-world applications. In fuzzy optimization, the objective can be optimized inexactly, and the constraints can be satisfied to varying degrees. Most solution approaches reported in the literature transform a fuzzy problem into a general (so-called "*crisp*") problem. The basic idea is that the objective function should be *essentially* smaller than or equal to some "*aspiration level*", and this can be regarded as a constraint. The fuzzy optimization is defined as maximizing the minimum degree of satisfaction among all the constraints (Yan and Luh (1997) [81]).

2.2.2 Robust Airline Schedule Planning

In part because traditional optimization models for airline schedule planning assume flights will arrive and depart as planned, airlines have incurred billions of dollars of revenue losses due to the disruptions in operations. Recognizing these, Yen and Birge (2001) [84] present a stochastic programming model for crew scheduling, and solve only a small problem instance. Ageeva (2000) [2] presents a robust aircraft maintenance routing model to provide opportunities to swap planes. The basic idea is to generate routes so that some routes will *meet*, that is, have a station in common for a given length of time, twice. This allows a controller to swap the routing of these airplanes during the period of disruptions and later swap them back to their original routes. This is helpful for airlines to recover from flight delays.

Chebalov and Klabjan (2002) [24] propose a similar idea for crew scheduling. They generate crew schedules with many opportunities for crew swapping, providing more opportunities for airlines to recover disrupted crews. Besides the objective of minimizing crew cost, they introduce the objective of maximizing the number of move-up crews, i.e., the crews that can potentially be swapped in operations. They define c_{opt} as the planned crew cost and r as the "robustness factor" measuring the additional crew cost airlines are willing to pay for robustness.. Then, they add a constraint limit the crew costs to $(1+r)c_{opt}$ when the maximum number of move-up crews are obtained. They develop a solution method based on Lagrangian decomposition in which they relax the constraints associated with counting the number of move-up crews, and add these constraints with penalty terms to the objective function.

Rosenberger et al. (2001) [67] realize that airline decision makers usually cancel a cycle when canceling a flight to achieve aircraft balance. Their goal then was to develop a robust fleet assignment and aircraft rotation model with many short cycles. Building a number of short cycles into the schedule makes it possible to isolate disruptions by avoiding cancellations of the large number of flights in long cycles. They define hub connectivity as the number of flight legs in a rotation beginning at a hub, ending at a different hub, and only stopping at spokes in-between. They then prove that fleet assignments with limited hub connectivity have more short cycles that begin and end at hubs. They build a string based fleet assignment model based on the idea proposed by Barnhart et al. (1998) [6]. Their model has two variants, one is to minimize fleet assignment costs with one added constraint to limit hub connectivity. Another is to minimize hub connectivity and add one constraint to constrain the fleet assignment cost. They use simulation to validate the model and show that the new fleeting and routing results perform better in operations than those generated by conventional models.

Schaefer et al. (2001) [70] propose a stochastic extension to the deterministic crew scheduling problem. By simulation, they obtained a linear approximation of the expected crew cost, then solve the resulting deterministic crew scheduling problem to minimize a lower bound on the total crew costs. In the simulation, they assume a specific crew recovery procedure called *push-back crew recovery.* In the push-back crew recovery procedure, a flight is delayed until all the resources are available. Their study is limited to "frictional disruptions", in which disruptions last for only limited duration, and do not include lengthy unscheduled maintenance disruptions and large-scale severe weather conditions.

Yen and Birge (2001) [84] develop a two-stage stochastic integer programming model to minimize total crew costs. The first stage is to determine the crew pairing using expected pairing costs. In the second stage, they consider future actions due to disruptions in the original schedule. Because a major source of delays caused by crews stems from crews who are scheduled to switch planes, they try to find out solutions in which crew plane changes are minimized. They have a nonlinear recourse model and develop a special branching algorithm that exploits problem structure to solve the problem.

Kang and Clarke (2002) [47] propose a methodology for deriving an airline schedule that is robust from a revenue perspective with reduced impact from unpredictable weather. This schedule is derived by partitioning a current airline schedule into several independent schedules that are prioritized on the basis of revenue. Disruptions are isolated in a layer, thus preventing disruptions from spreading over the network. The resulting *degradable airline schedule* provides airlines with a delay/cancellation policy and may enable airlines to segregate the market based on passenger preference for convenience and reliability. They model the problem using an integer programming model and solve it by linear approximations and column generation.

2.3 Definition of Robustness

In order to achieve robust airline schedules, robustness has to be defined in a systematic way. We propose the following definitions. A schedule plan is called robust if it

· minimizes some cost, such as the cost for the worst case among all possible realizations of uncertainties, the expected realized costs, the deviations from the optimal solutions under all realizations of uncertainties;

· minimizes aircraft/passenger/crew delays and/or disruptions; or

· is "easy" to recover aircraft/passenger/crew when disruptions occur; or

· isolates delays and schedule disruptions to avoid downstream impacts.

With the first criterion, the resulting fleet assignment results or crew schedules are robust in that overall operational fleeting costs or crew costs are minimized.

With the second criterion, the aircraft, crew or passenger delays and disruptions are minimized. The delay of one flight may cause the delays of downstream flights and cause passengers and crews to miss their connections. In this dissertation, we propose a method to reduce delay propagation without significantly increasing costs. Reducing delay propagation can improve aircraft on-time rates increase and reduce the numbers of crews and passengers delayed or disrupted. We also propose models to minimize passenger disruptions by adjusting flight departure times in a small time window. Crew delays or disruptions will also cause tremendous revenue loss, because crew cost is the second largest among all airline operating costs. In addition, one of the main reasons that cause flights to be delayed or cancelled is late arrival of crew members. Thus, avoiding crew delays and disruptions is also very important.

With the third criterion, the focus is to build schedule plans that are easier to recover from disruptions than those without robustness built in. When disruptions occur, airlines must reschedule flight operations. It would be good if the schedule plans are built in such a way that they can be easily rescheduled when disruptions occur. Looking at how Airline Operations Control Centers recover schedule plans from disruptions may help us find ways to build plans that are easy to recover. For example, airlines sometimes swap airplanes (and their routes) to help recovery from disruptions. Aircraft routes with a number of swap opportunities might be easier to recover than those without such opportunities. This is exactly what is proposed in Ageeva (2000) [2].

Finally, major U.S. airlines typically operate hub-and-spoke networks in which delays and disruptions easily propagate throughout the network. If we can build schedule plans that allow the isolation of delays and disruptions, then the spread of delays and disruptions will be reduced. For example, because long delays or cancellations of flights in one hub can seriously hurt the operation in another hub. When one flight has to be cancelled, airlines usually cancel a cycle of flights to avoid ferrying planes. If this cycle is very long, then lots of flights will be cancelled unnecessarily and other downstream flights may be impacted. Based on this observation, Rosenberger et al. (2001) [67], develop a model to generate fleet assignment and aircraft rotations with lots of short cycles.

	Min Cost	Min delays/ disruptions	Ease of recovery	Isolation of disruptions	
Schedule Design		This Thesis			Kang & Clarke
Fleet Assignment				Rosenberger, *et al.* (2001)	
Maintenance Routing		This Thesis	Ageeva & Clarke(2000)		Kang & Clarke This thesis
Crew Scheduling	Yen & Birge, Schaefer, *et al.* (2001)		Chebalov & Klabjan		

Figure 2-1: Robust Crietria and the Steps of Airline Schedule Planning

2.4 Robust Airline Schedule Planning Approaches

Based on these four alternative definitions for robustness, we propose the following mapping of airline schedule planning research to the alternative robustness definitions, as illustrated in Figure 2-1.

As can be seen from Figure 2-1, there are many open research directions to pursue. Our robust aircraft maintenance routing model combines *isolation of delay and disruption* and *maintenance routing* . The robust maintenance routing problem can be described as follows: Given a fleeted schedule, determine maintenance feasible aircraft routes that can reduce delay propagation and its impacts. The basic idea is to optimize the allocation of slack between flights using historical delay information. This will reduce delay propagation and its downstream impacts. In Chapter 3 we present our model, algorithmic solution approach, and computational results.

Because passenger disruptions occur frequently and cause tremendous revenue losses, it is important to develop methods in the planning stage to reduce the possible passenger disruptions. Our flight schedule retiming model to minimize passenger disruptions combines *minimal disruption* for passengers and *schedule generation*. Our idea is to allow a small time window for each flight and choose departure times that minimize the expected total number of disrupted passengers. Models, algorithmic solution approaches and computational results are presented in Chapter 4.

In recent years, more and more research has been done to integrate different steps of the airline schedule planning process. Our robustness definitions can be applied to each individual planning steps and also to integrated models. In Chapter 5, we present ideas on integrated robust schedule planning.

Chapter 3

Robust Aircraft Maintenance Routing

3.1 Introduction to Aircraft Maintenance Routing

3.1.1 Maintenance Requirements

The FAA mandates four main categories of airline safety checks: A-, B-, C-, and D-checks, varying in scope, duration, and frequency (Clarke et al.,1996 [26], [27]). These requirements are strict enforced and an aircraft will be grounded if any of the requirements are not met. To better illustrate the details of these safety checks, we summarize the maintenance procedure followed by American Airlines as follows (AMR Corporation (2002), [4]).

"PS" Daily Checks

Every aircraft is checked every day in its "PS" (Periodic Service) check. The aircraft is visually inspected and its maintenance log book is checked for entries and maintenance needs. The "PS" check can be performed overnight or during downtime in the flight day. It averages approximately two man-hours.

"A" Checks

The "A" check is more detailed than the "PS" check. "A" checks are performed roughly once a week (approximately 60 flight hours). The "A" check is performed at one of 40 stations around American's system. It averages 10 - 20 man-hours.

"B" Checks

The "B" check is an even more thorough maintenance check. The "B" check is done approximately once a month (roughly 300 - 500 flight hours). In addition to specific service performed on the aircraft, a detailed series of systems and operational checks are performed. American always performs "B" checks inside one of its hangars at seven different cities around its system. A "B" check requires approximately 100 man-hours on narrowbody aircraft (those with only one aisle) and approximately 200 - 300 man-hours on widebody aircraft (those with two aisles).

"C" Checks

The "C" check is the most thorough type of maintenance work performed by American. The airframe — virtually the entire aircraft — goes through an exhaustive series of checks, inspections and overhaul work. It is performed at either of American's heavy maintenance and engineering centers in Tulsa, Oklahoma or the Alliance Maintenance Facility in Fort Worth, Texas. There are different levels of "C" checks depending on the type of aircraft. These include:

Narrowbody "C" Checks

American does two types of "C" checks on its narrowbody planes. The first is a "Light C" check, which occurs approximately once a year. It requires approximately 2,100 man-hours and three days to accomplish. Every fourth "Light C" check becomes a "Heavy C" check. This check requires 20,000 - 30,000 man-hours and takes from three to five weeks to accomplish.

Widebody "C" Checks

Because of the complexity of widebody aircraft, all "C" checks are "Heavy C" checks. The complete airframe inspection and service is done every 15 - 18 months. It takes approximately 10,000 man-hours and from two to four weeks to accomplish a widebody

"C" check.

Jet Engine Overhauls

Modern jet engines are among the most reliable devices in aviation. American does not replace and overhaul jet engines at a specific number of hours. Instead, American uses a 24-hour-a-day "condition monitoring" process that scientifically tracks the condition of every engine on every aircraft. In addition to visual inspection, technicians monitor the internal condition of every engine, using such procedures as boroscope inspections and oil sample spectographs. The goal is to replace and overhaul an engine before a problem can occur. Engine overhauls are performed at the Tulsa and Alliance-Fort Worth Maintenance and Engineering facilities. The engine replacement is usually performed at one of the six "B" check hangar locations around the country.

3.1.2 Aircraft Maintenance Routing Problem

The goal of the aircraft maintenance routing problem is to determine a sequence of flights, called *routes*, to be flown by individual aircraft such that each flight is included in exactly one route, and all aircraft can be maintained as necessary. Usually, the maintenance routing problem considers only A checks. Among the four safety checks, A-checks are the only type that need to performed frequently. While the A-checkrequirement is that each aircraft be maintained after every 60 hours of flying, airlines typically enforce more stringent maintenance requirements and require an A check every 40-45 hours of flying (about three to four calendar days). Because the maintenance equipment requires large capital investment and some other constraints, these checks are done at a limited number of stations.

Some major airlines have an additional requirement that each aircraft have even wear and tear. The aircraft flying different segments of the flight schedule will encounter different conditions with respect to flight length, weather, other environmental conditions and maintenance schedules (Barnhart et al. (1998) [6]). One method by which airlines can ensure equal utilization of aircraft in a fleet over the long term is to require that every aircraft fly all the flights assigned to its fleet.

3.1.3 Literature Review

Recent work in the area of maintenance routing includes Feo and Bard (1989) [38], Kabani and Patty (1992) [46], Desaulniers et al. (1994) [35], Clarke et al. (1996) [27], Talluri (1998) [74], Gopalan and Talluri (1998) [40], Barnhart et al. (1998) [6]. These models assume that the fleeted schedule will repeat everyday and if the plane overnights at a maintenance base, it has the opportunity to undergo maintenance.

As pointed out by Gopalan and Talluri (1998) [40], it is possible to incorporate maintenance routing requirements into the fleet assignment model and formulate a joint model. The integer programming problem for such models, however, become computationally intractable, because the new constraints that have to be added to include maintenance requirements destroy the structure that makes it easy to solve the integer programming formulation for the fleet assignment problem.

Clarke et al. (1996) [26] add maintenance constraints to the fleet assignment model developed by Hane et al. (1995) [42] to ensure a sufficient number of maintenance opportunities for each fleet type. A maintenance opportunity exists when an aircraft overnights at one of its maintenance stations. While this approach ensures aircraft have enough maintenance on average, it does not guarantee that maintenance requirements will be satisfied for each individual aircraft. To do so, an aircraft maintenance routing problem has to be solved explicitly.

Aircraft maintenance routing models in the literature can be divided into two categories (Barnhart and Talluri (1997) [12]): those that use graph-theoretic heuristics and those that use mathematical programming models.

The work of Talluri (1998) [74] and Gopalan and Talluri (1998) [40] is representative of using graph-theoretic heuristics for maintenance routing. They model the problem as one that generates a *line-of-flight* graph, and finds a special Euler Tour in that directed graph. They fix all connections for each aircraft during the day to create lines of flying, specifying the origin at the start of the day and the destination at the end of the day for this aircraft. Using this idea, they present algorithms for finding 3-day and 4-day maintenance routes.

Feo and Bard (1989) [38] study the maintenance location problem which involves finding the minimum number of maintenance stations required to meet the specific 4-day A-check requirement for a proposed flight schedule. They assume that the intermediate stops during the day are not important. They formulate this problem as a min-cost, multicommodity network flow problem with integer restrictions on the variables. Because the size of the formulation is too large, they use heuristics to form maintenance routes.

Kabbani and Patty (1992) [46] study the maintenance routing problem for American Airlines, where each aircraft needs to have an A-check every three days. They formulate the problem as a set partitioning model where a column represents a possible week-long route and a row represents a flight. They develop a pseudo-cost to penalize routings with unfavorable characteristics such as violation of connection times, maintenance violations, and identification routes where aircraft are isolated, flying between a small subset of airports with no chance to receive maintenance. For some large size fleets, they separate the problem into two subproblems.

Desaulniers et al. (1997) [35] present a general modeling and algorithmic framework for fleeting, routing and scheduling problems. They try to determine a fleet schedule that maximizes profits, given a fleet of aircraft, a set of flight leg over a one-day horizon, departure time windows, and durations. They present two equivalent formulations: a multicommodity network flow formulation and a set partitioning formulation. They describe the network structure of the subproblem when a column generation technique is used to solve the LP relaxation of the set partitioning formulation. In addition, they present a Dantzig-Wolfe decomposition approach to solve the LP relaxation of the time constrained multicommodity network flow formulation. In both cases, they use a branch-and-bound algorithm to obtain integer solutions, and present optimal branching strategies compatible with these column generation approaches.

Clarke at al. (1996) [27] present a flight-based model for this problem and describe a Lagrangian relaxation solution approach that adds subtour and maintenance constraints as they are violated. They capture *through-value* in the objective function. Through-value is the revenue that would be expected to be gained from additional passengers who would be attracted to the service because of being able to stay on the same aircraft, rather than having to change airplanes at the stopover point.

Barnhart et al. (1998) [6] present a string-based model for fleet assignment and maintenance routing. A string is a sequence of connected flights that begins and ends at maintenance stations, satisfies flow balance and is maintenance feasible. The complicating maintenance constraints are easily modeled in this variable definition. Because the number of columns for this problem is huge, they exploit the network structure of the problem and present a branch-and-price solution approach to solve it. To ensure even wear-and -tear on each aircraft, they add constraints to force every aircraft to fly all the flights assigned to its fleet.

It is important to note that none of these methods considers possible delays and disruptions in the operation. Thus, while the solutions may appear optimal, in practice, they may be far from optimal.

3.2 Delay Propagation

As we described above, delay of one flight might propagate along the aircraft routes to downstream flights, which might further cause delays and disruptions of passengers and crews. Before we present our model to select aircraft routings with minimum delay propagation, we first provide some definitions.

Flight delays may be divided into two categories:

· *Propagated delay*: flight delay caused by waiting for incoming aircraft. This delay is a function of aircraft's routing. For the major U.S. airline for which we have data, the propagated delay is approximately 20-30% of the total delay.

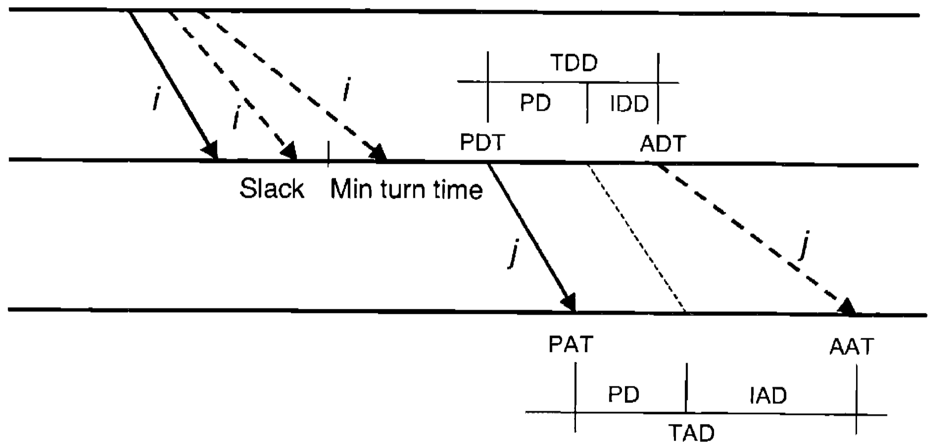

Figure 3-1: Departures, Arrivals and Delays

· *Non-propagated delay*: delay caused by all other reasons, not a function of routing. We also call this independent delay (independent of routing).

Figure 3-1 illustrates the relationship of departures, arrivals, and delays. The solid lines with arrow represent the original schedule for two flights i and j. The dotted lines with arrow represent the actual departures and arrivals of these flights. PDT refers to planned departure time, and ADT refers to actual departure time. PAT refers to planned arrival time and AAT refers to actual arrival time. The turn time is the time between the arrival of the aircraft at the gate and the time this aircraft is ready for the next flight. The minimum turn time is the least time possible to turn an aircraft. If PTT_{ij} is the planned turn time between flight i and flight j, and MTT is the minimum turn time, then the slack is the difference between planned turn time and minimum turn time, that is,

$$PTT_{ij} = PDT_j - PAT_i, \tag{3.1}$$

and

$$Slack_{ij} = PTT_{ij} - MTT. \tag{3.2}$$

TDD refers to total departure delay, comprised of independent departure delay (IDD) and propagated delay (PD). PD_{ij}, the delay propagated from flight i to flight j if both flights are flown by the same aircraft, can be determined as follows:

$$PD_{ij} = \max(TAD_i - Slack_{ij}, 0). \tag{3.3}$$

TAD, the total arrival delay, is also comprised of two parts, namely, propagated delay (PD) and independent arrival delay (IAD).

3.3 Modeling the Robust Aircraft Maintenance Routing Problem

As discussed above, delays may propagate along aircraft routes, but appropriately located slack between two consecutive flights can prevent delay propagation. One extreme case is if each flight were flown by one airplane, then there would be no delay propagation. Of course this is not possible in reality because airplanes are so expensive that airlines have only a limited number of them. To be cost effective, airlines try to fly as many flights as possible with the available fleet of aircraft. Even so, it is possible to "dampen" and reduce propagated delay and overall flight delays by intelligently routing the aircraft, allocating slack optimally to absorb the delay propagation as much as possible. This idea is to add slack where advantageous, while reducing slack where it is less needed.

Figure 3-2 illustrates the idea. Assume that flight f_1 and flight f_3 are in the same route (string) s_1, and flight f_2 and flight f_4 are in the same route (string) s_2. According to historical data, assume that we know flight f_1 is delayed, as shown in the figure, on average to the position of f_1'. This delay is longer than the slack between flight f_1 and flight f_3, causing delay to propagate from flight f_1 to flight f_3,and causing flight f_3 to be delayed or cancelled if the delay is too long. As a result, passengers connecting from flight f_3 to other flights will more likely be disrupted. Our goal is to consider the

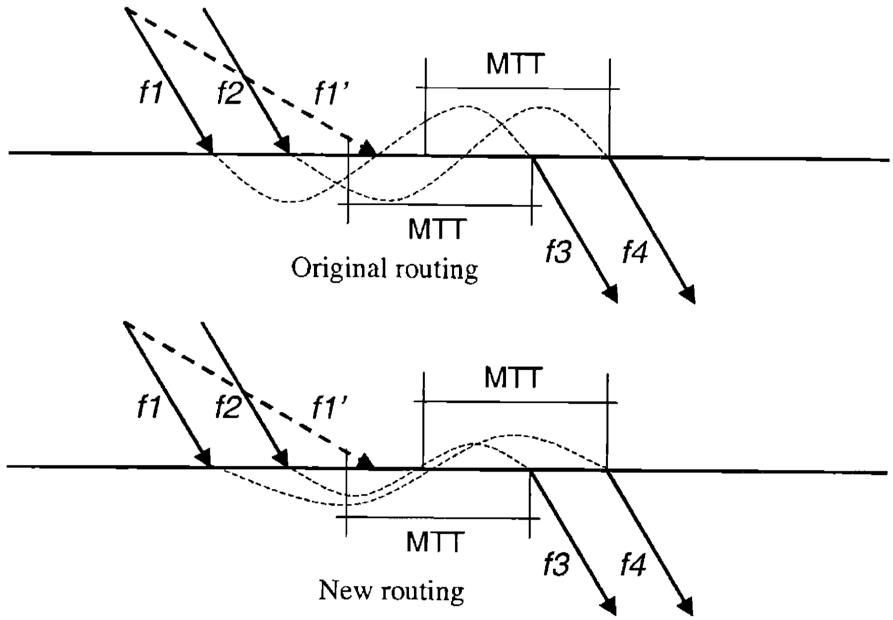

Figure 3-2: Illustration of the Idea

historical delay data in selecting aircraft routes, so that the delay and/or cancellation of flight f_3 and the resulting passenger disruptions can be reduced. To illustrate, assume that historical data shows that on average flight f_2 arrives on time. Then, a better way to construct the aircraft routes is illustrated in the "new routing" shown in the Figure 3-2, that is to put flight f_1 and flight f_4 in the same route and flight f_2 and flight f_3 in another route. The effect is to add more slack after flight f_1 to mitigate the downstream effects.

This problem can be solved separately for each fleet type. Because delays propagate along the aircraft routes, it is difficult to use leg-based models to track delay propagation. Thus, a route-based model is a more appropriate model for this problem.

We present in this section a string-based formulation for robust aircraft maintenance routing problem. A string is defined to be a sequence of connected flights that begins and ends at maintenance stations (not necessarily the same one). A string has the following properties (Shenoi (1996) [71]):

- The origin of the first flight and the destination of the last flight in the sequence are maintenance stations for that fleet;
- The destination of a flight is the origin of the next flight in the sequence;
- The flying time and elapsed time of the sequence does not exceed the maximum time-between-maintenance limits required by law and by airlines;
- The sequence satisfies any additional constraints imposed by the airline, such as maximum flying time per day, etc.;
- The starting time of the string is the starting time of the first flight in the sequence; the ending time of the string is the *maintenance ready time* of the last flight in the sequence for that fleet, where the maintenance ready time of a flight is defined as the time that the aircraft flying that flight will be ready after it has been serviced with an A-check.

Because the maintenance routing problem will be solved one or two months before the actual schedule inception date, the exact operational information including possible delays and disruptions is not known. We minimize the total *expected* propagated delay, meaning that the resulting routes might not minimize propagated delays for operations everyday, but should reduce propagated delay overall.

3.3.1 Time-Space Network

The underlying network structure of our model is a directed time-line network (see Figure 3-3), with the arc set representing the set of flights and the set of ground variables. The nodes in the graph correspond to either the departure or arrival of a flight. Arcs representing ground variables connect consecutive nodes at a single location, with a *wrap-around* arc connecting the first and last event at a location. *Count line* is an arbitrarily chosen point of time to count the number of aircraft on the ground and in the air.

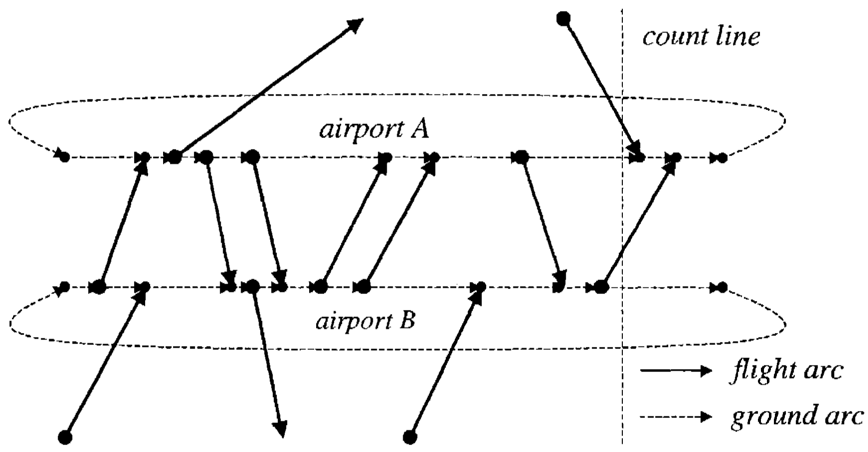

Figure 3-3: An Example of a Time-line Network Involving Two Airports

3.3.2 Determining Delays for Feasible Routes

One difficulty in modeling propagated delay is that both propagated delay and total arrival delay are a function of routing. While we have the historical data for propagated delay and total arrival delay for each flight based on existing routings, we do not have propagated delay and total arrival delay data for most of the feasible routes that have not been previously realized. Therefore, these delays in the historical data cannot be used directly to determine the objective function coefficients. Independent arrival delay, however, is not a function of routing. From historical data, the independent arrival delays can be calculated for each flight by tracking the routing of each individual aircraft. Given any feasible route, the total arrival delays and propagated delays of flights on this route can be generated using historical data as described below:

Algorithm 1 *generate delay data*

1. *Determine propagated delays (PD) in the historical data:*

   $PD_{ij} = \max(TAD_i - slack_{ij}, 0)$

2. *Determine independent arrival delays (IAD) for each flight from historical data:*

   $IAD_j = TAD_j - PD_{ij}$

3. Determine total arrival delay (TAD) and PD for each flight of any routing:

- *For the first flight on each string,* $TAD = IAD$
- $TAD_j = IAD_j + PD_{ij}$

3.3.3 Delay Distribution

We determine the distribution of delay with historical data from Airline Service Quality Performance (ASQP) database. The ASQP database provides flight information for all domestic flights served by jet aircraft by major airlines in the U.S. (those generating revenues of $1 billion or more annually). This database is available to the general public. ASQP provides the following flight operation information for each flight: planned departure time and arrival time, actual departure time and arrival time (including wheels-off and wheels-on time, taxi-out and taxi-in time, airborne time) and airplane tail number. For cancelled flights, reasons for cancellation and airplane tail number are not available.

We analyzed possible distributions for total arrival delays, including Normal, Exponential, Gamma, Weibull, Lognormal, etc. The arrival delays are usually strongly asymmetric. There are cases in which flights arrive early (the arrival delays are negative), but most flights arrive on time or late (the arrival delays are equal to or greater than zero). More specifically, most flights arrive around the scheduled arrival time, with very few of them arriving very early (more than 20 minutes), and some arriving very late (more than one hour). Therefore, the natural candidates for the arrival delay distributions are the Gamma, Lognormal and Weibull distribution.

SAS was used to estimate the parameters and calculate the test statistics. The χ^2 test and/or the Kolmogorov test were used to determine if the total arrival delays follow a specific distribution. We found that the log-normal distribution is the best fit among all possible distributions listed above.

From the Table 3.1, we can see that if a significant level 0.01 is chosen, the null hypothesis is accepted for 84% of all flights, implying that the actual arrival delays for

	$\alpha = 0.05$		$\alpha = 0.01$	
Num of flights	Num of flights with H_0 accepted	% of total flights	Num of flights with H_0 accepted	% of total flights
1448	1002	69%	1223	84%

Table 3.1: Test Results for H0: Total Arrival Delays Follow Lognormal Distribution.

84% of the flights follow a lognormal distribution. For these flights, the shape parameters are usually less than one and location parameters are less than 0.

A variable x is lognormally distributed if $y = \ln(x)$ is normally distributed with "ln" denoting the natural logarithm. The general formula for the probability density function (PDF) of the lognormal distribution is:

$$f(x) = \frac{e^{-\frac{(\ln \frac{x-\theta}{m})^2}{2\sigma^2}}}{(x-\theta)\sigma\sqrt{2\pi}}, \tag{3.4}$$

where σ is the shape parameter, θ is the location parameter, and m is the scale parameter. Examples of the lognormal PDFs are shown in Figure 3-4 (Engineering Statistics Handbook (2002) [36]).

A location parameter simply shifts the graph left or right on the horizontal axis. For these examples, the location parameter $\theta = 0$. Shape parameters allow a distribution to take on a variety of shapes, depending on the value of the shape parameter. Many probability distributions are not a single distribution, but are in fact a family of distributions. This is due to the distribution having one or more shape parameters. For example, the shapes of the Weibull distribution include an exponential distribution, a right-skewed distribution, and a relatively symmetric distribution. The effect of the scale parameter is to squeeze or stretch the PDFs.

The distribution parameters can be obtained using Maximum Likelihood Estimation (MLE) given historical delay data and new routing information. MLE is probably the most general and straight-forward procedure for finding estimators. A MLE estimator is the value of the parameters for which the observed sample is most likely to have occurred. For details of MLE, readers are referred to Ben-Akiva and Lerman (1985) [15], Pindyck

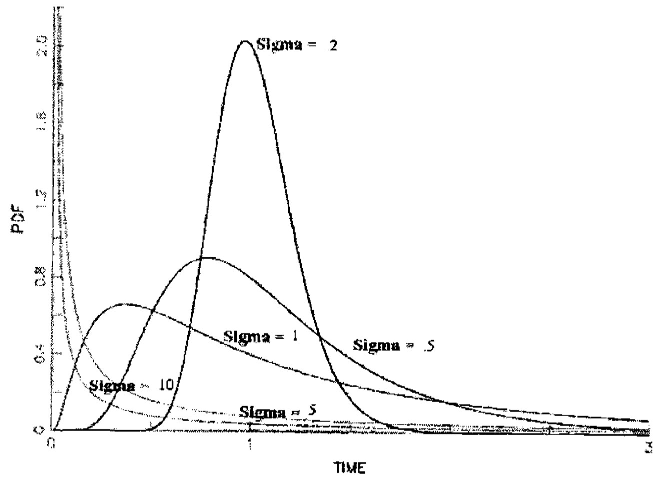

Figure 3-4: Examples of the Lognormal Probability Density Function

and Rubinfeld (1997) [63].

The maximum likelihood estimates for the scale parameter, m, and the shape parameter, σ , are

$$\hat{m} = \exp(\hat{u}), \text{ and}$$

$$\hat{\delta} = \sqrt{\frac{\sum_{i=1}^{N}(\ln(x_i) - \hat{u})^2}{N-1}},$$

where

$$\hat{u} = \frac{\sum_{i=1}^{N}\ln(x_i)}{N}.$$

If the location parameter is known, it can be subtracted from the original data points before computing the maximum likelihood estimates of the shape and scale parameters. Many commercial software such as SAS can automatically estimate the parameters based on the sample data.

3.3.4 Formulation of the Robust Aircraft Maintenance Routing Model

Let S be the set of feasible strings and F be the set of flights. We denote the set of ground variables as G, the set of strings ending with flight i as S_i^-, and the set of strings beginning with flight i as S_i^+. We have one binary decision variable x_s for each feasible string s. We have ground variables denoted by y, which are used to count the number of aircraft on the ground at maintenance stations. Let pd_{ij}^s be the delay propagated from flight i to flight j if flight i and flight j are in string s. Let a_{is} equal 1 if flight i is in string s, and equal 0 otherwise. Ground variables $y_{i,d}^-$ equal the number of aircraft on the ground before flight i departs at the maintenance station from which flight i departs, and ground variables $y_{i,d}^+$ equal the number of aircraft on the ground after flight i departs at the maintenance station from which flight i departs; Ground variables $y_{i,a}^-$ equal the number of aircraft on the ground before flight i arrives at the maintenance station, and ground variables $y_{i,a}^+$ equal the number of aircraft on the ground after flight i arrives at the maintenance station. r_s is the number of times string s crosses the *count time*, a point in time at which aircraft are counted, p_g is the number of times ground arc g crosses the count time, and N is the number of planes available.

The Robust Aircraft Maintenance Routing model (RAMR) is written as follows.

$$\min E(\sum_{s\in S}(\sum_{(i,j)\in s} pd_{ij}^s)x_s) \tag{3.5}$$

Subject to

$$\sum_{s\in S} a_{is}x_s = 1 \qquad \forall i \in F; \tag{3.6}$$

$$\sum_{s\in S_i^+} x_s - y_{i,d}^- + y_{i,d}^+ = 0 \qquad \forall i \in F; \tag{3.7}$$

$$-\sum_{s\in S_i^-} x_s - y_{i,a}^- + y_{i,a}^+ = 0 \qquad \forall i \in F; \tag{3.8}$$

$$\sum_{s\in S} r_s x_s + \sum_{g\in G} p_g y_g \leq N; \tag{3.9}$$

$$y_g \geq 0 \qquad \forall g \in G; \tag{3.10}$$

$$x_s \in \{0,1\} \qquad \forall s \in S \tag{3.11}$$

The objective is to minimize the expected total propagated delay of selected strings. The first set of constraints are cover constraints that ensure each flight is in exactly one string. The second and third sets of constraints are flow balance constraints, ensuring that the number of aircraft arriving at and departing from a location are equal. The fourth constraint is the count constraint to ensure that the total number of aircraft in use at count time T (and thus at any point in time) does not exceed the number of aircraft in the fleet. The last two sets of constraints force the number of aircraft on the ground to be non-negative and the number of aircraft assigned to a string to 0 or 1. Because variable y_g is a sum of binary x variables, the integrality constraints on the y variables can be relaxed, as discussed in Hane et al. (1995) [42].

3.4 Solution Approach

3.4.1 Overview of the Solution Approach

The robust aircraft maintenance routing (RAMR) problem is a stochastic discrete optimization problem. There is extensive literature addressing variants of this problem type. For a detailed literature review, the reader is referred to Kleywegt et al. (2001) [52], in which they also propose a Monte Carlo simulation-based approach for solving these problems. Their method is particularly applicable for the case that the expected value function in the objective cannot be written in closed form and/or its values cannot be easily calculated. Our model, however, is a stochastic discrete optimization problem without random variables in the constraints, and with an expected value function that can be calculated easily as explained in Section 3.4.2.

The objective function can be re-written as follows:

$$\min E[\sum_{s} x_s \times (\sum_{(i,j)\in s} pd_{ij}^s)] \quad = \min \sum_{s} x_s \times E[\sum_{(i,j)\in s} pd_{ij}^s] \quad =$$

$$\min \sum_{s\in S}(x_s \times \sum_{(i,j)\in s} E[pd_{ij}^s]) \tag{3.12}$$

Therefore, the objective function in original formulation 3.5 can be replaced by 3.12. The RAMR formulation is a deterministic mixed-integer linear program with a large number of 0-1 variables. For realistic problems, the complete generation of the corresponding instance, let alone its solution, requires prohibitive amounts of time and memory. The problem can be solved, however, using a *branch-and-price* approach. Branch-and-price is branch-and-bound with a linear programming relaxation solved at each node of the branch-and-bound tree using column generation. This is a non-trivial extension of branch-and-bound because the LP at each node of the enumeration tree has to be solved using implicit enumeration methods such as column generation, and the branching rules should not make the column generation pricing problem difficult to solve. This approach is particularly good for integer programming problems with huge numbers of variables. It works best when negative reduced cost columns can be generated (the pricing problem) without examining all variables (Barnhart et al. (1998) [9], Martin (1999) [59]).

3.4.2 Objective Function Coefficient

As shown in Section 3.3.3, we model total arrival delays using the lognormal distribution, and the distribution parameters can be obtained using Maximum Likelihood Estimation given historical delay data and new routing information. Based on it, the expected propagated delay for a flight pair can be determined as follows.

We know from Equation 3.3 in Section 3.3.3 that $PD_{ij} = \max(TAD_i - Slack_{ij}, 0)$. To simplify the notation, let y represent PD_{ij}, x represent TAD_i, c represent $Slack_{ij}$, we have $y = \max(x - c, 0)$, where x follows a lognormal distribution with the probability density function $f(x)$ described in Equation 3.4 . Because c is a constant and can be captured by location parameter θ, we have $y = g(x) = \max(x, 0)$.

$$E(PD_{ij}) = E(y) = E(g(x)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$$
$$= \int_0^{+\infty} xf(x)dx = \int_0^{+\infty} x\frac{e^{-\frac{(\ln\frac{x-\theta}{m})^2}{2\sigma^2}}}{(x-\theta)\sigma\sqrt{2\pi}}dx.$$

The expected propagated delays can be calculated as follows:

$$E(y) = (1 - \Phi(\frac{\ln(\frac{-\theta}{m})}{\sigma}))(\theta + me^{\frac{1}{2}\sigma^2}), \tag{3.13}$$

where $\Phi(\cdot)$ is the Cumulative Distribution Function of a standard normal distribution, which is calculated from the expression $\Phi(k) = \int_{-\infty}^{k} \frac{e^{-x^2/2}}{\sqrt{2\pi}}dx$. The values of $\Phi(\cdot)$ can be obtained in a normal distribution function table and by many commercial software.

3.4.3 Branch-and-Bound

Branch-and-bound is a divide-and-conquer method of solving integer programs (IP) (Bradley et al. (1977) [20]). This method first relaxes the integrality constraints of the IP, and solves the resulting linear program (LP). This LP is referred to as the root node of the branch-and-bound enumeration tree. If the optimal LP solution has no fractional values, the solution is optimal. Otherwise, the fractional solution is eliminated using a branching rule, that is, a set of constraints that partitions the feasible integer polyhedron into mutually exclusive subdivisions. This creates two new nodes in the enumeration tree, and an LP is solved at each of these nodes. If the LP is infeasible, then the corresponding IP must be infeasible, thus this node is fathomed. If the LP optimal solution is not better than the current best integer solution, this node is fathomed, because exploring this branch will not improve the IP solution. If the LP optimal solution is integral and better than the current best integer solution, then this node is fathomed and the current best integer solution is updated. If the LP optimal solution is fractional but the solution is better than the current best integer solution, this node is kept and further exploration is needed from this node, because exploring this branch might improve the IP solution.

3.4.4 LP Relaxation

Column generation is used to solve the linear programming relaxation of the RAMR problem, because it is impractical to explicitly enumerate all feasible strings. Starting with only a subset of variables (the corresponding problem, called the restricted master problem), the column generation algorithm that solves the pricing problem determines a set of optimal dual values to compute the reduced costs of nonbasic variables that might improve the solution. Variables with negative reduced cost should be added to the restricted master problem. The restricted master problem and the pricing problem need to be solved repeatedly until no variables have negative reduced costs. The steps to solve the LP relaxation are summarized as follows:

Algorithm 2 *LP relaxation of RAMR*

1. *Create initial feasible solution, form the restricted master problem (RMP), a LP with a restricted subset of the variables.*

2. *Solve the RMP to find an optimal primal and dual solution.*

3. *Solve the pricing problem. If one or more variables with negative reduced cost are identified, add them to the RMP and go to step 2; else stop: the LP is solved.*

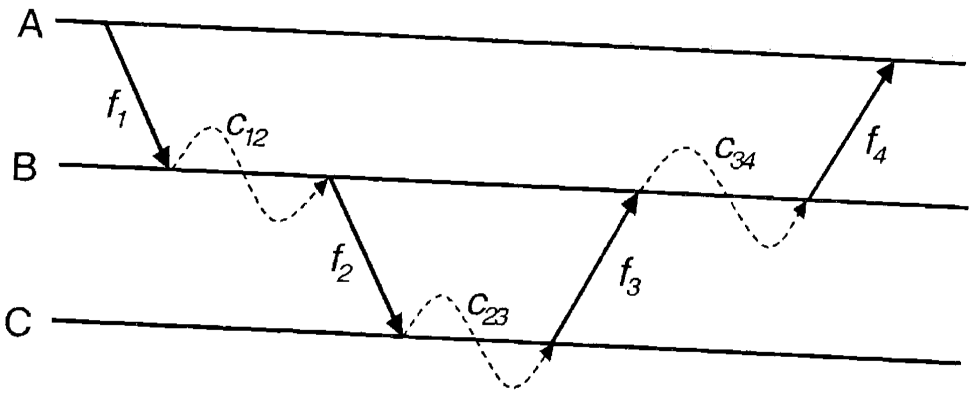

Figure 3-5: An Example of a Connection Network

3.4.5 The Pricing Problem

$d_s = \sum_{(i,j)\in s} E[pd^s_{ij}]$ represents the total propagated delay along string s. Let π_i be the dual variable associated with the cover constraint for flight i, δ be the dual variable corresponding to the count constraint, λ_i be the dual variable corresponding to the flow balance constraint for string s beginning or ending with flight i. The reduced cost of a string s beginning with flight m and ending with flight n is:

$$\overline{d_s} = d_s - \sum_i a_{is}\pi_i - r_s\delta - \lambda_m + \lambda_n.$$

Barnhart et al. (1998) [6] show that the pricing sub-problem of their string-based maintenance routing model can be cast as a constrained shortest path problem in a connection network. An example of a connection network is shown in Figure 3-5. For their model, each cost component in the reduced cost can be assigned to a flight arc or a connection arc so that the reduced cost of a string equals the total length of the arcs in that string. The variable with minimum reduced cost, therefore, corresponds to the path that has the least cost. Interested readers are referred to Barnhart et al. (1998) [6] for more details.

Figure 3-5 shows a connection network with three airports, A, B, and C, and four flights numbered 1 through 4. There are three connection arcs, c_{12}, c_{23} and c_{34}. Suppose airports A and C are maintenance stations. Examples of possible strings are string s_1: $f_3 - c_{34} - f_4$ and string s_2: $f_1 - c_{12} - f_2 - c_{23} - f_3 - c_{34} - f_4$.

For our model, the pricing problem cannot be cast as a shortest path problem. The reason is that d_s ($= \sum_{(i,j)\in s} E[pd_{ij}^s]$) cannot be assigned to each connection arc because the propagated delay for each pair of flights depends on which string they belong to. For example (see Figure 3-5), in string s_1: $f_3 - c_{34} - f_4$, if flight f_3 is delayed for 30 minutes(this delay is "independent delay", because flight f_3 is the start flight of this string and there is no delay propagated into flight f_3), there will be a 20 minute delay propagated to flight f_4 assuming the slack between these two flights is 10 minutes. If flight f_3 and flight f_4 belong to s_2: $f_1 - c_{12} - f_2 - c_{23} - f_3 - c_{34} - f_4$, and there is a 10 minute delay propagated from flight f_2 to flight f_3, then the delay of flight f_3 is increased to 40 minutes. In contrast, there will be a 30 minute delay propagated from flight f_3 to flight f_4 (the total delay for flight f_3 is 40 minutes, but the slack between flight f_3 and f_4 is 10 minutes).

We propose an approximate way to solve the pricing problem without explicitly evaluating the reduced cost for each possible string. We construct a connection network by assigning each component of $-\sum_i a_{is}\pi_i - r_s\delta - \lambda_m + \lambda_n$ to flight arcs and connection arcs. We solve shortest path problems for all OD pairs of the network. If the costs for all shortest paths are greater than or equal to zero, then no columns have negative reduced cost, because d_s is greater than or equal to zero, by definition. Thus, no columns will be added and the LP problem has been solved optimally. If the costs for some shortest paths are less than zero, then we add these costs to d_s. If the resulting total costs are less than zero, then the corresponding columns will be added to RMP. Otherwise, no columns will be added. We cannot claim optimality of the LP at this step because there might be unidentified paths with cost less than zero. Although this method does not guarantee optimality of its solutions, but is tractable because it doesn't require the enumeration of all variables.

3.4.6 IP solution

An integer solution to the robust aircraft maintenance routing problem can be obtained by using a special branching strategy called "branching on follow-ons" (Ryan and Foster (1981) [68], Barnhart et al. (1998) [6]). As proved in Barnhart et al. (1998) [9], this strategy will generate optimal integer solutions to the problem. We summarize this strategy as follows:

Algorithm 3 *Branching on Follow-ons*

1. If the solution is not fractional, the current maintenance routing problem is solved. If the solution is fractional, identify a fractional string s_1 with $0 < x_{s_1} < 1$. Denote the sequence of flights in s_1 as $f_1, f_2, f_3, ..., f_{n-1}, f_n$.

2. There must be another string s_2 containing flight f_i^* but not f_{i+1}^*, since $0 < x_{s_1} < 1$ and each flight must be covered exactly once, and because the LP solution cannot contain two identical variables.

3. Define S_L as the set of strings with each string containing flight f_i^* followed by f_{i+1}^*, and let S_R be the set of strings containing flight f_i^* and/or flight f_{i+1}^*, and f_i^* is not followed by f_{i+1}^*. We create left and right branch as follows:

   - On the left branch, we force flight f_i^* to be followed by flight f_{i+1}^* with $\sum_{s \in S_L} x_s = 1$. To ensure the pricing subproblem generates strings satisfying this rule, we eliminate from the connection network all arcs connecting flight f_i^* to any flight other than flight f_{i+1}^* and all arcs connecting to flight f_{i+1}^* from any flight other than flight f_i^*.

   - On the right branch, we do not allow flight f_i^* to be followed by flight f_{i+1}^*, that is, we require that $\sum_{s \in S_R} x_s = 1$. To ensure the pricing subproblem generates only strings satisfying this rule, we eliminate from the network all arcs connecting flight f_i^* to flight f_{i+1}^*.

Network	Num of flights	Num of strings
N1	20	7,909,144
N2	59	614,240
N3	97	6,354,384
N4	102	51,730,736

Table 3.2: Characteristics of Four Maintenance Routing Problems.

3.5 Proof-of-Concept

3.5.1 Underlying Networks

Table 3.2 presents the characteristics of four different maintenance routing problems, each represents a different fleet type. Column *Num of strings* represents all possible strings for each network. Although the number of flights in each fleet is relatively small, the number of possible strings is very large.

3.5.2 Data and Validation

We have the July and August 2000 data for a major U.S. airline consisting of:

- ASQP data: provides the following flight operation information for each flight: planned departure time and arrival time, actual departure time and arrival time (including wheels-off and wheels-on time, taxi-out and taxi-in time, airborne time) and airplane tail number

- Airline data: include number of passengers on each itinerary.

To validate our model, we build the model based on historical information, and generate aircraft routes, and then apply these routes to future operations. Suppose that we were at the end of July 2000, and we need to determine the aircraft routes for next month, August 2000. Thus, we estimate the objective function coefficients (propagated delays) based on July 2000 operational data, we solve our RAMR model to obtain aircraft routes that will be use for next month (August 2000), and then calculate the delays and the

Network	Old PD	New PD	PD reduced	% of PD reduced
N1	5723	4091	1632	29%
N2	3553	1388	2165	61%
N3	7128	3217	3911	55%
N4	9152	7108	4321	47%
Total	25556	15804	12029	47%

Table 3.3: Propagated Delays Based on July 2000 Data.

number of passenger misconnections based on the new routing solution from our model and existing routing from historical data. Finally we compare these numbers.

3.5.3 Computational Results

In this section we present results obtained by applying our robust maintenance routing model to data sets representing July and August 2000 operations of a major U.S. airlines. Our solution algorithm is implemented in C++ and CPLEX 6.5 on a HPC 3000 workstation.

Delay Analysis

The results based on July 2000 data are listed in Table 3.3. Column *Old PD* indicates the propagated delay in minutes in the historical data, column *New PD* indicates the propagated delay in minutes for our routing solution, column *PD reduced* indicates the reduction in minutes of propagated delay resulting from our new routing solution, and column *% of PD reduced* indicates the percentage reduction in propagated delay. On average, our routing solution can reduce total propagated delay for the four networks by 47%.

Our model parameters are determined using the July 2000 data, thus these results do not represent the value of our model when used in real-life setting in which the actual delays are unknown. In practice, the model will be built using historical data, and then

Network	Old PD	New PD	PD reduced	% of PD reduced
N1	6749	4923	1826	27%
N2	4106	2548	1558	38%
N3	8919	4113	4806	54%
N4	14526	9921	6940	48%
Total	34300	21505	15130	44%

Table 3.4: Propagated Delays Based on August 2000 Data.

P-delay	(0,30]	(30,60]	(60,90]	(90,120]	>120	>0
Old	4.8%	1.8%	1.2%	0.5%	0.7%	9.1%
New	2.6%	0.9%	0.7%	0.2%	0.6%	5.0%

Table 3.5: Distribution for propagated delays.

will be applied to future operations. Consistent with this, we applied our July 2000 routing solution to August 2000 and obtained the results presented in Table 3.4. On average, the RAMR model reduced total propagated delay in August by 44%.

The distributions of propagated delays for both the actual aircraft routings and our routings using August 2000 data for the four networks are summarized in Table 3.5. "*(a,b]*" indicates that the propagated delay is greater than a minutes and less than or equal to b minutes. The row *Old* represents the percentage of flights with propagated delay in specified ranges in historical data (what actually happened with the airline's aircraft routings), the row *New* represents the percentage of flights with propagated delay in specified ranges based on our new routing solution. These results show that our routing solution reduce number of flights with delay propagation in each time window.

The distributions of total delays for existing routing and new routing using August 2000 data for the four networks are summarized in Table 3.6. The Department of Transportation (DOT) defined on-time rate (delay less than 15 minutes) increases 1.6%, while the 60 and 120 minutes on-time rate (delay less than 60, and 120 minutes respectively) are also improved. The reduction in the number of flights suffering long delay should help to reduce the number of passengers and crews who have to be re-accommodated and reduce the number of flight cancellations.

	Total delay			on-time rates		
	>15 min	>60 min	>120 min	15 min	60 min	120 min
Old	22.3%	7.9%	2.9%	77.7%	92.1%	97.1%
New	20.7%	6.9%	2.6%	79.3%	93.1%	97.4%

Table 3.6: Distribution for Total Delays and On-time Performance.

Airlines	Northwest	Continental	Delta	TWA	Southwest
On-time rates	79.2%	77.7%	77.3%	76.7%	76.2%
Rank	1	2	3	4	5

Table 3.7: On-time Performance Rank for U.S. Major Airlines.

The on-time performance defined by DOT is an important indicator of airlines' performance and level of service and is publicly available, thus has significant impact on airlines' image. In August 2000, the 15 minute on-time performance for major U.S. airlines is shown in Table 3.7 (We only list the top five) (Bureau of Transportation Statistics [22]). A 1.6% increase of the on-time rates can improve the rank of any of these listed airlines. Most notable, it would move the second place airline, Continental, into first place.

Impact on Passengers

In our data, the average delay for disrupted passengers is approximately 419 minutes while the average delay for non-disrupted passengers is approximately 14 minutes (Bratu and Barnhart (2002)[21]). These lengthy delays cause tremendous revenue loss and operational challenges. Hence, we investigated the effect of our routing solution on passenger disruptions. We compare the number of disrupted passengers based on the existing routings and the our routings. To do so, we compute the number of disrupted passengers using a simple, approximate method. The first step is to re-generate the departure and arrival data for each flight:

Algorithm 4 *Re-generate the actual departure and arrival time for each flight for each day*

1. *Determine independent departure delays (IDD) and arrival delay (IAD) for each flight from historical data:*

   - $IDD_j = TDD_j - OLD\_PD_{ij}$, $IAD_j = TAD_j - OLD\_PD_{ij}$ *where* OLD_PD_{ij} *is the propagated delay based on existing routings, and can be determined by Equation 3.3*

2. *Determine New_ TDD and New_ TAD according to new routings:*

   - *For the first flight* j *on each string:* $New\_TDD_j = IDD_j$, *and* $New\_TAD_j = IAD_j$.

   - *For all other flights:* $New\_TDD_j = IDD_j + NEW\_PD_{ij}$, *and* $New\_TAD_j = IAD_j + NEW\_PD_{ij}$ *where* NEW_PD_{ij} *is the propagated delay based on our routings.*

3. *Determine the actual departure and arrival time of each flight for the our routings:*

   - *The* $New\_ACT_j = PDT_j + New\_TDD_j$

   - *The* $New\_AAT_j = PAT_j + New\_TAD_j$

The next step, calculating the number of disrupted passengers for a given routing solution, is achieved as follows:

- Passenger disruptions are calculated at the flight level. If a flight is cancelled, all passengers on that flight are disrupted.

  Suppose flight A is followed by flight B and both flights are operated, if $ADT_B - AAT_A < T_{\min}$, where $T_{\min}$ is the minimum connecting time for a passenger, then all passengers connecting from flight A to flight B are disrupted.

- For those flights without ASQP records, we don't have the data for the actual departure and arrival time. Therefore, we count only the disrupted passengers with connections for which both flights have ASQP records.

Network	Total num of D-pax	D-pax reduced	D-pax reduced (%)
N1	986	147	14.9%
N2	1070	79	7.4%
N3	1463	161	11.0%
N4	3323	355	10.7%
Total	6842	742	10.8%

Table 3.8: Results on Disrupted Passengers.

- Passengers are only counted as disrupted once. If a passenger is disrupted on any flight leg of his/her itinerary, that passenger is not counted as disrupted on any other flight legs.

We use the above method to estimate the number of disrupted passengers for both the historical routing and our routing for August 2000. The results are summarized in Table 3.8.

Column *D-pax reduced* represents the number of disrupted passengers reduced by using our routing solution, column *Total num of D-pax* represents the total number of disrupted passengers caused by flight delays (not by flight cancellations) for the historical routing, and column *D-pax reduced (%)* represents the percentage reduction in disrupted passengers. On average the RAMR routing reduces by about 11 percent the number of disrupted passengers caused by flight delays.

Chapter 4

Flight Schedule Retiming to Reduce Passenger Missed Connections

4.1 Passenger Delay and Disruption

As described in Chapter 1, flight delays and cancellations, leading to other flight, crew and passenger delays and disruptions, have increased significantly from 1995 to 2000. Recall, a passenger is "disrupted" if his or her planned itinerary is infeasible because one or more flights in the planned itinerary is cancelled or there is insufficient time for him or her to connect between flights. Studies show that the level of service experienced by passengers is correlated with airline profitability. For example, Flint (2000) [39] tries to make this point in his report stating that America West "found itself at the bottom of the DOT Consumer Report" in 1999, and "...business load fell 2 points in the second half of 2000, ...This contributed to a 98% fall in annual profits".

In this same time period, passenger dissatisfaction increased dramatically. From 1995 to 2000, the number of passenger complaints per 100,000 passengers for major U.S. airlines increased from 0.76 to 2.98 (almost 4 times) (Bratu and Barnhart (2002) [21]). Passenger delay and disruption are the main factors behind these complaints. Bratu and

	Ave. delay	% Pax	% Total pax delay
Disrupted pax	419 min	4%	51%
Nondisrupted pax	14 min	96%	49%
All pax	31 min		
Flights	16 min		

Table 4.1: Flight and Passenger Delays.

Barnhart (2002) [21] compute flight delays and passenger delays in August 2000 using data from a major U.S. airline. As shown in Table 4.1, the average flight delay is 16 minutes, while the average delay for passengers is 31 minutes. This table also shows that the average delay for disrupted passengers is 419 minutes while the average delay for nondisrupted passengers is only 14 minutes. Thus, while disrupted passengers are only a small proportion of all passengers (4%), they account for more than half of the total passenger delay minutes. Bratu and Barnhart (2002) [21] further estimate that the disrupted passengers represent on average 86.3% of the passengers who experience more than 4 hours arrival delay at their destination. These numbers show the significance of disrupted passengers.

While short delays, such as the average delay of 14 minutes, are not good, they are acceptable for most passengers. The delay experienced by disrupted passengers, however, are excessive. Significantly reducing disrupted passengers will lead to large reductions in overall passenger delay and to reductions in the "most critical" part of passenger delay. Moreover, models to minimize number of disrupted passengers are much simpler than those to minimize total passenger delays.

Figure 4-1 illustrates some definitions related to passenger disruption. In this chapter, we consider disrupted passengers to be those passengers who miss their connections because of flight delays. As defined in Section 3.2, PDT refers to planned departure time, and ADT refers to actual departure time. PAT refers to planned arrival times, and AAT refers to actual arrival time. MCT refers to the minimum connecting time needed by a passenger to connect to the next flight in his or her itinerary. In reality, MCT may vary for different passengers, as both the distance between arrival and departure gates and the walking speed of the individual passengers can vary significantly. In the planning stage, however, we only have the forecasted number of passengers for flights and do not have the characteristics of these passengers, nor the gate locations of each flight. Thus, we assume that MCT is a constant for all passengers. PCT refers to planned connecting time, and ACT refers to the actual connecting time. Slack is the difference between the planned connecting time and the minimum connecting time. The relationships between these definitions are summarized as follows.

$$PCT = PDT - PAT, \tag{4.1}$$

$$Slack = PCT - MCT, \tag{4.2}$$

and

$$ACT = ADT - AAT. \tag{4.3}$$

For any connecting passenger, he/she will be disrupted if

$$ACT < MCT. \tag{4.4}$$

4.2 Modeling Idea

Although many passengers are disrupted, causing significant losses to airlines and passengers, no existing planning models attempt to reduce passenger disruptions. Here, we present our flight schedule retiming model to minimize the number of passenger missed connections. In this section, we describe the basic idea of this model.

Figure 4-2 illustrates the idea of our flight schedule retiming model. Assume that in a month there are 100 passengers connecting from flight f_2 to flight f_3. The solid lines with arrows represent flights, and the head of the arrow represents the arrival time plus

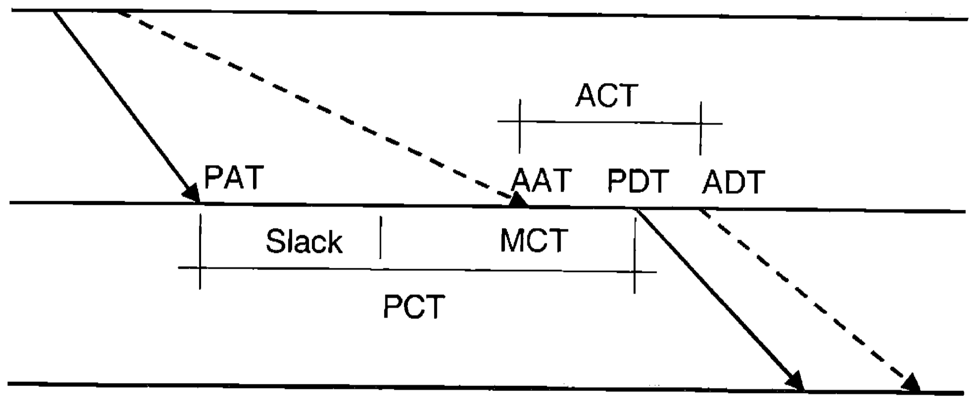

Figure 4-1: Definitions Related to Passenger Disruption

the 30 minute minimum aircraft turn time. We assume that the minimum connecting time for passengers is also 30 minutes. Suppose according to historical data, flight f_2 is frequently delayed to the position of f_2' with probability 0.3. Hence, the probability of a misconnection from flight f_2 to flight f_3 is 0.3, and on average, there are $0.3 \times 100 = 30$ passengers missing this connection. In the planning stage, if we move the departure time of flight f_3 to a later time represented by f_3^*, then even if flight f_2 is delayed to the position of f_2', passengers are not disrupted. Suppose flight f_2 is sometimes delayed to the position of f_2'' with probability 0.2. Now the probability of a misconnection from flight f_2 to flight f_3 is 0.2, and there are $0.2 \times 100 = 20$ passengers missing this connection, even if flight f_3 is in the position of f_3^*. Thus, by moving the departure time of flight f_3 to a later time, 10 connecting passengers are saved from disruptions. If we just consider flight f_2 and flight f_3, another solution to reach the similar effect is to move the departure time of flight f_2 to an earlier time, as shown by the position of f_2^*. However, if there is a flight f_1, as shown in the figure, with passengers connecting to flight f_2, and flight f_1 is also frequently delayed, then moving the departure time of flight f_2 to the position of

f_2^* might increase the total number of disrupted passengers.

As shown in Figure 4-1, if the slack is "eaten" by flight delay, passengers connecting between two flights will be disrupted. Adding more slack can be good for connecting passengers, but can result in reduced productivity of the fleet. The challenge then is to determine where to add this slack so as to maximize the benefit to passengers without requiring additional aircraft to fly the schedule. Moving flight departure times provides an opportunity to allocate slack to reduce passenger disruptions and maintain aircraft productivity. In practice, flight departure times are adjusted in small time windows beginning several weeks before the flights' departure up until the day of departure.

Levin (1971) [55] was the first to propose the idea of adding time windows to fleet routing and scheduling models. In that paper, time windows were modeled by allowing departure times to occur at discrete intervals within the time window. Desaulniers et al. (1997)[35] presented two formulations for the fleet assignment and aircraft routing problem with time windows, and building on the fleet assignment model developed by Hane et al. (1995) [42], Rexing et al. (2000) [65] presented a fleet assignment model that simultaneously selects departure times. Klabjan et al. (1999) [48] apply a similar idea to crew scheduling, and develop a crew scheduling model with time windows.

If we just consider passenger connections, more slack can result in fewer disrupted passengers. However, in practice, there are constraints on how much the departure times can be moved. If the departure times are moved too far from their original times, the demand for these flights can change. In addition, too much slack can require extra aircraft and crew, resulting in significant additional costs.

The time window, specifying how much time a given flight can be shifted, can be modeled with a simple extension of the basic flight network. By placing copies of a flight arc at specified intervals within that flight's time window and requiring only one of the flight arc copies to be used, we model the choice of flight departure time. Because the scheduled time of some flights is more flexible than others, the width of each time window is a parameter that can be different for every flight. Moreover, the interval between copies

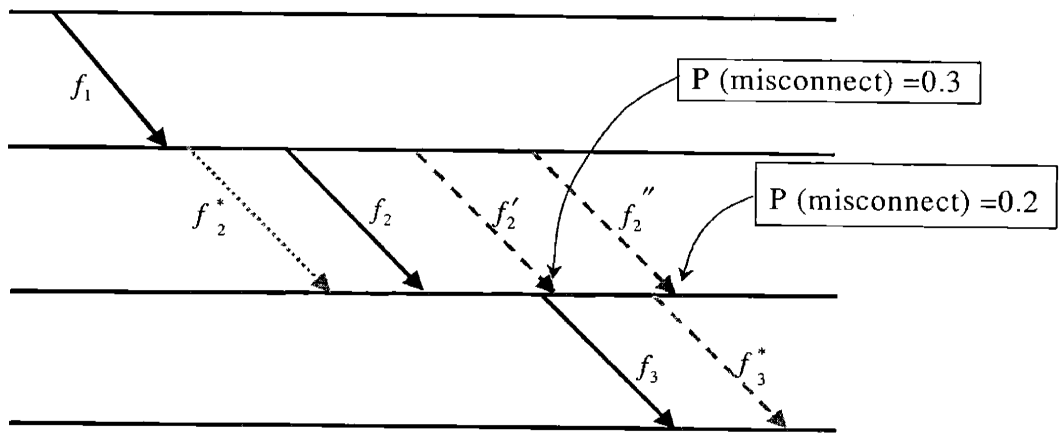

Figure 4-2: Illustration of Flight Schedule Retiming

is another parameter, one that can impact the tractability of the model and quality of the solution. To guarantee that flights are allowed to depart at any time within the time window, copies should be placed at one-minute intervals. However, it will be shown in Section 4.7 that using a narrow interval instead of a broader one causes an explosion in the problem size, but often fails to generate a substantially better solution.

By intelligently selecting flight departure times to *optimize* passengers' connecting time, the number of disrupted passengers can be reduced. We define a robust schedule as a planned schedule that, over a long time, minimizes the number of disrupted passengers in the operation.

To achieve this robust schedule, we will select flight departure times, given relatively small departure time windows, and solutions to the first three steps of the airline schedule planning process, namely, flight schedule design, fleet assignment and aircraft routing. Different fleet assignment and maintenance routing solutions cause differences in how delays propagate and passengers are disrupted. This can be illustrated using the same example shown in Figure 4-2. Suppose we solve the re-timing problem first, then solve the fleet assignment and maintenance routing problems. Based on historical data, flight f_1 was often on-time, and flight f_2 was delayed with various probabilities as described before. If flight f_2 is retimed to the position of f_2^*, and the fleet assignment and maintenance routing problems are re-solved, likely creating different aircraft routes, delays for these flights will also change. As a result, if flight f_1 is delayed frequently, then moving flight f_2 to the position of f_2^* is not the best choice anymore. Instead, flight f_3 should be moved to the position of f_3^*.

If this schedule retiming is done after solving fleet assignment and maintenance routing, this problem is avoided. Hence, in our work, we consider the fleet assignment and maintenance routes as fixed, and ensure that any schedule retiming solution does not violate the fleeting and routing solution.

4.3 Flight Schedule Retiming Models and Their Properties

4.3.1 A Connection-based Flight Schedule Retiming Model

We define a variable for each pair of flights with connecting passengers (see Figure 4-3). We have one binary decision variable $f_{i,n}$ for each flight i copy n, which is equal to one if flight i copy n is selected, and zero otherwise; one binary decision variable $x_{i,n}^{j,m}$ for the connection between flight i copy n and flight j copy m, which is equal to one if the connection between flight i copy n and flight j copy m is selected, and zero otherwise. Let $dp_{i,n}^{j,m}$ be the number of disrupted passengers between flight i and flight j if flight i copy n and flight j copy m are selected. We denote the set of all flights as F, the set of all flights with connecting passengers as F^C, the set of all flights to which passenger connecting as F^I, the set of all flights from which passenger connecting as F^O. Let N_i be the number of copies generated for flight i. We denote the set of flights with passengers connecting from flight i as $C^+(i)$, then $|C^+(i)|$ is the number of flights with passengers connecting from flight i. Similarly, we denote the set of flights with passengers connecting to flight i as $C^-(i)$, then $|C^-(i)|$ is the number of flights with passengers connecting to flight i. Our objective is to minimize the expected total number of disrupted passengers, subject to the following constraints:

1. For each flight, exactly one copy must be selected.

2. For each connection, exactly one copy will be selected, and this selected copy must connect the selected flight-leg copies. For example, in Figure 4-3, if flight i copy 2 and flight j copy 3 are selected, then the copy of the connection from flight i copy 2 to flight j copy 3 must be selected, that is, $x_{i,2}^{j,3} = 1$.

3. The current fleeting and routing solutions can not be altered.

Objective Function

The objective function of our retiming model to minimize the expected number of disrupted passengers can be written as:

$$\min E\left[\sum_{i\in F^O}\sum_{n\in N_i}\sum_{j\in C^+(i)}\sum_{m\in N_j} dp_{i,n}^{j,m}x_{i,n}^{j,m}\right] = \min \sum_{i\in F^O}\sum_{n\in N_i}\sum_{j\in C^+(i)}\sum_{m\in N_j} x_{i,n}^{j,m}\times E[dp_{i,n}^{j,m}]$$

In order to compute $E[dp_{i,n}^{j,m}]$, we need to know the distribution of $dp_{i,n}^{j,m}$, the number of disrupted passengers connecting from flight leg i copy n to leg j copy m, for all flight legs i and j and all copies n and m. We assume that if the difference between the actual departure time of flight j and the actual arrival time of flight i is less than the minimum connecting time, all passengers connecting from flight i to flight j are disrupted. Based on this, the distribution of $dp_{i,n}^{j,m}$ is a binary distribution, namely

$$dp_{i,n}^{j,m} = \left\{ \begin{array}{cc} c_{ij} & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{array} \right\}, \tag{4.5}$$

where c_{ij} is the number of passengers connecting from flight i to flight j. Probability p is determined as follows:

$$p = \text{prob}(ADT_{j,m} - AAT_{i,n} < MCT), \tag{4.6}$$

where $ADT_{j,m}$ is the actual departure time of flight j if copy m is selected, and $AAT_{i,n}$ is the actual arrival time of flight i if copy n is selected. As described in Chapter 3, once the robust aircraft maintenance routing problem is solved, the distribution of ADT and AAT for each flight leg can be determined. Then the $E[dp_{i,n}^{j,m}]$ can be determined for each connection between two flights.

Model Formulation

The Connection-Based Flight Schedule Retiming Model (CFSR) is written as follows.

$$\min \sum_{i \in F^O} \sum_{n \in N_i} \sum_{j \in C^+(i)} \sum_{m \in N_j} x_{i,n}^{j,m} \times E[dp_{i,n}^{j,m}] \tag{4.7}$$

subject to

$$\sum_{n \in N_i} f_{i,n} = 1 \qquad \forall i \in F^C; \tag{4.8}$$

$$\sum_{n \in N_i} \sum_{m \in N_j} x_{i,n}^{j,m} = 1 \qquad \forall i \in F^O, j \in C^+(i); \tag{4.9}$$

$$\sum_{m \in N_j} x_{i,n}^{j,m} = f_{i,n} \qquad \forall i \in F^O, n \in N_i, j \in C^+(i); \tag{4.10}$$

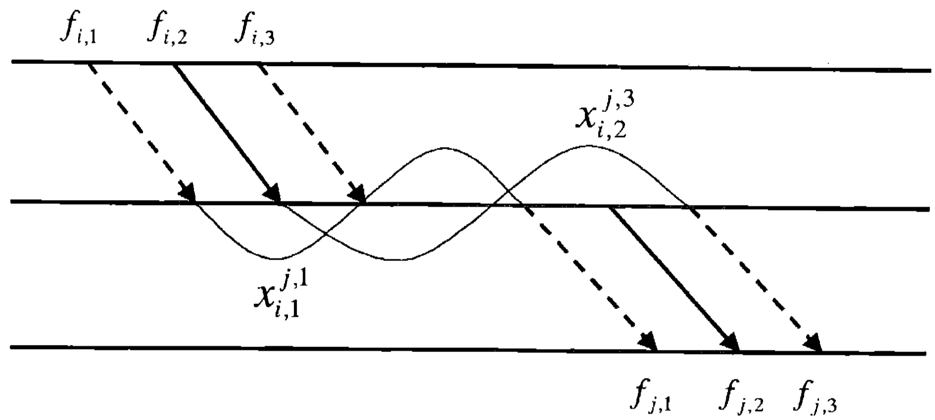

Figure 4-3: Illustration of the Variable Definition

$$\sum_{n \in N_i} x_{i,n}^{j,m} = f_{j,m} \qquad \forall j \in F^I, m \in N_j, i \in C^-(j); \tag{4.11}$$

$$f_{i,n} \in \{0,1\} \qquad \forall i \in F^C, n \in N_i; \tag{4.12}$$

$$x_{i,n}^{j,m} \in \{0,1\} \qquad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j. \tag{4.13}$$

The objective function minimizes the expected total number of disrupted passengers. Constraints 4.8 are cover constraints, in conjunction with the integrality requirements of variable f (the constraints 4.12) ensure that for each flight exactly one copy will be selected. Constraints 4.9 in conjunction with the integrality requirements of variable x (the constraints 4.13), ensure that for each connection exactly one copy will be selected. Constraints 4.10 and 4.11 jointly ensure that variables f and s are selected consistently. As we explained above, this problem will be solved after solving the fleet assignment and aircraft maintenance routing problems. Therefore, we need to add constraints to

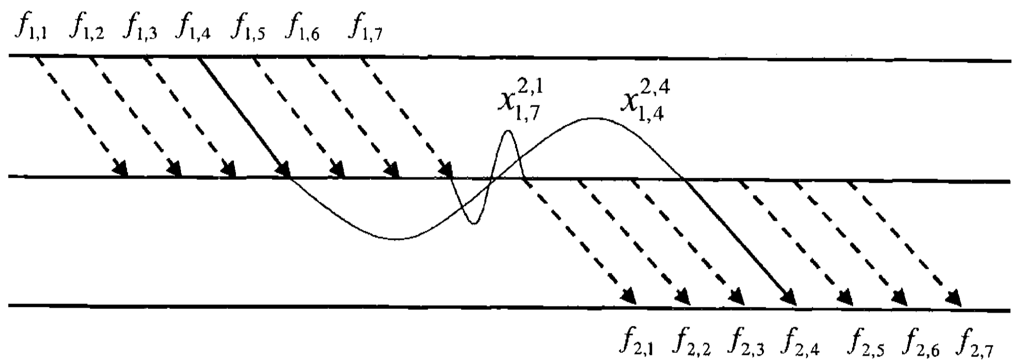

Figure 4-4: Example: How to Keep Current Routing Solution

maintain the current fleeting and routing solution, which is discussed next.

Enabling the Current Fleeting and Routing Solution

To maintain the current fleeting and routing solution while selecting flight departure times, we must ensure that the time for each aircraft to turn always exceeds the minimum turn time. For example, in Figure 4-4, suppose that flights 1 and 2 are in an aircraft route. If the time between the arrival of flight copy $f_{1,7}$ and the departure of flight copy $f_{2,1}$ is less than the minimum turn time, then flight 1 copy 7 and flight 2 copy 1 cannot be selected together, implying that $x_{1,7}^{2,1}$ must equal zero. In general, for any pair of flight legs $i-j$ in an aircraft route, we can keep the current routing solution feasible by forcing $x_{i,n}^{j,m} = 0$, if the time between the arrival of flight copy $f_{i,n}$ and the departure of flight copy $f_{j,m}$ is less than the minimum turn time. This can be easily implemented by setting to zero the upper and lower bounds for each x variable corresponding to a connection violating the minimum turn time requirement.

Model Properties

In this section, we analyze the model properties. Specifically, there are some constraints in this formulation can be eliminated or relaxed. In fact, the second set of constraints can be eliminated, because constraints 4.8, 4.10, 4.11 and 4.12 imply them. It is also not necessary to enforce the constraints 4.13,the integrality of the connection variables, as shown in the following.

Theorem 1 *The constraints* $\sum_{n \in N_i} \sum_{m \in N_j} x_{i,n}^{j,m} = 1, \forall i \in F^O, j \in C^+(i)$ *in the CFSR model are redundant and can be eliminated.*

Proof. Consider any pair of flights i_1 and j_1 such that flight i_1 is followed by flight j_1 in an aircraft routing. From 4.10, we have:

$$\sum_{m \in N_{j_1}} x_{i_1,n}^{j_1,m} = f_{i_1,n}, \forall n \in N_{i_1},$$

which implies

$$\sum_{n \in N_{i_1}} \sum_{m \in N_{j_1}} x_{i_1,n}^{j_1,m} = \sum_{n \in N_{i_1}} f_{i_1,n}.$$

Together with $\sum_{n \in N_{i_1}} f_{i_1,n} = 1$, this implies that for any pair of flights i_1 and j_1 in sequence in an aircraft route,

$$\sum_{n \in N_{i_1}} \sum_{m \in N_{j_1}} x_{i_1,n}^{j_1,m} = 1.$$

Hence, constraints 4.9 are redundant and can be relaxed. ■

Theorem 2 *The integrality of the connection variables 4.13 can be relaxed.*

Proof. Let us consider flights i_1 and j_1 such that flight i_1 is followed by flight j_1 in an aircraft routing. We will show that constraints 4.8, 4.10, 4.11, 4.12 and a relaxation of constraints 4.13 imply the satisfaction of the integrality requirements for x.

Constraints 4.8 and 4.12 ensure that, for every flight, exactly one copy will be selected. Suppose copy n_1 of flight i_1 and copy m_1 of flight j_1 are selected, then

$$f_{i_1,n_1} = 1;\ f_{i_1,n} = 0, \forall n \in N_{i_1} \text{ and } n \neq n_1;$$

$$f_{j_1,m_1} = 1;\ f_{j_1,m} = 0, \forall m \in N_{j_1} \text{ and } m \neq m_1.$$

From constraints 4.10, we have

$$\forall n \in N_{i_1} \text{ and } n \neq n_1,\ \ f_{i_1,n} = 0 = \sum_{m \in N_{j_1}} x_{i_1,n}^{j_1,m}.$$

Because $x \geq 0$, this implies

$$x_{i_1,n}^{j_1,m} = 0,\ \forall n \in N_{i_1} \text{ and } n \neq n_1,\ m \in N_{j_1}.$$

Similarly, we have

$$x_{i_1,n}^{j_1,m} = 0,\ \forall n \in N_{i_1},\ \forall m \in N_{j_1} \text{ and } m \neq m_1,$$

which implies

$$\sum_{m \in N_{j_1}, m \neq m_1} x_{i_1,n}^{j_1,m} = 0,\ \forall n \in N_{i_1},$$

and

$$\sum_{m \in N_{j_1}, m \neq m_1} x_{i_1,n_1}^{j_1,m} = 0.$$

Thus, together with constraints 4.10, we have:

$$
\begin{aligned}
f_{i,n_1} = 1 = \sum_{m \in N_{j_1}} x_{i_1,n_1}^{j_1,m} \\
= x_{i_1,n_1}^{j_1,m_1} + \sum_{m \in N_{j_1}, m \neq m_1} x_{i_1,n_1}^{j_1,m} \\
= x_{i_1,n_1}^{j_1,m_1}
\end{aligned}
$$

Hence, for any pair of flights i_1 and j_1

$$x_{i_1,n_1}^{j_1,m_1} = 1 \text{ and } x_{i_1,n}^{j_1,m} = 0,\ \forall n \in N_{i_1} \text{ and } m \in N_{j_1} \text{ and } n \neq n_1 \text{ or } m \neq m_1.$$

■

For an integer programming problem, even a relatively small number of additional binary variables can greatly impact the tractability of the problem. By relaxing the integrality of the connection variables, we reduce significantly the number of binary variables.

The CFSR model can be re-written equivalently by replacing 4.13 with:

$$0 \leq x_{i,n}^{j,m} \leq 1 \quad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j. \tag{4.14}$$

4.3.2 The Aggregate Connection-Based Model

The relationship between the variables x and the variables y can be modeled in different ways, leading to different,but equivalent models, with different properties. One way to model the relationship between the variables x and y is to aggregate constraints 4.10 and 4.11 in the CFSR model. Instead of considering the relationship between x and y for each pair of flights with connecting passengers, we consider the relationship for a flight and all other flights with passengers connecting to or from this flight. This Aggregated Connection-based Flight Schedule Retiming Model (ACFSR) is formulated as:

$$\min \sum_{i \in F^O} \sum_{n \in N_i} \sum_{j \in C^+(i)} \sum_{m \in N_j} x_{i,n}^{j,m} \times E[dp_{i,n}^{j,m}] \tag{4.15}$$

subject to

$$\sum_{n \in N_i} f_{i,n} = 1 \qquad \forall i \in F^C; \tag{4.16}$$

$$\sum_{j \in C^+(i)} \sum_{m \in N_j} x_{i,n}^{j,m} = \left|C^+(i)\right| f_{i,n} \qquad \forall i \in F^O, n \in N_i; \tag{4.17}$$

$$\sum_{i \in C^-(j)} \sum_{n \in N_i} x_{i,n}^{j,m} = \left|C^-(j)\right| f_{j,m} \qquad \forall j \in F^I, m \in N_j; \tag{4.18}$$

$$f_{i,n} \in \{0,1\} \qquad \forall i \in F^C, n \in N_i; \tag{4.19}$$

$$0 \leq x_{i,n}^{j,m} \leq 1 \quad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j. \tag{4.20}$$

This model, as shown below, is equivalent to CFSR.

Theorem 3 *ACFSR model is equivalent to CFSR model.*

Proof. We first prove that the constraints 4.10 and 4.11 are equivalent to the constraints 4.17 and 4.18.

Simply aggregating the $\sum_{m \in N_j} x_{i,n}^{j,m} = f_{i,n}$, over all $j \in C^+(i)$, we obtain

$$\sum_{j \in C^+(i)} \sum_{m \in N_j} x_{i,n}^{j,m} = \left|C^+(i)\right| f_{i,n}, \forall i \in F^O, n \in N_i.$$

Similarly, aggregating 4.11 we obtain 4.18. This means that the constraints 4.10 and 4.11 imply the constraints 4.17 and 4.18.

Next, we show that any f and x satisfying 4.17 also satisfy 4.10:

If $f_{i,n} = 0$, we have $x_{i,n}^{j,m} = 0, \forall j \in C^+(i), m \in N_j$, according to constraints 4.17 and $x_{i,n}^{j,m} \geq 0$. Hence, for any flight leg $j \in C^+(i)$, we have

$$\sum_{m \in N_j} x_{i,n}^{j,m} = f_{i,n}, \forall i \in F^O, n \in N_i.$$

According to constraints 4.17, if $f_{i,n} = 1$, then

$$\sum_{j \in C^+(i)} \sum_{m \in N_j} x_{i,n}^{j,m} = \left| C^+(i) \right|.$$

Because $x_{i,n}^{j,m} \in \{0, 1\}$, we have $\sum_{m \in N_j} x_{i,n}^{j,m} = 1$, $\forall j \in C^+(i)$. This implies

$$\sum_{m \in N_j} x_{i,n}^{j,m} = f_{i,n}, \forall i \in F^O, n \in N_i.$$

Similarly, any f and x satisfying 4.18 also satisfy 4.11.

Because the constraints 4.10 and 4.11 are equivalent to the constraints 4.17 and 4.18, and all other constraints and the objective function are the same in both models, ACFSR model is equivalent to CFSR model. ■

4.3.3 The Disaggregate Connection-based Model

An alternative retiming model disaggregate constraints 4.10 and 4.11 and consider each individual connection copy. The model, the Disaggregate Connection-Based Flight Schedule Retiming Model (DCFSR), is formulated as follows:

$$\min \sum_{i \in F^O} \sum_{n \in N_i} \sum_{j \in C^+(i)} \sum_{m \in N_j} x_{i,n}^{j,m} \times E[dp_{i,n}^{j,m}] \tag{4.21}$$

subject to

$$\sum_{n \in N_i} f_{i,n} = 1 \qquad \forall i \in F^C; \tag{4.22}$$

$$x_{i,n}^{j,m} \geq f_{i,n} + f_{j,m} - 1 \qquad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j; \tag{4.23}$$

$$f_{i,n} \in \{0,1\} \qquad \forall i \in F^C, n \in N_i; \tag{4.24}$$

$$0 \leq x_{i,n}^{j,m} \leq 1 \qquad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j. \tag{4.25}$$

This model, as shown below, is also equivalent to CFSR.

Theorem 4 *DCFSR model is equivalent to CFSR model.*

Proof. We first prove that the constraints 4.10 and 4.11 are equivalent to the constraints 4.23.

Consider any pair of flights i_1 and j_1 with connecting passengers. Combining $\sum_{n \in N_i} f_{i_1,n} = 1$ with $\sum_{m \in N_j} x_{i_1,n}^{j_1,m} = f_{i_1,n}$ implies that, for any solution of CFSR,

$$\sum_{n \in N_i} \sum_{m \in N_j} x_{i_1,n}^{j_1,m} = \sum_{n \in N_i} f_{i_1,n} = 1.$$

Moreover, for any copy n_1 of flight i_1 and copy m_1 of flight j_1, we have

$$f_{i_1,n_1} + f_{j_1,m_1} = \sum_{m \in N_j} x_{i_1,n_1}^{j_1,m} + \sum_{n \in N_i} x_{i_1,n}^{j_1,m_1},$$

and

$$f_{i_1,n_1} + f_{j_1,m_1} \leq \sum_{n \in N_i} \sum_{m \in N_j} x_{i_1,n}^{j_1,m} + x_{i_1,n_1}^{j_1,m_1} = 1 + x_{i_1,n_1}^{j_1,m_1}.$$

Thus, for any solution to CFSR, we have that

$$x_{i_1,n_1}^{j_1,m_1} \geq f_{i_1,n_1} + f_{j_1,m_1} - 1, \forall i_1 \in F^O, n_1 \in N_{i_1}, j_1 \in C^+(i_1), m_1 \in N_{j_1}.$$

Next, we show that any integer solution satisfies 4.23 also satisfies 4.10 and 4.11.

Both CFSR and DCFSR models ensure that, for each flight, only one copy is selected. Without losing generality, assume that $f_{i_1,n_1} = 1$, $f_{j_1,m_1} = 1$, then $f_{j_1,m\neq m_1} = 0, \forall m_1 \in N_{j_1}$, and $f_{i_1,n\neq n_1} = 0, \forall n_1 \in N_{i_1}$. According to constraints 4.23 , we have

$$x_{i_1,n_1}^{j_1,m_1} \geq f_{i_1,n_1} + f_{j_1,m_1} - 1 = 1.$$

In conjunction with $0 \leq x_{i_1,n_1}^{j_1,m_1} \leq 1$, we have $x_{i_1,n_1}^{j_1,m_1} = 1$ and $x_{i_1,n}^{j_1,m} = 0$, $\forall n \neq n_1$ or $m \neq m_1$,because this is a minimization problem with positive objective function coefficients. For flight $i_1$copy n_1, and flight j_1, we have

$$\begin{aligned} \sum_{m\in N_j} x_{i_1,n_1}^{j_1,m} &= x_{i_1,n_1}^{j_1,m_1} + \sum_{m\in N_j, m\neq m_1} x_{i_1,n_1}^{j_1,m} \\ &= 1 = f_{i_1,n_1}. \end{aligned}$$

For flight i_1 copy $n \neq n_1$, and flight j_1, we have $f_{i_1,n\neq n_1} = 0$ and

$$\sum_{m\in N_j} x_{i_1,n\neq n_1}^{j_1,m} = 0.$$

Thus, we have $\sum_{m\in N_j} x_{i_1,n}^{j_1,m} = f_{i_1,n}$, for any copy n. Similarly, we have $\sum_{n\in N_i} x_{i_1,n}^{j_1,m} = f_{j_1,m}$, for any copy m.

Because the constraints 4.10 and 4.11 are equivalent to the constraints 4.23, and all other constraints and the objective function are the same in both models, ACFSR model is equivalent to CFSR model. ■

4.4 Sizes of Models

We compute the sizes of CFSR, ACFSR, DCFSR models for a major US airline with 2,044 flight legs and 76,641 itineraries. Suppose that 7 copies are generated for each flight leg with 5 minute intervals (7 copies correspond to a 30 minute time window) and

Models	Num. of Var.	Num. of Integer Var.	Num. of Constraints
CFSR	1,216,180	14,308	345,436
ACFSR	1,216,180	14,308	30,660
DCFSR	1,216,180	14,308	1,203,916

Table 4.2: Comparasion of the Problem Sizes for Flight-Based Formulations.

on average every flight leg has passengers that are connecting to 12 other flights legs (This is typical for this airline). The sizes of the models are summarized in Table 4.2. As expected, compared to CFSR model, ACFSR model has far fewer constraints and DCFSR model has far more.

4.5 Quality of the bounds of the models' LP Relaxations

The LP relaxation of integer or mixed integer programming problems is a lower bound on the optimal solution value. In a Branch-and-bound algorithm, the quality of this lower bound is very important. The sharper the bound, the better the algorithm. In fact, the quality of this bound can greatly impact the computational performance of the solution algorithm, especially when solving a large-scale problem.

We first define the *strength* of LP relaxations. Consider two LP relaxations A and B of two minimization integer programming formulations. A is at least as strong as B, if $Z_A \leq Z_B$, where Z is the objective function value.

Theorem 5 *The LP relaxation of CFSR is at least as strong as that of ACFSR, and can be stronger in some instances.*

Proof. First, we show that the LP relaxation of CFSR is at least as strong as that of ACFSR. We define the following two polyhedra, which are the feasible sets of the LP relaxations for CFSR and ACFSR:

$$P^{LP}_{CFSR} = \{(\mathbf{x},\mathbf{f})| \quad \begin{array}{c} \sum_{n\in N_i} f_{i,n} = 1, \quad \forall i \in F^C; \\ \sum_{m\in N_j} x^{j,m}_{i,n} = f_{i,n}, \forall i \in F^O, n \in N_i, j \in C^+(i); \\ \sum_{n\in N_i} x^{j,m}_{i,n} = f_{j,m}, \forall j \in F^I, m \in N_j, i \in C^-(j); \\ 0 \leq f_{i,n} \leq 1 \quad \forall i \in F^C, n \in N_i; \\ 0 \leq x^{j,m}_{i,n} \leq 1 \quad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j.\} \end{array}$$

$$P^{LP}_{ACFSR} = \{(\mathbf{x},\mathbf{f})| \quad \begin{array}{c} \sum_{n\in N_i} f_{i,n} = 1, \quad \forall i \in F^C; \\ \sum_{j\in C^+(i)} \sum_{m\in N_j} x^{j,m}_{i,n} = |C^+(i)|\, f_{i,n} \quad \forall i \in F^O, n \in N_i; \\ \sum_{i\in C^-(j)} \sum_{n\in N_i} x^{j,m}_{i,n} = |C^-(j)|\, f_{j,m} \quad \forall j \in F^I, m \in N_j; \\ 0 \leq f_{i,n} \leq 1 \quad \forall i \in F^C, n \in N_i; \\ 0 \leq x^{j,m}_{i,n} \leq 1 \quad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j.\} \end{array}$$

It is straightforward to show that any feasible solution in P^{LP}_{CFSR} is feasible for P^{LP}_{ACFSR}.Simply aggregate the constraints

$$\sum_{m\in N_j} x^{j,m}_{i,n} = f_{i,n}, \ \forall i \in F^O, n \in N_i, j \in C^+(i)$$

over all $j \in C^+(i)$ and the constraints

$$\sum_{n\in N_i} x^{j,m}_{i,n} = f_{j,m}, \forall j \in F^I, m \in N_j, i \in C^-(j)$$

over all $i \in C^-(j)$ in P^{LP}_{CFSR} to obtain the corresponding constraints in P^{LP}_{ACFSR}. Let Z^{LP}_{CFSR} and Z^{LP}_{ACFSR} denote the optimal solution values of the LP relaxations of CFSR and ACFSR respectively. The objective function for both minimization models is the same, and $P^{LP}_{CFSR} \subseteq P^{LP}_{ACFSR}$, it follows that

$$Z^{LP}_{ACFSR} \leq Z^{LP}_{CFSR}.$$

Next, we show that the LP relaxation of CFSR can be stronger than that of ACFSR for some instances. Consider a network with four flights l, i, j, and k, each with two

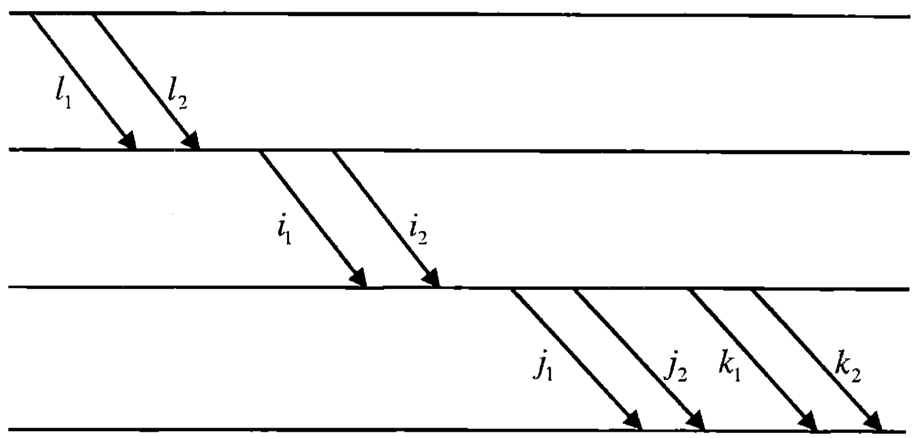

Figure 4-5: A Network to Prove Bounds of LP Relaxations of CFSR and DCFSR

Connection	l_1, i_1	l_1, i_2	l_2, i_1	l_2, i_2
$E[DP_{i,n}^{j,m}]$	8	2	10	8
Connection	i_1, j_1	i_1, j_2	i_2, j_1	i_2, j_2
$E[DP_{i,n}^{j,m}]$	6	3	8	6
Connection	i_1, k_1	i_1, k_2	i_2, k_1	i_2, k_2
$E[DP_{i,n}^{j,m}]$	2	1	3	2

Table 4.3: The Expected Numbers of Disrupted Passengers.

copies (See Figure 4-5). The expected numbers of disrupted passengers between each pair of flights are summarized in Table 4.3.

An optimal solution for the LP relaxation of ACFSR is $x_{l,1}^{i,2} = x_{l,1}^{i,1} = 0.5$, $x_{i,1}^{j,2} = x_{i,2}^{k,2} = 1$ and all other $x = 0$, $f_{l_1} = f_{j_2} = f_{k_2} = 1$, $f_{i_1} = f_{i_2} = 0.5$ and all other $f = 0$. The optimal solution value is 10. This solution, however, is infeasible for CFSR. In fact,the optimal solution for CFSR is $f_{l_1} = f_{i_1} = f_{j_2} = f_{k_2} = 1$, $x_{l,1}^{i,1} = x_{i,1}^{j,2} = x_{i,1}^{k,2} = 1$,and all other variables equal zero. The corresponding optimal solution value is 12. ■

Theorem 6 *The LP relaxation of CFSR is at least as strong as that of DCFSR and can* *be stronger in some instances.*

Proof. First, we show that the LP relaxation of CFSR is at least as strong as that of DCFSR. Denote the feasible set of the LP relaxation of DCFSR as follows:

$$P^{LP}_{ACFSR} = \{(\mathbf{x},\mathbf{f})| \quad \textstyle\sum_{n\in N_i} f_{i,n} = 1, \quad \forall i \in F^C;$$
$$x^{j,m}_{i,n} \geq f_{i,n} + f_{j,m} - 1 \quad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j;$$
$$0 \leq f_{i,n} \leq 1 \quad \forall ii \in F^C, n \in N_i;$$
$$0 \leq x^{j,m}_{i,n} \leq 1 \quad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j.\}$$

Consider any pair of flights i_1 and j_1 with connecting passengers. Combining $\sum_{n\in N_i} f_{i_1,n} = 1$ with $\sum_{m\in N_j} x^{j_1,m}_{i_1,n} = f_{i_1,n}$ implies that, for any feasible solution to the CFSR LP,

$$\sum_{n\in N_i}\sum_{m\in N_j} x^{j_1,m}_{i_1,n} = \sum_{n\in N_i} f_{i_1,n} = 1.$$

Moreover, feasible solution to the CFSR LP satisfies

$$f_{i_1,n_1} + f_{j_1,m_1} = \sum_{m\in N_j} x^{j_1,m}_{i_1,n_1} + \sum_{n\in N_i} x^{j_1,m_1}_{i_1,n}$$

Considering these, we get

$$f_{i_1,n_1} + f_{j_1,m_1} \leq \sum_{n\in N_i}\sum_{m\in N_j} x^{j_1,m}_{i_1,n} + x^{j_1,m_1}_{i_1,n_1} = 1 + x^{j_1,m_1}_{i_1,n_1}.$$

Thus,for any feasible solution to the CFSR LP, we have that

$$x^{j_1,m_1}_{i_1,n_1} \geq f_{i_1,n_1} + f_{j_1,m_1} - 1, \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j.$$

Thus, $P^{LP}_{CFSR} \subset P^{LP}_{DCFSR}$. Z^{LP}_{DCFSR} denote the optimal solution value of the DCFSR LP relaxation. Similar to the proof of Theorem 5, both CFSR and DCFSR have the same LP objective function, hence:

$$Z_{DCFSR}^{LP} \leq Z_{CFSR}^{LP}.$$

Next, we show that the LP relaxation of CFSR can be stronger than that of DCFSR for some instances. Consider a network with two flights i and j, each with two copies and passengers are connecting from flight i to flights j. A solution with $f_{i,1} = f_{i,2} = f_{j,1} = f_{j,2} = 0.5$ and all $x = 0$ is feasible for ACFSR, the corresponding solution value of the LP relaxation of DCFSR is zero, an optimal solution. This solution, however, is infeasible for the LP relaxation of CFSR, because $\sum_{m \in N_j} x_{i,n}^{j,m} = 0 \neq f_{i,n} = 0.5$. An optimal solution for the LP relaxation of CFSR must have some $x > 0$, because $\sum_{m \in N_j} x_{i,n}^{j,m} = f_{i,n}$ and $\exists n$ such that $f_{i,n} > 0$. This leads to a solution value greater than zero, because the objective function coefficients are positive. ■

Note that if we add the constraints $\sum_{n \in N_i} \sum_{m \in N_j} x_{i_1,n}^{j_1,m} = 1$ to DCFSR model, the bound can be improved. These results show that the CFSR model is dominates the DCFSR model. It has a LP relaxation that is at least as strong or stronger than that of the DCFSR model and it has same number of variables and many fewer constraints. Therefore, we implement only the CFSR and the ACFSR models.

4.6 Solution Approach

4.6.1 Overview of the Solution Approach

The CFSR and ACFSR formulations are deterministic mixed-integer programs with a large number of variables. For practical problems, complete generation of all variables will require prohibitive amounts of time and memory. Thus we solve these problems using branch-and-price(see Section 3.4 for a detailed description).

4.6.2 Branching Strategy

After solving an LP relaxation at a node of the branch-and-bound tree, we must decide which branching constraints should be added. We branch based on the cover constraints:

$$\sum_{n \in N_i} f_{i,n} = 1 \qquad \forall i \in F^C.$$

Building on the results of Hane et al. (1995) [42], we employ special ordered set branching in which we divide the set of variables $f_{i,n}$ for each flight leg i into two sets. We force the sum of the variables in the first set to equal one on one branch and the sum of the variables in the second set to equal one on the other branch. For fleet assignment problem, Hane et al. (1995) [42] show that this is a more effective branching strategy than branching on individual variables.

4.6.3 Column Generation

At an iteration of the column generation algorithm, let $\pi^j_{i,n}$ be the optimal dual variables associated with constraints 4.9 in CFSR, $\pi^j_{i,n}$ be the optimal dual variables associated with constraints 4.10 and $\pi^{j,m}_i$ be the optimal dual variables associated with constraints 4.11. Then the reduced cost for each connection copy between flights i and j is

$$\overline{dp^{j,m}_{i,n}} = dp^{j,m}_{i,n} - \pi^j_{i,n} - \pi^{j,m}_i. \tag{4.26}$$

Because the number of columns is just over one million for a typical airline problem, a large but manageable number, the reduced cost for each copy of each connection can be calculated explicitly and all columns with negative reduced costs are added to the restricted master problem at each iteration.

4.7 Proof-of-Concept

4.7.1 Underlying Networks

For the computational experiments with our retiming models, we combine the four networks (described in Table 3.2) to form one network with a total of 278 flights. because there are many passengers connecting to or from this network, we also consider flights in the full airline network to or from which passengers using the flight legs in the 278 flight network connect. For these additional flights, we fix the current schedule. The total number of flight legs considered in this expanded network is 1067.

4.7.2 Connection Variables

There are many variables and constraints associated with flights without copies. It would be good to generate variables only for flights with copies, which can reduce the problem size significantly. In fact, we can just generate variables only for flights with copies, by adding objective function coefficients to f variables. Let $O(i)$ be the set of flights without copy but with passengers connecting out from flight i, $I(i)$ be the set of flights without copy but with passenger connecting into flight i. Let $dp_{i,n}$ be the number of disrupted passengers associated with flight i copy n, and these passengers are either connecting into flight i from another flight without copy or connecting out from flight i to another flight without copy. Let $dp^k_{i,n}, k \in O(i)$ be the number of disrupted passengers connecting from flight i copy n to flight k without copy; $dp^{i,n}_k, k \in I(i)$ be the number of disrupted passengers connecting from flight k with copy to flight i copy n. Then we have

$$dp_{i,n} = \sum_{k \in O(i)} dp^k_{i,n} + \sum_{k \in I(i)} dp^{i,n}_k.$$

The objective function can be re-written as follows:

$$\min \sum_{i,n,j,m} x^{j,m}_{i,n} \times E[dp^{j,m}_{i,n}] + \sum_{i,n} f_{i,n} \times E[dp_{i,n}].$$

Because $E[dp_{i,n}]$ is a constant for each flight i copy n, we can determine it off-line before solving the problem.

4.7.3 Data and Validation

We have the July and August 2000 data for a major U.S. airline consisting of:

- ASQP data: provides the following flight operation information for each flight: planned departure time and arrival time, actual departure time and arrival time (including wheels-off and wheels-on time, taxi-out and taxi-in time, airborne time) and airplane tail number

- Airline data: include number of passengers on each itinerary.

To validate our model, we build the model based on historical information, and generate schedules and apply them to future operations. Suppose that we were at the end of July 2000, and we need to determine the schedule plans for next month, August 2000. We use our robust aircraft maintenance routing model (see Chapter 3) to obtain routing solution and delays for each flight are then determined based on this routing solution (recall that propagated delays are a function of routing). Based on these, the expected number of disrupted passengers for each connection copy is estimated using July 2000 data as described in Section 4.3.1. The sample average of the number of disrupted passengers is used as an approximation of the mean. Then we solve our flight schedule retiming models to obtain the final flight departure times for August 2000. Hence, given these routings and flight leg departure times generated from our model based on July 2000 data and delay information, we calculate the numbers of disrupted passengers for August 2000 based on the new schedule generated by our model and the existing routing from the historical data. We then compare these numbers.

Models	Num of constraints	Num of variables	Num of non-zeros
ACFSR	1,990	27,013	54,320
CFSR	7,506	27,013	59,836

Table 4.4: Comparasion of the Sizes of the RFRS Model and the RAFRS Model.

Models	Value at node 0	Optimal value	Num of nodes searched	Time to solve
ACFSR	10,437	10,899	≥43,590 (out of memory)	$\geq 386,756$ sec
CFSR	10,899	10,899	1	13 sec

Table 4.5: Comparasion of the Strength of the CFSR Model and the ACFSR Model.

4.7.4 Computational Results

The results obtained by applying our flight schedule retiming models to the network of a major U.S. airline (described in Section 4.7.1)are presented below. Problems are solved using CPLEX 6.5 on a HPC 3000 machine with 1G RAM.

Sizes and Bounds

First, we compare the size of the CFSR and ACFSR models. Using a 30 minute time window allowing flights to depart at most 15 minutes earlier or later than originally scheduled, we generate copies for flight arcs every five minutes, for a total of 7 copies in each flight leg's time window. The numbers of constraints, variables and non-zeros in the CFSR and ACFSR models, shown in table 4.4, are consistent with our earlier analysis. Also consistent are the results reported in table 4.5. The LP bound of the ACFSR model is very loose, with a very fractional solution. After searching 43,590 nodes in the branch-and-bound tree, solution algorithm fails to find an optimal solution because of memory limitations. In contrast, the LP relaxation of the CFSR model is very tight. For this problem instance, an optimal solution is found at the root node of the branch-and-bound tree. For the same data, the ACFSR model did not find an optimal even after a runtime of 5-days, while the CFSR model found an optimal solution within 13 seconds.

Time window	Old D-pax	New D-pax	D-pax reduced	D-pax reduced (%)
±15min(7 copies)	17,459	10,899	6,560	37.6%
±10min(5 copies)	17,459	12,070	5,389	30.9%
±5min(3 copies)	17,459	14,069	3,390	19.4%

Table 4.6: Effects of Retiming on Numbers of Disrupted Passengers (July 2000 Data).

Misconnections and Time Window Width

We also determined the number of passenger misconnections that can be avoided through retiming. The results are presented in Tables 4.6 and 4.7. *Time Window* indicates the total amount time (in minutes) flight legs are allowed to shift and the number of copies of flight legs generated in this time window. For example, ±15*min(7 copies)* represents each flight leg to depart at most 15 minutes earlier or later than originally scheduled. Because we generate copies for flight arcs every five minutes, there are 7 copies in this time window. *Old D-pax* indicates the total number of passenger misconnections in the original schedule and *New D-pax* indicates the number of passenger misconnections in our new schedule. *D-pax reduced* and *D-pax reduced (%)* indicate the difference in the number (and percentage) of passenger misconnections between the old and new schedule. Note that, in our computational experiment, we consider only those passengers whose itineraries have at least one flight leg included in the subnetwork with 278 flights. The disruption status of all other passengers is unchanged by our retiming solution.

For July 2000 data, allowing flight leg departures to shift within 30 minute time windows (15 minutes earlier or later), the total percentage of passenger misconnections is reduced by 37.6% using our new schedule rather than the original schedule. Moreover, even if time windows are reduced to 10 minutes, about 20% fewer passengers miss their connections.

There is a caveat associated with these results, because our model parameters are determined using the July 2000 data. The above results are not replicatable in practice because they rely on perfect knowledge of relevant future events. In practice, models must be built using historical data, and then applied for future operations. Thus, we take our

Time window	Old D-pax	New D-pax	D-pax reduced	D-pax reduced (%)
±15min(7 copies)	18,808	11,348	7,460	39.7%
±10min(5 copies)	18,808	12,732	6,076	32.3%
±5min(3 copies)	18,808	15,042	3,766	20.0%

Table 4.7: Effects of Retiming on Numbers of Disrupted Passengers (August 2000 Data).

retiming decisions based on July 2000 data and apply them to the August 2000 flight network. Our results are summarized in Table 4.7. If flight departure times are allowed to shift in a thirty-minute time window, about 40% fewer passengers miss their connections. A twenty-minute time window reduce the number of passenger misconnections by over 30%, while a ten-minute time window reduce it by 20%.

Effects of Minimum Connection Time

All the previous results are based on the assumption that the minimum connection time for passengers is 30 minutes. We summarize the effects of connection time and time window width on the numbers of passenger misconnections in Tables 4.8 and 4.9. Using July 2000 data and retiming the minimum connection time equal 20 and then 25 minutes, the results are summarized as:

1. The numbers of passenger misconnections for both the actual and retimed schedules are reduced, for all time window widths. This is an expected result given the definition of a disrupted passenger (in Section 4.1). If the required connection time is shortened, a passenger is more likely to make his/her connection.

2. With shortened connection times, the percentage reduction in the number of passenger misconnections in the retimed schedule is reduced. This likely follows because many of the disrupted passengers in the original schedule misconnection by more than 30 minutes and hence are unrecoverable through these retiming models. Nonetheless, there is a reduction in the number of passenger misconnections between 10% to 30%, even with these "short" time windows and connection times.

Time window	Old D-pax	New D-pax	D-pax reduced	D-pax reduced (%)
±15min(7 copies)	14,199	9,866	4,333	30.5%
±10min(5 copies)	14,199	10,778	3,421	24.1%
±5min(3 copies)	14,199	12,026	2,173	15.3%

Table 4.8: 25 minute Minimum Connection Time (July 2000 Data).

Time window	Old D-pax	New D-pax	D-pax reduced	D-pax reduced (%)
±15min(7 copies)	12,090	9,148	2,942	24.3%
±10min(5 copies)	12,090	9,812	2,278	18.8%
±5min(3 copies)	12,090	10,767	1,323	10.9%

Table 4.9: 20 minute Minimum Connection Time (July 2000 Data).

For August 2000 data, we obtained similar results presented in Tables 4.10 and 4.11, for our routing and retiming decisions based on July data. There is a significant reduction in the number of passenger misconnections (from 15% to 33% for 25 minute minimum connection times, and between 12% to 27% for 20 minute minimum connection time).

Effects of Copy Interval

For given time window width, another parameter to determine is the optimal number of copies to generate for each flight leg, that is, the optimal time interval between flight copies. These results are obtained by assuming a 30 minute minimum connection time and based July 2000 data.

In Table 4.12, we provide results of our analysis in which we assumed a minimum of

Time window	Old D-pax	New D-pax	D-pax reduced	D-pax reduced (%)
±15min(7 copies)	15,102	10,144	4,958	32.8%
±10min(5 copies)	15,102	11,237	3,865	25.6%
±5min(3 copies)	15,102	12,753	2,349	15.6%

Table 4.10: 25 minute Minimum Connection Time (August 2000 data).

Time window	Old D-pax	New D-pax	D-pax reduced	D-pax reduced (%)
±15min(7 copies)	12,724	9,275	3,449	27.1%
±10min(5 copies)	12,724	10,054	2,670	21.0%
±5min(3 copies)	12,724	11,107	1,617	12.7%

Table 4.11: 20 minute Minimum Connection Time (August 2000 data).

Time window	Num of constrs	Num of vars	Num of non-zeros	Increase
±15min(7 copies)	7,506	27,013	59,836	1.0
±15min(31 copies)	32,514	507,253	1,040,236	17.4
±10min(5 copies)	5,422	14,085	32,320	1.0
±10min(21 copies)	22,094	234,213	485,856	15.0
±5min(3 copies)	3,338	5,325	13,140	1.0
±5min(11 copies)	11,674	65,373	139,876	10.6

Table 4.12: Comparison of the Problem Sizes (5 min minute copy interval vs. 1 minute copy interval).

Time window	Old D-pax	New D-pax	D-pax reduced	Improve (%)
±15min(7 copies)	17,459	10,899	6,560 (37.6%)	0.0
±15min(31 copies)	17,459	10,865	6,594 (37.8%)	0.52%
±10min(5 copies)	17,459	12,070	5,389 (30.9%)	0.0
±10min(21 copies)	17,459	12,056	5,403 (30.9%)	0.26%
±5min(3 copies)	17,459	14,069	3,390 (19.4%)	0.0
±5min(11 copies)	17,459	14,058	3,401 (19.5%)	0.28%

Table 4.13: Comparison of numbers of Disrupted Passengers (5 minute copy Interval vs. 1 minute copy Interval).

connection time of 30 minutes and varied the flight leg copy interval in time windows of various width. *Increase* indicates the factor increase in the numbers of non-zeros in the model compared to a base case with 5 minute copy interval. In Table 4.13, *Improve* indicates the percentage reduction in the number of disrupted passengers, again compared to a 5 minute copy interval. Generating copies for flight legs every minute results in dramatically increased problem sizes and modest benefit. By placing copies more sparsely, we improve model tractability considerably and obtain solutions that are nearly as good.

Given this result, we consider the case of using longer copy intervals of 15 minutes for a 30 minute time window and 10 minutes for a 20 minute time window. The results are summarized in Tables 4.14 and 4.15. Generating fewer copies further reduce problem size, while achieving about the same number of disrupted passengers as achieved for shorter copy intervals. This is because in most cases the model selects flight copies that either represent the original departure time or push the flight departure time as far as possible. This is what we expect. Imagining that for two flights with passenger connecting, the model will push the departure time of the first flight to an earliest possible time and the departure time of the second flight to a latest possible time, which provides more connecting time for passengers.

In summary, if the problem size is a concern (sometimes just loading a large-scale problem to a computer may be impossible because of memory limitations), then by generating copies at a broad interval (15 or 10 minutes depending on the time window), we can solve a much smaller problem and still obtain good solutions. Of course, if the problem size is not a concern, for example, when using our RFRS model and the problem instance is moderate, then narrower time interval will lead to more savings.

Estimating the Impact on Passenger Delays

In this section, we estimate roughly the impact of our model on total passenger delays. As shown in Table 4.1, the delay experienced by disrupted passengers is 51% of total

Time window	Num of constrs	Num of vars	Num of non-zeros	Increase
±15min(7 copies)	7,506	27,013	59,836	4.6
±15min(3 copies)	3,338	5,325	13,140	1.0
±10min(5 copies)	5,422	14,085	32,320	2.5
±10min(3 copies)	3,338	5,325	13,140	1.0

Table 4.14: Comparasion of the Problem Sizes (5 min, 10 min, 15 min intervals).

Time window	Old D-pax	New D-pax	D-pax reduced	Improve
±15min(7 copies)	17,459	10,899	6,560 (37.6%)	0.76%
±15min(3 copies)	17,459	10,949	6,510 (37.3%)	0.0
±10min(5 copies)	17,459	12,070	5,389 (30.9%)	0.33%
±10min(3 copies)	17,459	12,088	5,371 (30.8%)	0.0

Table 4.15: Comparasion of the Results (5 min, 10 min, 15 min Intervals).

passenger delays (in minutes). In applying our model and assuming the minimum connecting time for passengers is 30 minutes; flight copies are generated every five minutes within 30 minute time window; and every disrupted passenger is delayed for 419 minutes, the average delay for disrupted passengers; then a reduction of about 40% in the total number of disrupted passengers leads to roughly 20% decrease in total passenger delay. Moving from thirty to twenty-minute time window reduces the decrease in delay minutes to roughly 16%, while a ten-minute time window achieve a reduction of roughly 10%. Importantly, our models especially reduce delays for passengers who would otherwise be excessively delayed, that is, disrupted. As stated earlier, the average delay for nondisrupted passengers is only 14 minutes while the average delay for disrupted passengers is 419 minutes.

Chapter 5

Extensions

5.1 Integrated Robust Aircraft Maintenance Routing and Fleet Assignment

The string-based model proposed by Barnhart et al. (1998) [6] can solve fleet assignment and maintenance routing problems at the same time. Similarly, one extension for our robust aircraft maintenance routing model is to adopt it to solve integrated fleet assignment and maintenance routing. Adding fleeting decisions provide more feasible strings, potentially leading to an improved solution with reduced delay propagation. However, although the financial benefit of aircraft maintenance routing is very limited, the fleet assignment costs are very significant. Therefore, when solving integrated fleet assignment and maintenance routing, it is inappropriate to minimize delay propagation without considering fleet assignment costs. Instead, one can estimate the costs for delay propagation, add them to the cost of fleet assignment, and minimize total cost. However, the cost for delay propagation is hard to determine accurately. Applying an idea similar to that proposed by Rosenberger et al. (2001) [67], we present two integrated models for robust aircraft maintenance routing and fleet assignment.

The notation we use here is the same as used in Section 3.3.4 except the following.

Let K be the set of fleets and G^k be the set of ground variables for fleet type k. We have one binary decision variable x_s^k for each feasible string s, which is equal to 1 if this string is flown by fleet type k; and 0 otherwise. c_s^k is the cost (both operating and spill) of flying string s with fleet type k. We have ground variables, denoted y^k, which are used to count the number of aircraft of fleet type k on the ground at maintenance stations. Ground variables $y_{i,d}^{-,k}$ and $y_{i,d}^{+,k}$ equal the number of aircraft of fleet type k on the ground just before and after, respectively, flight i departs. Similarly ground variables $y_{i,a}^{-,k}$ and $y_{i,a}^{+,k}$ equal the number of aircraft of fleet type k on the ground just before and just after flight i arrives. r_s^k is the number of times string s with fleet type k crosses the count time; p_g^k is the number of times ground arc g with fleet type k crosses the count time; and N_k is the number of planes available for fleet type k. pd_{ij}^{sk} is the delay propagated from flight i to flight j if flight i and flight j are in string s and string s is assigned to fleet type k. pd^U is an upper bound on the total expected propagated delay, and c^U is an upper bound on the total fleet assignment and aircraft routing cost.

The string-based integrated fleet assignment and aircraft maintenance routing model (SFAMMR) is written as (Barnhart et al. (1998) [6]):

$$\min \sum_{k \in K} \sum_{s \in S} c_s^k x_s^k \tag{5.1}$$

Subject to

$$\sum_{k \in K} \sum_{s \in S} a_{is} x_s^k = 1 \qquad \forall i \in F; \tag{5.2}$$

$$\sum_{s \in S_i^+} x_s^k - y_{i,d}^{-,k} + y_{i,d}^{+,k} = 0 \qquad \forall i \in F, k \in K; \tag{5.3}$$

$$-\sum_{s \in S_i^-} x_s^k - y_{i,a}^{-,k} + y_{i,a}^{+,k} = 0 \qquad \forall i \in F, k \in K; \tag{5.4}$$

$$\sum_{s\in S} r_s^k x_s^k + \sum_{g\in G} p_g^k y_g^k \leq N^k \qquad \forall k \in K; \tag{5.5}$$

$$y_g^k \geq 0 \qquad \forall g \in G^k, k \in K; \tag{5.6}$$

$$x_s^k \in \{0,1\} \qquad \forall s \in S, k \in K. \tag{5.7}$$

The first set of constraints are the cover constraints, forcing every flight to be contained in exactly one string, and hence assigning each flight to exactly one fleet type. The second and third set of constraints are the aircraft flow balance constraints for each fleet. The fourth set of constraints are the fleet count constraints that ensure the number of aircraft used in each fleet type is not more than that available.

Let X be the set of feasible solutions satisfying the above constraints. Two integrated models for robust aircraft maintenance routing and fleet assignment are presented as follows.

$$\min \sum_{k\in K}\sum_{s\in S} c_s^k x_s^k \tag{5.8}$$

Subject to

$$\sum_{k\in K}\sum_{s\in S}\sum_{(i,j)\in s} E[pd_{ij}^{sk}]x_s^k \leq pd^U; \tag{5.9}$$

$$x \in X. \tag{5.10}$$

In this model, we minimize the total fleet assignment and maintenance routing costs but constrain the total expected propagated delay to pd^U, an upper bound on total propagated delay. Our next model minimizes the total expected propagated delay and constrains fleet assignment and maintenance routing costs to c^U, an upper bound.

$$\min \sum_{k\in K}\sum_{s\in S}(x_s^k \times \sum_{(i,j)\in s} E[pd_{ij}^{sk}]) \tag{5.11}$$

Subject to

$$\sum_{k\in K}\sum_{s\in S} c_s^k x_s^k \leq c^U; \tag{5.12}$$

$$x \in X. \tag{5.13}$$

The upper bound pd^U can be obtained by

$$pd^U = (1+\delta)Z_{pd}^{\min} \tag{5.14}$$

where δ is a small positive parameter value; and $Z_{pd}^{\min}$ is the minimum expected propagated delay. $Z_{pd}^{\min}$ can be obtained by solving the following problem:

$$\min \sum_{k\in K}\sum_{s\in S}\sum_{(i,j)\in s} E[pd_{ij}^{sk}]x_s^k \tag{5.15}$$

Subject to

$$x \in X. \tag{5.16}$$

The upper bound c^U can be obtained by

$$c^U = (1+\epsilon)Z_c^{\min} \tag{5.17}$$

where ϵ is a small positive parameter value; and $Z_c^{\min}$ is the minimum fleet assignment and maintenance routing cost. It can be obtained by solving the SFAMMR model.

By controlling the values of δ and ϵ, we can find robust fleet assignment and maintenance routing solutions within specified ranges of the minimum costs or the minimum total expected propagated delay. These models can be solved using the same approach as proposed in Section 3.4.

5.2 Robust Aircraft Maintenance Routing with Time Window

In solving aircraft maintenance routing problems, allowing flights to be rescheduled within a small time window might produce a more robust routing solution, one that further reduces delay propagation. To model this, the string-based model with copies of each flight leg can be used.

The notation we use here is the same as that in Section 3.3.4 except that there are many more string variables, with potentially several strings containing the same flight legs, but different copies. Let $a^s_{i,n}$ equal one if flight i copy n is in string s; and equal zero otherwise, then an extension of the robust maintenance routing model presented in Chapter 3 is described as follows.

$$\min E(\sum_{s\in S}\sum_{(i,j)\in s} pd^s_{ij}x_s) \tag{5.18}$$

Subject to

$$\sum_{s\in S}\sum_{n\in N_i} a^s_{i,n}x_s = 1 \qquad \forall i \in F; \tag{5.19}$$

$$\sum_{s\in S_i^+} x_s - y^-_{i,d} + y^+_{i,d} = 0 \qquad \forall i \in F; \tag{5.20}$$

$$-\sum_{s\in S_i^-} x_s - y^-_{i,a} + y^+_{i,a} = 0 \qquad \forall i \in F; \tag{5.21}$$

$$\sum_{s \in S} r_s x_s + \sum_{g \in G} p_g y_g \leq N; \tag{5.22}$$

$$y_g \geq 0 \qquad \forall g \in G; \tag{5.23}$$

$$x_s \in \{0, 1\} \qquad \forall s \in S. \tag{5.24}$$

One difficulty of this model is that adding flight copies has an explosive effect on the number of strings. For example, consider a string consisting of four flight legs. For each flight leg, there are five copies generated including the original one. In this case, there are $5^4 = 625$ distinct strings, each containing the same set of flight legs. To reduce problem size, begin by observing that many of these copies of a string have the same propagated delay, because they have the same connection times. To illustrate this, let's take a look at a simple example. Suppose there is one string that consists of two flights (see Figure 4-4), and each flight has 7 copies. Then there are $7 \times 7 = 49$ strings with the same flight legs. Among these strings, many have the same delay propagation because the time between the two flight legs, that is the planned aircraft turn times is the same for some strings. For example, the string consisting of copy 1 of flight leg 1 and copy 1 of flight leg 2 has the same turn time as the one consisting of copy 2 of flight leg 1 and copy 2 of flight leg 2. For this string, there are $6+5+4+3+2+1+5+4+3+2+1 = 36$ copies that are duplicate for our purpose. This means there are only $49 - 36 = 13$ copies that have different values of propagated delays. In fact, we can prove that for any string consisting of two flight legs, each with n copies, the number of different strings is

$$n^2 - \sum_{i=1}^{n-1} i - \sum_{j=1}^{n-2} j. \tag{5.25}$$

We can also determine the number of different strings consisting of more than two flights.

We can integrate robust maintenance routing with time windows with fleet assignment, to add more robustness into the schedule plan. Likely, such a model will have tractability issues when solving large-scale problems. Research in this direction should focus on better formulations of the problem and/or new ways to reduce problem size and exploit problem structure.

5.3 Fleet Assignment with Time Window and Passenger Disruption Considerations

Rexing et al. (2000) [65] propose the idea of allowing flight leg departure time to be rescheduled within small time windows simultaneously with fleet assignment. They develop a model called the Fleet Assignment Model with Time Windows (FAMTW). It can lead to reductions in fleet assignment costs in two ways:

- a more appropriate aircraft type might be assigned to a flight leg because more aircraft connections are possible.
- aircraft can be utilized more efficiently, which can result in fewer aircraft required to fly the schedule.

They report annual savings of 23.9 million to 46.2 million dollars if twenty-minute time windows are allowed and 35.0 million to 77.7 million dollars if forty-minute time windows are allowed. Furthermore, by removing slack between flight connections, the schedule can be flown with two fewer aircraft, and fleet assignment cost is still less than the cost based on the original schedule. Obviously, the savings are significant, but one of the main sources of the frequent delays and disruptions in airline operations is that the schedule is planned very tightly, without sufficient slack to recover effectively from disruptions.

Recall in Chapter 4, we minimize the number of disrupted passengers by adding a time window for each flight. Integrating this model and the FAMTW model allows fleeting decisions to be affected by their impact on passenger disruptions. The difficulty is in dtermining the costs of passenger disruptions. Passenger disruptions result not only in re-accommodation costs but also costs associated with loss of goodwill. Thus, similar to what we have done in Section 5.1, we present the following two integrated models for balancing fleet assignment costs with improvements in passenger travel times.

Let's first define some new variables. Let O be the set of airports, K be the set of different fleet types, T be the sorted set of all event (departure or arrival) times at all airports, $CL(k)$ be the set of flight legs that pass the count time when flown by fleet type k, and $I(k,o,t)$ and $O(k,o,t)$ be sets of flights arriving and departing respectively from airport o at time t for fleet type k. $f_{i,n,k}$ is the binary variable that takes on value 1 if copy n of flight i is covered by fleet type k, and 0 otherwise. $c_{i,n,k}$ is the cost of assigning fleet type k to flight i copy n. y_{k,o,t^+} and y_{k,o,t^-} are the variables that count the number of aircraft of fleet type k at airport o just after and just before time t respectively; y_{k,o,t_n} are the variables that count the number of aircraft for fleet type k at airport o at the count time t_n.

The FAMTW model (Rexing et al. (2000), [65])can be written as follows:

$$\min \sum_{i\in F}\sum_{k\in K}\sum_{n\in N_i} c_{i,n,k} f_{i,n,k} \tag{5.26}$$

Subject to

$$\sum_{n\in N_i}\sum_{k\in K} f_{i,n,k} = 1 \qquad \forall i \in F \tag{5.27}$$

$$y_{k,o,t^-} + \sum_{i\in I(k,o,t)}\sum_{n\in N_i} f_{i,n,k} - y_{k,o,t^+} - \sum_{i\in O(k,o,t)}\sum_{n\in N_i} f_{i,n,k} = 0 \qquad \forall k \in K, o \in O, t \in T \tag{5.28}$$

$$\sum_{o\in O} y_{k,o,t_n} + \sum_{i\in CL(k)}\sum_{n\in N_i} f_{i,n,k} \le N_k \qquad \forall k \in K \tag{5.29}$$

$$f_{i,n,k} \in \{0,1\} \quad \forall i \in F, n \in N_i, k \in K \tag{5.30}$$

$$y_{k,o,t} \geq 0 \quad \forall k \in K, o \in O, t \in T. \tag{5.31}$$

The first set of constraints are cover constraints, ensuring that each flight leg is assigned to exactly one departure time(that is, flight leg copy) and one fleet type. The second set of constraints are aircraft flow balance constraints, ensuring aircraft balance at each location for each fleet type. The third set of constraints are the count constraints, which ensure that the assignment does not require more aircraft than that available.

Let Y be the set of feasible solutions that satisfy all the above constraints. Two integrated models for robust scheduling and fleet assignment are presented as follows.

$$\min \sum_{i\in F} \sum_{k\in K} \sum_{n\in N_i} c_{i,n,k} f_{i,n,k} \tag{5.32}$$

Subject to:

$$\sum_{i\in F^O} \sum_{n\in N_i} \sum_{j\in C^+(i)} \sum_{m\in N_j} E[dp_{i,n}^{j,m}] x_{i,n}^{j,m} \leq dp^U \tag{5.33}$$

$$\sum_{m\in N_j} x_{i,n}^{j,m} = \sum_{k\in K} f_{i,n,k} \quad \forall i \in F^O, n \in N_i, j \in C^+(i); \tag{5.34}$$

$$\sum_{n\in N_i} x_{i,n}^{j,m} = \sum_{k\in K} f_{j,m} \quad \forall j \in F^I, m \in N_j, i \in C^-(j); \tag{5.35}$$

$$0 \leq x_{i,n}^{j,m} \leq 1 \quad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j. \tag{5.36}$$

$$f \in Y \tag{5.37}$$

In this model, we minimize the fleet assignment cost but constrain the expected number of disrupted passengers to not exceed an upper bound dp^U. Constraints 5.34 and 5.35 reflect the relationship between x variables and f variables. The upper bound dp^U is defined as

$$dp^U = (1+\delta)Z_{dp}^{\min} \tag{5.38}$$

where $Z_{dp}^{\min}$ is the minimum expected number of disrupted passengers without considering the cost for fleet assignment, obtained by solving the model presented in Chapter 4. δ is a small positive parameter value.

Our next model minimizes the expected number of disrupted passengers and limit fleet assignment costs to an upper bound c^U.

$$\min \sum_{i\in F^O} \sum_{n\in N_i} \sum_{j\in C^+(i)} \sum_{m\in N_j} E[dp_{i,n}^{j,m}]x_{i,n}^{j,m} \tag{5.39}$$

Subject to:

$$\sum_{i\in F} \sum_{k\in K} \sum_{n\in N_i} c_{i,n,k} f_{i,n,k} \leq c^U \tag{5.40}$$

$$\sum_{m\in N_j} x_{i,n}^{j,m} = \sum_{k\in K} f_{i,n,k} \qquad \forall i \in F^O, n \in N_i, j \in C^+(i); \tag{5.41}$$

$$\sum_{n\in N_i} x_{i,n}^{j,m} = \sum_{k\in K} f_{j,m} \qquad \forall j \in F^I, m \in N_j, i \in C^-(j); \tag{5.42}$$

$$0 \leq x_{i,n}^{j,m} \leq 1 \quad \forall i \in F^O, n \in N_i, j \in C^+(i), m \in N_j. \tag{5.43}$$

$$f \in Z \tag{5.44}$$

The upper bound c^U is defined as

$$c^U = (1+\epsilon) Z_c^{\min} \tag{5.45}$$

where $Z_c^{\min}$ is the minimum fleet assignment cost obtained by solving the FAMTW model. ϵ is a small positive parameter value.

Some of the techniques proposed in Rexing et al. (2000) and Hane et al. (1995) to reduce problem size for fleet assignment might not be useful for the above models. For instance, we should be very careful when deleting flight copies even if they have the same fleet assignment costs because these copies might lead to different numbers of passenger disruptions. According to Rexing et al. (2000) [65], for a network with 2037 flights and 7 fleet types, if 20 minute time windows are allowed and copies are generated every 5 minutes, their problem had 105,524 rows, 155,219 columns, and 363,542 non-zero elements before applying any techniques to reduce problem size. Hence, problem size for FAMTW is large, but not prohibitive. The basis for this statement is our experience, reported in Table 4.12, in solving a flight schedule retiming problem with 32,514 rows, 507,253 columns and 1,040,236 non-zero elements.

Chapter 6

Summary and Future Research Directions

6.1 Summary

In airline operations, delays and cancellations of flights are common phenomena. Because passengers, crews and aircraft are interdependent, especially in hub-and-spoke networks, delays and cancellations of flights often cause delays and disruptions for crews and passengers. Crew and passenger delays and disruptions can then cause further flight delays and disruptions.

Fundamentally, there are two approaches to address this problem. One is to reschedule flights, crews and passengers after disruptions occur; and another is to build robust plans by considering possible delays and disruptions. A more robust plan can reduce delays and disruptions during operations.

In Chapter 2, we first show that building robustness into flight schedules is a difficult task and requires innovative ideas and methodologies. We review methodologies for robust planning in airline schedule planning and other application domains. We provide various definitions of robustness in the context of airline schedule planning, and present a framework for robust airline schedule planning. We categorize robust airline schedule planning methods proposed in the literature within this framework. Moreover, we identify some open research topics that can also fit into this framework.

One of the difficulties of building robust airline schedule planning models is that it is hard to balance the trade-off between schedule robustness and costs. In fact, it is difficult to determine the value of schedule robustness. In our research, we look for ways to develop robust airline schedule planns without significantly added costs. As a case in point, in Chapter 3 we present a robust aircraft maintenance routing model and its solution approach. Because flight delays propagate along an aircraft's route, our model minimizes delay propagation by intelligently routing aircraft. Using data from a major US airline, we show that this method can generate robust maintenance routes that can reduce propagated delays by more than 40%, improve 15-minute on-time rates by 1.6% and reduce the number of disrupted passengers by approximately 10%.

In Chapter 4, we present another approach, flight schedule retiming, to build robustness into airline schedules without significantly increasing costs. Our idea here is to allow flight departures to be rescheduled within small time windows in order to reduce the number of passengers who miss their connections because of flight delays. Our approach places slack in the schedule where it is more advantageous to passengers. We present various ways to model this problem and discuss their relative strength and weaknesses. Using data from a major US airline, we show that our method generates schedules that can reduce the number of disrupted passengers by as much as 40%.

In Chapter 5, we describe possible some extensions to the models we proposed in Chapter 3 and 4. The robust aircraft maintenance routing model can be extended to solve integrated fleet assignment and robust aircraft maintenance routing problems using either of the two models we present. We also present two models extending our flight schedule retiming models to incorporate fleet assignment decisions, in an attempt to further reduce passenger disruptions, without substantive increase in fleet assignment costs. These integrated models will likely enhance the performance of airline schedules and improve passenger travel times.

6.2 Directions for Future Research

Examples of future research directions are described below.

6.2.1 Adding Weights to Objective Function Coefficients

Our models presented earlier minimize the expected total number of propagated delays and passenger misconnections. We treat all propagated delays the same no matter how they are distributed and which flights they impact, and treat all passenger misconnections the same whether they are business travelers or not. In practice, airlines might want to protect certain flights and passengers that are important to them. In this case, various weights can be added to the objective function coefficients. For example, in RAMR model, we could add higher weights for longer delays, because longer delays might cause far more serious downstream delays and disruptions; we could add higher weights for delays propagated to hubs; we could also add higher weights for delays propagated to flights with many business passengers. In our flight schedule models, we could add higher weights for passenger misconnections involving business passengers or international flights.

6.2.2 Considering Variance in Objective Functions

Although our models minimize the expected values, it is equally important to consider variances. If variances are very high, then airlines might incur many disruptions and high costs to recover even if their plans minimize expected costs. One approach is to minimize the expectation plus a constant times the variance. This constant can be viewed as a parameter to balance the expectation and variance.

6.2.3 Crew Considerations

In our models, impacts on crews are not considered explicitly. Crew scheduling, however, is very important in airline schedule planning, because crew cost is the second largest cost among all operation costs for airlines and crew delays and disruptions have impact on flight and passenger delays and disruptions. One way to take crew into consideration is to use simulation to evaluate the impacts of our models on crews. Another approach is to consider robust aircraft maintenance routing and flight schedule retiming when solving crew scheduling problems. Cohn and Barnhart (2003) [31] solve an integrated model for aircraft maintenance routing and crew scheduling without considering the possible delays and disruptions in the operations. We could develop an integrated model for robust aircraft maintenance routing and crew scheduling.

6.2.4 Applications to Airline Operations

Although the models presented in this dissertation are developed for planning, they might also be used in airline operations with some changes. For example, the flight schedule retiming approach can be used to determine the exact flight departure times when delays and disruptions occur to minimize passenger misconnections. In this case, the objective function coefficients are exact numbers instead of expected value.

6.2.5 Using Simulation to Evaluate Schedules

Simulation is a good way to evaluate comprehensively the future performance of schedule plans in operations. We can use or develop a simulator to further test schedules generated by our models. In order to do it, the simulator must be accurate and consider all the components of an airline's operations such as recovery of aircraft, crews and passengers. In fact, one challenge is to foresee how recovery will take place, because different recovery policies will give different performance results even for the same schedule.

6.2.6 Fleet Assignment with Minimum Expected Cost

Conventional fleet assignment models do not consider extra costs incurred during irregular operations. In airline operations, when delays or disruptions occur, airlines often swap aircraft between flights to recover and reduce added costs cause by these delays or disruptions (Yan and Tu (1997) [82]). Costs associated with this swap include costs for switching gates and/or switching crew members, extra spill costs, etc.. To generate a fleeting solution that can facilitate recovery, one can build a model to minimize expected costs. An example is a two-stage stochastic programming model with recourse that minimizes fleet assignment costs plus the expected cost of future recovery actions caused by disruptions in the original schedule.

6.2.7 Fleet Assignment Under Demand Uncertainty

Demand is considered to be an constant input for each flight or itinerary in the conventional fleet assignment models. However, demand is a random variable and it is very difficult to forecast demand accurately. Without considering demand uncertainty, actual load factors can be lower than expected, leading to excess capacity and operating costs, or higher than expected, leading to overloading costs. Therefore, existing fleet assignment models can be improved by considering demand uncertainty.

Barnhart et al. (2001) [10] develop an itinerary-based fleet assignment model to improve FAM models (Hane et al. (1995) [42]) by more accurately determining spill costs (the lost revenue due to inability of airlines to accommodate total demand). A two-stage stochastic programming model could be built to consider demand uncertainty. The first stage would be to determine fleet assignment, based on expected spill costs. The second stage would be to determine passenger flows; however, solving such model for a large-scale real-life problem might be very challenging, because the itinerary-based fleet assignment model already has very large size. Another alternative is to model the demand uncertainty in the Subnetwork-based Fleet Assignment (Barnhart et al. (2001) [7], Lohatepanont (2001) [56]). This model can capture fleet assignment cost as accurately as the itinerary-based model without compromising tractability. Hence, it is suitable for further developing integrated models.

In addition, most of the general robust planning methodologies discussed in Section 2.2.1 have been used successfully to solve various problems under demand uncertainty. We believe these methodologies can also be applied to fleet assignment under demand uncertainty.

6.2.8 Aircraft Routes with Swap Opportunities

Ageeva (2000) [2] proposes to build aircraft routes that intersect more than once within a certain time window. This allows controllers more flexibility to swap planes during operations. She generates a set of optimal or near-optimal routing solutions, then selects the one with most swap opportunities. In fact, it is not necessary to require pairs of routes to intersect twice in order to provide a swap opportunity. If two routes intersect once but have the same number of days left before required maintenance, then there is a swap opportunity. To develop this further, one could develop a model to maximize the number of swap opportunities. One idea is to define a variable for each pair of strings, then the number of swap opportunities can be determined for this pair of strings. Such a model might be intractable with a prohibitively large number of variables. Thus, new modeling ideas and solution approaches that exploit special structure of the problem are needed.

6.2.9 Aircraft Routes with Short Cycles

Rosenberger et al. (2001) [67] present a string-based model to generate routes with short cycles. Such routes are less sensitive to flight cancellations than those without short cycles. They determine a lower bound for the number of short cycles using the hub connectivity of a fleet assignment and maintenance routing. They define hub connectivity as the number of legs in a route that begins at a hub, ends at a different hub, and only stops at spokes in between. Then they present a model to minimize hub connectivity.

Another idea is to develop a string-based model to maximize directly the number of short cycles, because once a string is built, the number of short cycles can be easily calculated. One problem with this idea, however, is that the pricing problem can not be cast as a shortest path problem.

Bibliography

[1] Abara, J. (1989). "Applying Integer Linear Programming to the Fleet Assignment Problem." *Interfaces* 19, July-August.

[2] Ageeva, Y. (2000). "Approaches to incorporating robustness into airline scheduling." Master's Thesis. Massachusetts Institute of Technology.

[3] Air Transport Association website (2003): http://www.airlines.org/.

[4] AMR Corporation website (2003): http://www.amrcorp.com/.

[5] Anbil, R., R. Tanga, and E. Johnson (1992). "A Global Approach to Crew-Pairing Optimization." *IBM Systems Journal* 31: 71 - 78.

[6] Barnhart, C., N. Boland, L. Clarke, E. Johnson, G. Nemhauser, R. Shenoi (1998). "Flight String Models for Aircraft Fleeting and Routing." *Transportation Science* 32: 208 - 220.

[7] Barnhart, C., A. Farahat, and M. Lohatepanont (2001). "Airline Fleet Assignment: An Alternative Model and Solution Approach Based on Composite Variables", Submitted to *Operations Research.*

[8] Barnhart, C., E. Johnson, R. Anbil, and L. Hatay (1994). "A Column Generation Technique, for the Long-Haul Crew Assignment Problem." *Optimization in Industry: Volume II*, T. Ciriano and R. Leachman (eds.). John Wiley and Son, England, 7 - 22.

[9] Barnhart, C., E. Johnson, G. Nemhauser, M. Savelsbergh, and P. Vance (1998). "Branch-and-Price: Column Generation for Solving Huge Integer Programs". *Operations Research* **46** (3): 316-329.

[10] Barnhart, C., T. Kniker, and M. Lohatepanont (2001). "Itinerary-Based Airline Fleet Assignment". *Transportation Science*, to appear.

[11] Barnhart C. and R. Shenoi (1998). "An Approximate Model and Solution Approach for the Long-Haul Crew Pairing Problem." *Transportation Science* 32: 221 - 231.

[12] Barnhart, C., K. Talluri (1997). Chapter 10: Airline Operations Research. *Design and Operation of Civil and Environmental Engineering Systems*: 435-469.

[13] Beasley J. and B. Cao (1996). "A Tree Search Algorithm for the Crew Scheduling Problem." *European Journal Of Operational Research* 94: 517 - 526.

[14] Bellman, R., and L.A. Zadeh (1970). "Decision-making in a Fuzzy Environment". *Management Science* 17, 141-164.

[15] Ben-Akiva, M. and S. Lerman (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press, Cambridge, Massachusetts.

[16] Berge, M. (1994). "Timetable Optimization: Formulation, Solution Approaches, and Computational Issues". AGIFORS proceeding, 341-357.

[17] Bertsimas, D., John N. Tsitsiklis (1997). *Introduction to Linear Optimization.* Athena Scientific, Inc.

[18] Birge, J. (1995). "Current Trends in Stochastic Programming Computation and Applications." Technical Report 95-15, Department of Industrial and Operations Engineering, University of Michigan.

[19] Birge, J., F. Louveaux (1997). *Introduction to Stochastic Programming.* Springer, Inc.

[20] Bradley, S. P., A. Hax, and T. L. Magnanti (1977). *Applied Mathematical Programming.* Addison-Wesley Publishing Company, Reading, Massachusetts.

[21] Bratu, S., C. Barnhart (2002). "A Study of Passenger Delay for a Major Hub-and-Spoke Airline." Working paper, Center for Transpotation and Logistics, Massachusetts Institute of Technology.

[22] Bureau of Transportation Statistics (2003). http://www.bts.gov/oai/.

[23] Cao, J. and A. Kanafani (1997). "Real-time Decision Support for Integration of Airline Flight Cancellations and Delays." *Transportation Planning and Technology* 20: 183-217.

[24] Chebalov, S., and Klabjan, D. (2002). "Robust Airline Crew Scheduling: Move-up Crews." Proceedings of the 2002 NSF Design, Service, and Manufacturing Grantees and Research Conference.

[25] Chu, H., E. Gelman, and E. Johnson (1997). "Solving Large Scale Crew Scheduling Problems." *European Journal of Operational Research* 97: 260 - 268.

[26] Clarke, L., C. Hane, E. Johnson, and G. Nemhauser (1996). "Maintenance and Crew Considerations in Fleet Assignment." *Transportation Science* 30: 249 - 260.

[27] Clarke, L., E. Johnson, G. Nemhauser, and Z. Zhu (1996). "The Aircraft Rotation Problem." *Annals of Operations Research* 69: 33 - 46.

[28] Clarke, M. and B. Smith (1999). "The Development of the Airline Operations Control Center." Research paper, Sabre Inc.

[29] Clarke M. and B. Smith (2000). "The Impact of Operations Research on the Evolution of the Airline Industry: A Review of the Airline Planning Process." Research paper, Sabre Inc.

[30] Cohn, A. (2002). "Composite-Variable Modeling for Large-Scale Problems in Transportation and Logistics." Doctoral Thesis. Massachusetts Institute of Technology.

[31] Cohn A. and C. Barnhart (2003). "Improving Crew Scheduling by Incorporating Key Maintenance Routing Decisions." *Operations Research* (To appear).

[32] Cordeau, J., G. Stojkovic, F. Soumis, and J. Desrosiers (2000). "Benders Decomposition for Simultaneous Aircraft Routing and Crew Scheduling." Technical report G-2000-37, GERAD, Ecole Polytechnique de Montreal.

[33] Daskin, M. S., and N. D. Panayotopoulos (1989). "A Lagrangian Relaxation Approach to Assigning Aircraft to Routes in Hub and Spoke Networks." *Transportation Science* 23: 91 - 99.

[34] Desaulniers, G., J. Desrosiers, Y. Dumas, S. Marc, B. Rioux, M. Solomon, and F. Soumis (1997). "Crew Pairing at Air France." *European Journal of Operational Research* 97: 245 - 259.

[35] Desaulniers, G., J. Desrosiers, M. M. Solomon, and F. Soumis (1997). "Daily Aircraft Routing and Scheduling." *Management Science 43*, 841-854.

[36] NIST/SEMATECH e-Handbook of Statistical Methods (2002). http://www.itl.nist.gov/div898/handbook/.

[37] Etschmaier, M. and D. Mathaisel (1985). "Airline Scheduling: An Overview." *Transportation Science* 19: 127 - 138.

[38] Feo, T. A., and J. F. Bard (1989). "Flight Scheduling and Maintetance Base Planning." *Management Science* 35, 1415-1432.

[39] Flint, P. (2000). "Back on Schedule." *Air Transport World*, p47-51.

[40] Gopalan R. and K. Talluri (1998). "The Aircraft Maintenance Routing Problem." *Operations Research* 46: 260 - 271.

[41] Gu, Z., E. L. Johnson, G. L. Nemhauser, and Y. Wang (1994). "Some Properties of the Fleet Assignment Problem." *Operations Research Letters* 15, 59-71.

[42] Hane, C., C. Barnhart, E. Johnson, R. Marsten, G. Nemhauser, and G. Sigismondi (1995). "The Fleet Assignment Problem: Solving a Large-Scale Integer Program." *Mathematical Programming* 70: 211 - 232.

[43] Hoffman, K. and M. Padberg (1993). "Solving Airline Crew Scheduling Problems by Branch-and-Cut." *Management Science* 39: 657 - 682.

[44] Jarrah, A. I.Z., G. Yu, N. Krishnamurthy, and A. Rakshit (1993). "A Decision Support Framework for Airline Flight Cancellations and Delays." *Transportation Science* 27, 266-280.

[45] Jin, H. (1998). "Designing Robust Railroad Blocking Plans." Doctoral Thesis. Massachusetts Institute of Technology.

[46] Kabbani, N. M., and B. W. Patty (1992). "Aircraft Routing at American Airlines." Proceedings of the Thirty-Second Annual Symposium of AGIFORS, Budapest, Hungary, 1992.

[47] Kang, L. S., and Clarke, J. P. (2002). "Degradable Airline Scheduling." working paper, Operations Research Center, Massachusetts Institute of Technology.

[48] Klabjan, D., E. Johnson, G. Nemhauser, E. Gelman, and S. Ramaswamy (1999). "Airline Crew Scheduling with Time Windows and Plane Count Constraints." Technocal Report TLI/LEC-99-12, Georgia Institute of Technology.

[49] Klabjan, D., E. Johnson, and G. Nemhauser (1999). "Solving Large Airline Crew Scheduling Problems: Random Pairing Generation and Strong Branching." Technical Report TLI/LEC-99-11, Georgia Institute of Technology.

[50] Klabjan, D., E. Johnson, G. Nemhauser, E. Gelman, and S. Ramaswamy (1999). "Airline Crew Scheduling with Time Windows and Plane Count Constraints." Technical Report TLI/LEC-99-12, Georgia Institute of Technology.

[51] Klabjan, D. and K. Schwan (1999). "Airline Crew Pairing Generation in Parallel." Technical Report TLI/LEC-99, Georgia Institute of Technology.

[52] Kleywegt, A. J., A. Shapiro, and T. Homem-de-mello (2001). "The Sample Average Approximation Method for Stochastic Discrete Optimization." To appear in *SIAM Journal on Optimization.*

[53] Kouvelis, P., and G. Yu (1997). *Robust Discrete Optimization and Its Applications.* Kluwer Academic Publishers.

[54] Lettovsky, L. (1997). "Airline Operations Recovery: An Opmitization Approach." Ph.D. Dissertation, Georgia Institute of Technology.

[55] Levin, A. (1971). "Scheduling and Fleet Routing Models for Transportation Systems." *Transportation Science* 5, 232-255.

[56] Lohatepanont, M. (2001). "Fleet Assignment and Schedule Design: Integrated Models and Algorithms". Doctoral Thesis. Massachusetts Institute of Technology.

[57] Lohatepanont, M. and C. Barnhart (2001). "Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment." Accepted by *Transportation Science.*

[58] Luo, S. and G. Yu (1997). "On the Airline Schedule Perturbation Problem Caused by the Ground Delay Program." *Transportation Science* 31, 298-311.

[59] Martin R. C. (1999). *Large Scale Linear and Integer Optimization: A Unified Approach.* Kluwer Academic.

[60] Mathaisel, D. F. X (1996). "Decision Support Airline System Operations Control and Irregular Operations." *Computers and Operations Research* 23, 1083-1098.

[61] MIT Global Airline Industry Program Website: http://web.mit.edu/airlines/.

[62] Mulvey, J.M., R.J.Vanderbei, and S.A.Zenios (1995). "Robust Optimization of Large-Scale Systems." *Operations Research* 43, 264-281.

[63] Pindyck, R.S. and D. L. Rubinfeld (1998). *Econometric Models and Economic Forecasts.* Irwin McGraw-Hill.

[64] Rakshit, A., N. Krishnamurthy, and G.Yu (1996). "System Operations Advisor: A Real-time Decision Support System for Managing Airline Operations at United Airlines." *Interfaces* 26, 50-58.

[65] Rexing, B., C. Barnhart, T. Kniker, A. Jarrah, and N. Krishnamurthy (2000). "Airline Fleet Assignment with Time Windows." *Transportation Science* 34: 1 - 20.

[66] Rosenberger, J. M., E. L. Johnson, and G. L. Nemhauser (2001). "Rerouting Aircraft for Airline Recovery." white papers and reports, the Logistics Institute, Georgia Institute of Technology. http://www.tli.gatech.edu/research/papers/files/tli0104.pdf.

[67] Rosenberger, J. M., E. L. Johnson, and G. L. Nemhauser (2001). "A Robust Assignment Model with Hub Isolation and Short Cycles." white papers and reports, the Logistics Institute, Georgia Institute of Technology. http://www.tli.gatech.edu/research/papers/files/tli0112.pdf.

[68] Ryan, D.M. and B.A.Foster (1981). "An integer Programming Approach to Scheduling." *Computer Scheduling of Public Transport: Urban Passenger Vehicle and Crew Scheduling*, A. Wren (ed.), North Holland, Amsterdam, 269-280.

[69] Sakawa, M. (1993). *Fuzzy Sets and Interactive Multiobjective Optimization.* Plenum Publishers.

[70] Schaefer, A. J., E. L. Johnson, A. J. Kleywegt, and G. L. Nemhauser (2001). "Airline Crew Scheduling Under Uncertainty." white papers and reports, the Logistics Institute, Georgia Institute of Technology. http://www.tli.gatech.edu/research/papers/files/tli0101.pdf.

[71] Shenoi (1996). "Integrated Airline Schedule Optimization: Models and Solution Methods." Doctoral Thesis. Massachusetts Institute of Technology.

[72] Stochastic Programming Community Website (2003). http://stoprog.org/.

[73] Subramanian, R., R. P. Scheff, J. D. Quillinan, D.S. Wiper and R.E. Marsen (1994). "ColdStart - Fleet Assignment at Delta Airlines." *Interfaces* 24, 104-120, Jan-Feb.

[74] Talluri, K. (1998). "The Four-Day Aircraft Maintenance Routing Problem." *Transportation Science* 32: 43 - 53.

[75] Teodorovic, D. and S. Guberinic (1990). "Optimal DispatchingStrategy on an Airline Network after a Schedule Perturbation." *European Journal of Operational Research* 15: 178-182.

[76] Teodorovic, D. and E. Krcmar-Nozic (1989). "Multicariteria Model to Determine Flight Frequencies on an Airline Network under Competitive Conditions." *Transportation Science* 23: 14-25.

[77] Teodorovic, D. and G.Stojkovic (1995). "Model to Reduce Airline Schedule Disturbances." *Journal of Transportation Engineering* 121: 324-331.

[78] Thengvall, B. G., J. F. Bard, and G. Yu (2000). "Balancing User Preferences for Aircraft Schedule Recovery." *IIE Transactions* 32: 181-193.

[79] Vance, P., C. Barnhart, E. Johnson, and G. Nemhauser (1997). "Airline Crew Scheduling: A New Formulation and Decomposition Algorithm." *Operations Research* 45: 188 - 200.

[80] Watanable, T., and H. Ellis (1993). "Robustness in Stochastic Programming Models." *Applid Mathematical Modelling* 17, 545-554.

[81] Yan, H., and P.B. Luh (1997). "A Fuzzy Optimization-Based Method for Integrated Power System Scheduling and Inter-Utility Power Transaction with Uncertainties." *IEEE Transactions on Power Systems* 12, 756-763.

[82] Yan, S., and Y. Tu (1997). "Multifleet Routing and Multistop Flight Scheduling for Schedule Perturbation." *European Journal of Operational Research* 103: 155-169.

[83] Yan, S. and D.Yang (1996)."A Decision Support Framework for Handling Schedule Perturbation." *Transportation Research*, Part B, 30: 405-419.

[84] Yen, J. W. and J. R. Birge (2001). "A Stochastic Programming Approach to the Airline Crew Scheduling Problem." submitted to *Operations Research.*

[85] Yu, G., M. Arguello, G. Song, S. M. McCowan, and A. White (2003). "A New Era for Crew Recovery at Continental Airlines." to appear in *Interfaces.*

[86] Zimmermann, H. J. (1991). *Fuzzy Set Theory and Its Applications.* Kluwer Academic Publishers, Boston.