Routing and Scheduling Models for Robust Allocation of Slack

by

Virot Chiraphadhanakul

B.Eng. (Computer Engineering)
Chulalongkorn University (2007)

Submitted to the Department of Civil and Environmental Engineering
and the Sloan School of Management
in partial fulfillment of the requirements for the degrees of

Master of Science in Transportation

and

Master of Science in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

Author . . . .
Department of Civil and Environmental Engineering
and the Sloan School of Management
May 18, 2010

Certified by . . . .
Cynthia Barnhart
Professor of Civil and Environmental Engineering
Associate Dean for Academic Affairs, School of Engineering
Thesis Supervisor

Accepted by . . . .
Dimitris J. Bertsimas
Boeing Professor of Operations Research
Co-Director, Operations Research Center

Accepted by . . . .
Daniele Veneziano
Chairman, Departmental Committee for Graduate Students

Routing and Scheduling Models for Robust Allocation of Slack

by

Virot Chiraphadhanakul

Submitted to the Department of Civil and Environmental Engineering
and the Sloan School of Management
on May 18, 2010, in partial fulfillment of the
requirements for the degrees of
Master of Science in Transportation
and
Master of Science in Operations Research

Abstract

A myriad of uncontrollable factors in airline operations make delays and disruptions unavoidable. Most conventional scheduling models, however, ignore the presence of uncertainties in actual operations in order to limit the complexity of the problem. This leads to schedules that are prone to delays and disruptions. As a result, there has been wide interest recently in building robustness into airline schedules. In this work, we investigate slack allocation approaches for robust airline schedule planning. In particular, we propose three models: aircraft re-routing model, flight schedule re-timing model, and block time adjustment model, together with their variants. Using data from an international carrier, we evaluate the impacts of the resulting schedules on various performance metrics, including passenger delays. The results show that minor modifications to an original schedule can significantly improve the overall performance of the schedule. Through empirical results, we provide a comprehensive discussion of model behaviors and how an airline's characteristics can affect the strategy for robust scheduling.

Thesis Supervisor: Cynthia Barnhart
Title: Professor of Civil and Environmental Engineering
Associate Dean for Academic Affairs, School of Engineering

Acknowledgments

It has been my honor to work with Professor Cynthia Barnhart. I would like to express my deepest gratitude to her for her constant understanding and support. This thesis could not be completed without her invaluable advice. Thank you, Cindy, for being my great mentor and my trusted friend.

I am grateful to Jeppesen who made this project possible. Thanks are due to Sergey Tiourine, Bo Vaaben, and Erik Andersson for their assistance and insightful advice.

I would like to thank my colleagues– Vikrant, Lavanya, and Niklaus Eggenberg for their suggestions and comments that have contributed greatly to my thesis.

Many thanks go to all my MST and ORC fellows, notably, Clara, Cristian, Joline, Matthieu, Sarvee, and Ying, whose friendship made my stay at MIT much more enjoyable and memorable. Also, I would like to thank all Thai students in Boston and Cambridge areas for building a home-like atmosphere for me.

I am deeply indebted to all of my teachers. Without any of them, I would not be where I am today. My special appreciation goes to Dr. Manoj Lohatepanont, Dr. Proadpran Punyabukkana, and Dr. Atiwong Suchato who had endorsed me for admission to MIT.

More importantly, I owe my deepest gratitude to my family for their love and support. They are my everything, and I love them all. Finally, I would like to thank Noey for her love and her patience with our long-distance relationship.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Airline Schedule Planning

The airline schedule planning process involves considerable complexities in airline operations. A large number of interconnected elements in an airline schedule, such as origin and destination airports, gates, airport slots, aircraft types, crew restrictions, aircraft maintenance requirements and passenger demands, need to be taken into account. As a result, the problem size becomes too large to be solved in a single optimization model. Conventionally, the schedule planning process is decomposed into four subproblems: (1) schedule design, (2) fleet assignment, (3) aircraft maintenance routing, and (4) crew scheduling. We provide here a brief description of each subproblem. For interested readers, comprehensive overviews of the airline schedule planning process and detailed literature reviews are presented in Barnhart et al. [4] and Belobaba et al. [6].

1. Schedule Design The goal of this subproblem is to select and schedule a set of flight legs to be operated by an airline, given the desired markets to serve. Each flight leg in an airline flight schedule is specified by an origin airport, a destination airport, a scheduled departure time, and a scheduled arrival time. Because an airline's flight schedule is very critical to its competitive position and profitability, a schedule development process requires collaboration between many business units to resolve all tactical and operational issues that may arise from the resulting schedule, and therefore the use of optimization models to generate "optimized" flight schedules is still limited.

2. Fleet Assignment Given a flight schedule from the previous subproblem, the fleet assignment problem is to assign a specific aircraft type to each flight leg such that the cost of assignment is minimized. This cost includes the cost of operating a flight leg with a specific aircraft type and the *spill* cost– the opportunity cost of having insufficient seating capacity to satisfy passenger demands. Additionally, an assignment is feasible only if it requires no more aircraft of each type than is available, and only if the flow of each aircraft type is balanced across the airline network, and thus allowing the schedule to be repeated periodically.

3. Aircraft Maintenance Routing Given a flight schedule and a fleet assignment obtained in the previous steps, the aircraft maintenance routing problem is to assign specific aircraft (i.e., tail number) to each flight leg such that each aircraft's routing–the sequence of flight legs it operates–allows periodic maintenance checks after a certain number of flying hours. In general, the main purpose of this problem is to obtain a feasible solution, rather than an optimal solution with respect to some objective function.

4. Crew Scheduling Given the solutions to the three preceding steps, the crew scheduling problem is to assign cockpit and cabin crews to each flight leg such that the crew cost is minimized. A crew schedule must satisfy numerous work-rule restrictions as well as collective bargaining agreements between the airline and its employees. Because of the complexity of this problem, it is typically divided further into two sequential subproblems, namely, the *crew paring problem* and the *crew assignment problem*. In the crew paring problem, minimum cost multi-day sequences of flight legs, called *pairings*, are created. These pairings must satisfy all the work-rule restrictions. The crew assignment problem then combines the pairings into month-long crew schedules, called *bidlines* or *rosters* and assigns them to each crew according to their preferences and priorities.

Because of the dependency among the subproblems, sequentially solving these subproblems typically yields a suboptimal solution. In the past few decades, numerous efforts have been made to integrate some of these subproblems in order to improve the solution quality. Nonetheless, for large airlines, no single optimization model capturing every element in the airline scheduling process has been deployed successfully.

Additionally, these subproblems are in fact still so large and complex that optimization models are typically solved deterministically, i.e., assuming that every flight will be operated as planned. Ignoring the presence of uncertainties in actual operations results in schedules that are vulnerable to delays and disruptions, thereby incurring higher operational costs.

1.2 Delays and Disruptions in Airline Operations

Delays and Disruptions are inevitable in airline operations due to many unforeseeable factors such as congested airports, adverse weather conditions, crew sickness, and aircraft mechanical problems. The impact of delays is exacerbated when they propagate to subsequent flights through an airline's interconnected network. Large delays can also lead to flight cancellations and passenger misconnections, causing passengers to wait for several hours for the next available flight. Fearing, Vaze, and Barnhart [17] estimate passenger delays for U.S. domestic flights in 2007. The results show that just over 50% of the passenger delays are caused by flight delays; a third of the passenger delays are caused by flight cancellations; and nearly a sixth of the passenger delays are caused by missed connections.

To emphasize the impact of delays in the U.S., we present here some statistics from the report by the Joint Economic Committee (JEC) [32]. In 2007, 4.3 million hours of flight delays cost as much as $41 billion to the U.S. economy. These delays lead to $19 billion of direct operating costs to airlines. About 740 million extra gallons of jet fuel were consumed by the delayed flights, resulting in an additional 7.1 million metric tons of carbon dioxide emissions into the atmosphere. Assuming a value of $37.60 per delay hour per passenger, the delay cost to passengers was valued at up to $12 billion. Note that this calculation does not include the passenger delays due to misconnections and flight cancellations.

1.3 Robust Airline Schedule Planning

The significant impacts of delays and disruptions have motivated wide interests in building *robustness* into airline schedules, i.e., proactively making them more resilient to delays and disruptions, as opposed to the conventional deterministic scheduling models which ignore the presence of uncertainties in actual operations.

The key challenge of robust airline schedule planning is to define *robustness* of a schedule such that it well reflects desired characteristics and can be captured in a tractable mathematical model. We provide a detailed review of robust airline scheduling literature in Section 2.1.

In this thesis, we investigate slack allocation approaches for robust airline scheduling. Slack is defined as additional time allocated beyond the minimum time required for each aircraft connection, passenger connection, or expected flight duration. To minimize operating costs, airlines have made numerous efforts to increase the utilization of all resources in their operations, often resulting in the minimization of schedule slack. Slack, however, is desirable in robust schedules as it can potentially absorb delays in an airline network, reduce the likelihood of operational disruptions, and provide flexibility to recover once the operation is disrupted. Therefore, we seek to re-allocate, rather than simply increase, existing slack in schedules in order that the resulting distribution of slack is more effective in absorbing delays and minimizing disruptions.

For a more general discussion of irregular operations and control in air transportation, readers are referred to Ball et al. [3] and Kohl et al. [22].

1.4 Contributions

Our contributions in this thesis can be summarized as follows:

1) We propose a new visualization method that facilitates understanding of robust

schedule behaviors.

2) We provide a modeling framework for robust slack allocation in airline schedule planning. In particular, we propose three robust slack re-allocation models: the robust aircraft re-routing model, the robust flight schedule re-timing model, and the robust block time adjustment model, together with their variants. These models have different sets of underlying assumptions and affect the distribution of slack in the system in different manners. Additionally, we introduce a notion of *effective slack*, which is proved to serve as a good robustness proxy in many cases.

3) We present the proof-of-concept results obtained using data from an international carrier. The results show that minor modifications to an original schedule can significantly improve the overall performance of the schedule. Unlike many works in the literature that focus mainly on improving robustness of a schedule with respect to a particular objective, we evaluate the impacts of the resulting schedules on different performance metrics, including passenger delays and delay propagation. The results exhibit the trade-offs among different performance evaluation metrics. Along with the proof-of-concept results, we thoroughly examine the behaviors of the robust routing and scheduling models presented in this work; and we provide a comprehensive discussion of how different characteristics of an airline can affect the strategy for robust schedule planning.

1.5 Thesis Outline

The structure of the rest of the thesis is as follows. In Chapter 2, we provide background information on robust airline schedule planning, including a survey of related literature and some performance evaluation metrics that we will use extensively in this work. To facilitate understanding of a schedule's performance, we introduce the visualization that we develop and use in addition to data tables and aggregate statistics. At the end of the chapter, we define different types of slack and describe how an airline can strategically re-allocate slack in its schedules to minimize delays and disruptions.

In Chapter 3, we present the optimization models for different slack re-allocation schemes, namely, the robust aircraft re-routing model, the robust flight schedule re-timing model, and the robust block time adjustment model. For each slack re-allocation model, we also provide alternative objective functions that can potentially result in more robust solutions or different solutions that are robust with respect to different performance evaluation metrics.

In Chapter 4, the proof-of-concept results, obtained using data from an international carrier, are presented. We analyze the performances of the solutions to the robust slack re-allocation models proposed in this work. In parallel with the performance analysis, we provide a discussion of model behaviors and how the strategy for robust schedule planning depends on the characteristics of an airline's flight network.

In Chapter 5, we discuss possible future research topics that build upon the research and results of this thesis.

Chapter 2

Background

2.1 Robustness in Airline Scheduling: Literature Review

Ehrgott and Ryan (2000) [16] develop a bicriteria optimization framework for solving a crew scheduling problem. Two objective functions are cost and robustness. Robustness of a schedule is obtained by penalizing a pairing that has a connecting time less than the expected delay and requires crews to change aircraft. A violation is measured by the amount by which the expected delay of an incoming flight exceeds the scheduled connecting time. The detail of an expected delay calculation, however, is not provided. They use the ϵ-constraint methods to solve the bicriteria optimization problem. In particular, one objective is incorporated as a constraint with an upper bound of ϵ on its value in the optimization problem of the other objective. Although only preliminary results are presented, they sufficiently demonstrate the trade-off between cost and robustness in airline crew scheduling.

Ageeva (2000) [1] constructs robust schedules by maximizing the number of aircraft swap opportunities. Two aircraft "meet" if their routes contain a common station within a specific period of time. If two aircraft meet twice along their routes, then they can be swapped at the first meeting point and swapped back to their original routes later. These overlapping routes provide flexibility to recover when the planned schedule is disrupted. The traditional string-based aircraft routing model is used to obtain the optimal cost of the problem. Then, among the solutions with the optimal cost, the most robust one is selected. The results show that the robustness of the final solution can be improved as much as 35% compared to the original solution. Nonetheless, the operational performance of the approach is not provided.

Rosenberger, Johnson, and Nemhauser (2004) [26] develop a robust fleet assignment model that includes many short cancellation cycles and reduces hub connectivity. A cancellation cycle is a sequence of flight legs in an aircraft rotation that begins and ends at the same airport. A rotation with short cancellation cycles can decrease the number of flight legs that need to be cancelled when a rotation is disrupted. Hub connectivity is defined as the number of flight legs in an aircraft rotation that begins at one hub, ends at a different hub, and only stops at spokes. Having limited hub connectivity in a rotation can mitigate the impact of propagated disruptions from one hub to others. SimAir, a simulation of airline operations, together with a recovery module, is used to evaluate the resulting assignments. Crews, however, are excluded from the simulation. The results show that the performance of this approach is more robust than that of the traditional fleet assignment model.

Kang (2004) [21] presents a new approach to obtain robust schedules by incorporating degradability into the existing schedule. She divides the existing schedule into independent sub-schedules or layers, so that disruptions in one layer will not propagate to other layers. Moreover, each layer is prioritized based on revenue. When disruptions occur, airlines can take appropriate actions based on the priority of each layer. Degradability can be incorporated into the existing schedule in three stages: schedule design, fleet assignment, and maintenance routing. It is shown that considering degradable scheduling earlier in a schedule planning process yields a better objective function value. MEANS, the MIT Extensible Air Network Simulation, is used to simulate the airline operations. Even though the average flight delay and passenger delay of the degradable schedules are about the same as those of the traditional schedule, the distributions of delays among different layers are as expected, i.e., a higher priority layer suffers much less delay and fewer cancellations. The difference between each layer is even more significant under bad weather conditions. In addition, the degradable airline schedules also lead to lower recovery cost in most cases.

Schaefer et al. (2005) [28] solve a crew scheduling problem using a new approach that takes uncertainty into account. They modify the cost coefficients in a traditional crew scheduling model, which is modeled as a set partitioning problem, to reflect the actual costs in operations with disruptions. Two methods are proposed: 1) using the expected cost of a pairing, and 2) using a penalty parameter. In the former method, they use SimAir, a Monte Carlo simulation of airline operations, to find the expected cost of a pairing. The penalty method penalizes a pairing with particular properties that may bring about poor performance when disruptions exist. Note that this penalty is not limited to an additive penalty, as implemented in Ehrgott and Ryan (2000). The results from simulations show that the performance of crew schedules from both proposed methods are better than a traditional approach, which does not take uncertainty into account.

Lan, Clarke, and Barnhart (2006) [24] solve a robust aircraft maintenance routing problem (RAMR) by formulating a traditional string-based model with an objective of minimizing total expected propagated delay. The key idea is that the impact of propagated delay can be mitigated by cleverly routing aircraft and allocating slack to absorb delay propagation. They solve a deterministic mixed-integer linear program using a combination of column generation and branch-and-price technique. The results show that this approach can improve on-time performance, and decrease total propagated delay as well as the number of disrupted passengers– ones that miss connections. In addition to the robust aircraft maintenance routing model, they propose a robust flight schedule re-timing model (RFSR) that minimizes the expected number of disrupted passengers. In this model, the departure time of each flight is allowed to move within a small time window. Using the basic flight network, the solution of this model contains exactly one of the flight arc copies placed within the flight's time window. They again solve the problem using a branch-and-price technique. The results show that the optimal solution decreases the number of disrupted passengers by about 40% for a 30-minute time window.

Yen and Birge (2006) [33] model a crew scheduling problem as a two-stage stochastic integer program minimizing the expected cost. Their work can be considered as an extension of Schaefer et al. (2005). The major difference is that this model also captures disruption interactions among crew pairings. In particular, the first stage of the model solves the crew scheduling problem, minimizing expected costs, and the recourse problem minimizes the expected costs owing to crews switching planes. In order to solve this stochastic integer programming problem, they propose the flight-pair branching algorithm, which produces smaller branching trees than the traditional crew-pairing branching. The results exhibit the trade-off between planned crew costs and recourse delay costs. Moreover, the algorithm leads to pairings with smaller numbers of crew plane changes and larger connection times. These results are consistent with the work by Ehrgott and Ryan (2000).

Shebalov and Klabjan (2006) [29] present a new approach to constructing robust crew schedules that increases flexibility for recovery. They introduce the idea of move-up crews– ones that can potentially be swapped in operations. The model maximizes the number of move-up crews with a restriction on an increase in crew costs. They solve this problem by using a combination of column generation and Lagrangian relaxation. The results suggest that increasing the number of move-up crews brings about a decrease in operational crew costs because robust solutions require fewer deadheads, need fewer stand-by crews, and yield fewer uncovered legs. Nevertheless, a saving in crew costs only materializes when irregular operations occur. Otherwise, robust solutions result in higher crew costs.

Smith and Johnson (2006) [31] introduce a notion of *station purity* in the fleet assignment problem. By imposing station purity, the number of fleet types serving each station (or airport) in the schedule is limited to the station's *purity level.* Station purity is beneficial to airlines because it increases swap opportunities for both aircraft and crews to recover from delays and disruptions. Moreover, it can potentially reduce an airline's maintenance costs because each station does not have to hold fleet-specific equipment, stock additional spare parts, and maintain qualified maintenance crews for many fleet types. They develop a *station decomposition* approach and a *fix-and-price* heuristic in order to solve this problem efficiently. The results show that station purity can significantly reduce maintenance costs and crew costs, resulting in an estimated increased profit of $129 million per year. This figure does not include potential savings in actual operations.

The idea of station purity is extended to the crew planning in Gao, Johnson, and Smith (2009) [18], where the number of *crew bases* serving each airport is limited.

AhmadBeygi, Cohn, and Lapp (2008) [2] propose a flight schedule re-timing model that minimizes the propagation of *root flight delays*. In order to approximate propagated delays in the objective function, they simulate propagation of delays from multiple resources (aircraft, crews, passengers, etc.) using *propagation trees*. Because of the interconnectivity among different resources, any given flight might experience multiple flight delays from different sources simultaneously. In this case, their algorithm does not accurately estimate propagated delays. In contrast to the robust flight schedule re-timing model (RFSR) proposed by Lan, Clarke, and Barnhart (2006) [24], their formulation does not make use of flight copies. Instead, the changes in flight departure times are allowed to take any integral value within the specified time windows. Additionally, they show that the formulation is integral, given integral input data, and thus can be solved efficiently by relaxing the integrality constraints. In the computational experiments, they consider only delay propagation due to aircraft and cockpit crews, and the solutions are evaluated based on the objective function. The propagated delay reductions achieved by their solutions range from approximately 5% to 50%, depending on the time window sizes.

Eggenberg (2009) [15] proposes a general framework for solving an optimization problem subjected to uncertainty, called the *Uncertainty Feature Optimization* (UFO) framework. The underlying idea of this framework is to model the desired characteristics of the solution as *Uncertainty Features* (UF) for which we want to maximize. He shows that, with appropriate uncertainty features, the UFO framework is equivalent to a general stochastic programming approach and the robust optimization formulation proposed by Bertsimas and Sim (2004) [8]. To apply the UFO framework to the aircraft re-routing model and the flight re-timing model, the uncertainty feature is defined as the total aircraft connection slack and the sum of minimum aircraft connection slack in each aircraft route. He compares the performances of his models to the early versions of our robust aircraft re-routing model and robust flight schedule re-timing model. The computational results show that without an explicit use of historical delay data, his models tend to generate solutions that perform adequately well with respect to every performance metric; while our models that make extensive use of historical delay data can significantly improve the performance metrics that are positively correlated with the objective function but might perform worse with respect to the other metrics.

Burke et al. (2009) [11] propose a multi-objective approach for robust airline schedule planning. In particular, they consider two robust objective functions– *schedule reliability* and *schedule flexibility*. The value of schedule reliability is defined by the non-linear function of probabilities of on-time departures, and the value of schedule flexibility is defined by the number of aircraft swap opportunities. Moreover, their model allows aircraft re-routing and flight re-timing simultaneously. They approximate the Pareto optimal front using a *multi-meme memetic algorithm*– a hybrid of genetic algorithms with multiple local search operators. The computational results, obtained using KLM's simulation model [20], demonstrate the trade-off between reliability and flexibility in the schedules, and the increased reliability and flexibility result in the better on-time performance.

Marla and Barnhart (2010) [25] apply general robust optimization approaches, namely, the extreme-valued based approach of Bertsimas and Sim [7, 8] and the chance-constrained programming approach of Charnes and Cooper [12, 13], to the aircraft routing problem and compare to the problem-specific approach– the robust aircraft maintenance routing model proposed by Lan, Clarke, and Barnhart (2006) [24]. Through the empirical results, they investigate the behaviors of each model and provide a comprehensive discussion of the trade-offs among the different models. In essence, an appropriate choice of model depends on underlying delay distributions and performance metrics of interest. They also present the extended formulations that are applicable to the general network-based resource allocation problem.

2.1.1 Robust Schedules by Objective Function

Objectives of robust schedules proposed in the literature can be classified into three main categories:

1) **Preventing delays and disruptions**

   This type of robust schedule is aimed at minimizing an occurrence of delays and disruptions. To construct a robust schedule of this type, the optimization models require historical data such as

   - average arrival delays of each flight;
   - expected propagated delays at each connection;
   - distributions of delays at each airport;
   - probabilities of passenger misconnection at each connection.

   The objective of the model is to minimize expected delays and likelihood of disruptions by means of proxies derived from the given historical data. An alternative objective is to maximize slack on a day of operation by generating plans with:

   - aircraft routes with long aircraft connection times;
   - crew pairings with long rest times between duties and long sit times between plane changes; or
   - flight schedules with long scheduled flying times.

2) **Minimizing the impact of delays and disruptions, once a schedule gets disrupted**

   This type of robust schedule is constructed such that, once a schedule gets disrupted, the impact of delays and disruptions is minimal. There are two broad ways to achieve this goal.

2.1) **Maximizing recovery flexibility**

Recovery flexibility embedded in this type of robust schedule provides airline operations controllers many recovery options to alleviate delays and disruptions. Hence, an airline is likely to obtain a recovery solution that requires only a modest change and is more economical. For example, aircraft swap opportunities and move-up crews can be used to prevent further delay propagation, rather than delaying other flights awaiting the originally assigned crews or aircraft.

2.2) **Isolating delays and disruptions**

A schedule of this kind partitions an airline network into isolated subnetworks such that the delays and disruptions arising in one subnetwork are contained within that subnetwork. Thus, the impact of delays and disruptions is limited, and the other subnetworks can still be operated as planned.

3) **Minimizing the expected costs of a schedule**

As opposed to the traditional schedule planning approach, which assumes the schedule will be operated as planned, this type of schedule takes into account uncertainty that may arise in a day of operation and potentially increase the total operating cost. Note that in order to minimize the expected costs, the optimization model may involve minimizing delays and disruptions or their impact, so this objective is generally used in conjunction with the other objectives discussed earlier. The key advantage of this objective is that it can capture the trade-off between costs of robustness and savings in recovery costs.

Table 2.1 summarizes the preceding robust scheduling approaches in the literature by objective functions.

2.1.2 Robust Schedules by Schedule Planning Phase

Robust airline scheduling approaches proposed in the literature also apply to different phases of the traditional airline schedule planning process (Section 1.1) and are

	Min Expected Delay/Disruption	Max Recovery Flexibility	Delay/Disruption Isolation	Min Cost
Ehrgott and Ryan (2000)	•			•
Ageeva (2000)		•		•
Rosenberger, Johnson, and Nemhauser (2004)		• (Short Cycles)	• (Hub Isolation)	
Kang (2004)			•	
Schaefer et al. (2005)	• (Penalty Method)			• (Min Expected Cost)
Lan, Clarke, and Barnhart (2006)	•			
Yen and Birge (2006)	•			•
Shebalov and Klabjan (2006)		•		
Smith and Johnson (2006)		•		
AhmadBeygi, Cohn, and Lapp (2008)	•			
Gao, Johnson, and Smith (2009)		•		
Eggenberg (2009)	•			
Burke et al. (2009)	• (Reliability)	• (Flexibility)		
Marla and Barnhart (2010)	•			

Table 2.1: Robust schedules by objective function

	Schedule Design	Fleet Assignment	Aircraft Routing	Crew Scheduling
Ehrgott and Ryan (2000)				•
Ageeva (2000)			•	
Rosenberger, Johnson, and Nemhauser (2004)		•	•	
Kang (2004)	•	•	•	
Schaefer et al. (2005)				•
Lan, Clarke, and Barnhart (2006)	•		•	
Yen and Birge (2006)				•
Shebalov and Klabjan (2006)				•
Smith and Johnson (2006)		•		
AhmadBeygi, Cohn, and Lapp (2008)	•			
Gao, Johnson, and Smith (2009)		•		•
Eggenberg (2009)	•		•	
Burke et al. (2009)	•		•	
Marla and Barnhart (2010)			•	

Table 2.2: Robust schedules by planning phase

summarized in Table 2.2.

2.2 Performance Metrics for Airline Schedules

As in any complex system, there is no single *best* metric that captures every aspect of an airline's intricate operations. Different metrics cannot be used interchangeably. Also, different stakeholders (airlines, passengers, government, etc.) may be interested in different performance metrics. Passengers are concerned about an airline's service level, while some airlines may pay the most attention to the total operating cost of a schedule. We discuss here three performance metrics on which we will focus extensively throughout this work.

2.2.1 15-Minute On-Time Performance

15-minute On-Time Performance (15-OTP) measures a percentage of flights that departed/ arrived at the gate no later than 15 minutes after the scheduled departure/ arrival time as indicated in the Computerized Reservations Systems (CRS). It is a standard metric widely used in the airline industry because it is simple to compute and easy to understand, not only for people in the industry but also the general public. Additionally, the U.S. Department of Transportation (US DOT) also uses 15-OTP to evaluate airline performance and regularly publishes the rankings. As a result, many airlines focus on 15-OTP.

Despite the wide use of 15-OTP, it is not a very good metric for evaluating overall performance of an airline because:

1) It does not provide any information about the delay distribution. Several hours of delay is treated the same as a 15-minute delay. Therefore, given two airlines with the same 15-OTP, one can have a much larger average delay than the other.

2) It does not capture delay propagation in an airline network. One of the most important causes of delays in the U.S. is aircraft arriving late. Without sufficient

slack in a schedule, the delays due to late-arriving aircraft can propagate to subsequent flights in the same route.

3) It does not reflect passenger delays. Suppose a flight arrives 10 minutes late. Even though it is *on time* based on 15-OTP, a connecting passenger on that flight might miss his/her connection and have to wait for several hours for the next available flight. As a result, given two airlines with the same 15-OTP, one cannot tell which airline performs better with respect to passenger delays.

2.2.2 Delay Propagation[1]

The impact of delays in an airline network can be exacerbated when delays propagate. Because of airline network interconnectivity, a small delay caused by one flight leg can propagate and potentially lead to large delays on subsequent flight legs. It thus suggests that delay propagation might be a good measure to indicate the robustness of airline schedules.

A delay of each flight leg can be decomposed into two components.

1) **Propagated delay** This component of delay occurs when the aircraft to be used for a flight leg is delayed on its prior flight legs, and there is insufficient slack between the two flights to turn the aircraft. Note that propagated delay is a function of an aircraft routing.

2) **Nonpropagated delay** This component of delay captures all delays caused by reasons such as airborne delay or taxi delay. Because it is independent of an aircraft routing, we call it *an independent delay.*

Note that this definition of propagated delay only takes into account the delays due to aircraft arriving late. In reality, a flight may also experience propagated delays caused by crews or passengers.

Figure 2-1 illustrates the relationship between departures, arrivals, and delays of two flights i and j in the same routing. A solid arrow represents a planned departure

[1]The notion of delay propagation presented in this section is introduced in Lan, Clarke, and Barnhart (2006) [24].

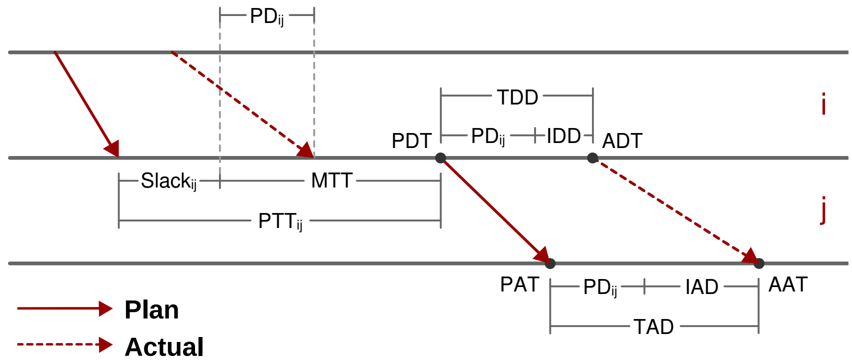


Figure 2-1: Flight delay breakdown

time (PDT) and a planned arrival time (PAT) of each flight. A dashed arrow represents an actual departure time (ADT) and an actual arrival time (AAT) of each flight. A planned turn time between flights i and j (PTT_{ij}) is the time between PAT_i and PDT_j. PTT_{ij} must be larger than the minimum turn time (MTT_{ij}) required for turning an aircraft. MTT_{ij} depends on the connection airport, fleet type, and other requirements for flights i and j. The additional time in PTT_{ij} in excess of MTT_{ij} is called slack ($Slack_{ij}$).

If the arrival delay of flight i is larger than $Slack_{ij}$, some portion of the delay cannot be absorbed and consequently propagates to flight j. Thus, the total departure delay (TDD) of flight j is composed of the propagated delay from flight i to flight j (PD_{ij}) and the independent departure delay (IDD) of flight j itself. Similarly, the total arrival delay (TAD) of flight j comprises PD_{ij} and the independent arrival delay (IAD).

Note that IDD captures only the independent delay before a flight is airborne (taxi-out delay, ground delay, etc.) whereas IAD includes both IDD and the additional independent delay in the air or at the destination airport. Also, IDD and IAD may take negative values if an airline expedites the ground process, flies a flight faster, or pads the schedule by increasing the block time to account for potential delays

Mathematically, we have the following relationships:

$$TDD = \text{Max}(ADT - PDT, 0) \tag{2.1}$$

$$TAD = \text{Max}(AAT - PAT, 0) \quad (2.2)$$

$$PTT_{ij} = PDT_j - PAT_i \quad (2.3)$$

$$Slack_{ij} = PTT_{ij} - MTT_{ij} \quad (2.4)$$

$$PD_{ij} = \text{Max}(TAD_i - Slack_{ij}, 0) \quad (2.5)$$

$$TDD_j = \text{Max}(PD_{ij} + IDD_j, 0) \quad (2.6)$$

$$TAD_j = \text{Max}(PD_{ij} + IAD_j, 0) \quad (2.7)$$

2.2.3 Passenger Delay

A passenger delay is measured by the difference between the planned arrival time and the actual arrival time at a passenger's final destination. A passenger's itinerary is called *disrupted* if one or more flights in his/her itinerary are canceled, or some connecting time between consecutive flights becomes less than the minimum connecting time (MCT) required for the passenger to proceed from the arrival gate to the departure gate of his/her subsequent flight leg.

Typically, flight delays underestimate passenger delays because a small flight delay may cause a passenger to miss his/her connection, and the passenger has to wait for several hours for the next available flight. Additionally, flight delay statistic does not reflect the extent of flight cancellations, which cause many disrupted passengers. Although the number of disrupted passengers might be very small, these disrupted passengers generally represent a large proportion of total passenger delay. For a detailed discussion about flight delays and passenger delays, readers are referred to Barnhart and Bratu (2005) [10].

According to the U.S. Airline Passenger Trip Delay Report 2008 by the Center for Air Transportation Systems Research at George Mason University [30], passengers on scheduled domestic U.S. airline flights were delayed a total of 299 million hours (34 thousand years) in 2008. Despite the 10% decrease in passenger delays from

2007, about one out of four passengers underwent a misconnection, cancelled flight, diverted flight, or denied boarding due to overbooking. The report suggests that it takes significantly longer with today's operations to re-book disrupted passengers because of the extensive use of smaller aircraft with high load factors. The average delay for passengers on cancelled flights is as long as 15 hours.

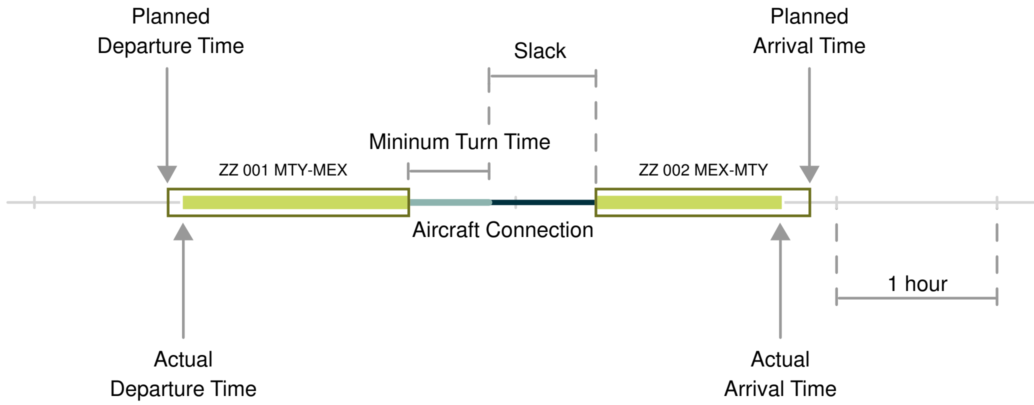
Consequently, it is increasingly important for airlines to pay attention to passenger delays and make the effort to cut down costs due to passenger re-accommodation, including compensation for failing to re-accommodate passengers in a timely manner and, importantly, strive to elevate passenger satisfaction.

2.3 Visualization

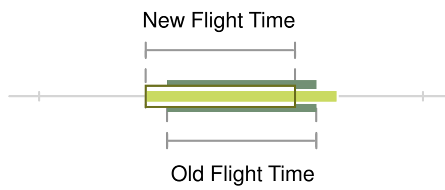
Performance metrics used to evaluate a schedule are typically in the form of aggregate statistic– a total, an average, a maximum/minimum, etc. Although these numbers can provide a good idea of a schedule's performance, most of the time they obscure many other useful details (distributions, patterns, etc.), and they are thus not very helpful in understanding the causes and effects of delays in operations.

To facilitate understanding of a schedule's performance, we develop and make use of visualization in addition to data tables and aggregate statistics. Showing all key pieces of information on the same page, the visualization allows us to easily see interactions among all components and understand how delays and disruptions affect aircraft and passenger connections. Additionally, the visualization is of great help in characterizing and comparing the robust schedules obtained from various optimization models with different formulations, objective functions, or other parameters.

Figure 2-2 illustrates the notations we use in the visualization. A blank outer rectangle denotes a planned flight time of each flight; a filled inner rectangle denotes an actual flight time of each flight; a line connecting two flights represents an aircraft connection. As discussed in Section 2.2.2, each aircraft connection is composed of a minimum turn time and slack. Later in this thesis, we will discuss robust flight re-timing models. The notation for flight re-timing is illustrated in Figure 2-2b.



(a) Flights and aircraft connections



(b) Flight re-timing

Figure 2-2: Visualization notation

To demonstrate our use of visualization, we present a small example in Figure 2-3. Visualization allows us to easily observe many attributes of a schedule at the same time.

We can make the visualization more informative by introducing lines in different styles to represent crew connections or passenger connections. This enables us to see how flight delays affect crew and passenger connections in an airline network. To further facilitate analysis, we allow users to view historical data of a flight (delay distribution, planned block time distribution, etc.) or some key information such as disrupted passenger connections by clicking or pressing some specific key.

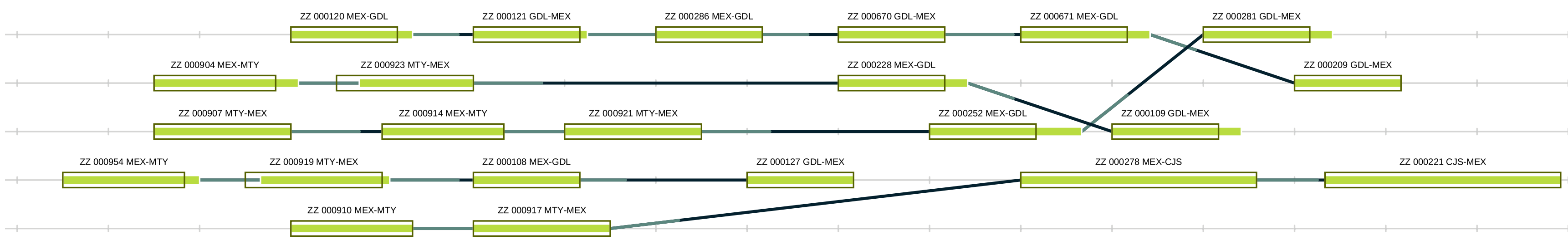


Figure 2-3: Visualization example

This figure is a visualization of a subnetwork operated with five aircraft. A planned routing for each aircraft is a sequence of flights along each horizontal line. An actual routing is a sequence of flights connected by blue lines. We can easily observe

- independent delays, e.g., ZZ 252, which departs on time but arrives late;
- delay propagation, e.g., from ZZ 904 to ZZ 923 at MTY airport;
- aircraft swaps indicated by inclined aircraft connection lines. For example, the first aircraft, originally assigned to fly the flights along the top line, flies the first five flights as planned, then it is swapped to fly the last flight of the second aircraft's route;
- tight connections– ones with little or no slack, e.g., the connection between ZZ 910 and ZZ 917;
- utilization of block time slack, e.g., ZZ 923, although the flight departs late, it still manages to arrive on time;
- utilization of ground time slack at the connection, e.g., between ZZ 120 and ZZ 121, although ZZ 120 arrives late, there is sufficient ground time to absorb the delay, and ZZ 121 can still depart on time.

2.4 Slack Re-Allocation in Robust Airline Scheduling

Slack in an airline schedule is additional time allocated beyond the minimum time required for each aircraft connection, passenger connection, or expected flight duration. Slack is desirable in robust schedules as it can 1) potentially absorb delays in an airline network; 2) reduce the likelihood of operational disruptions; and 3) provide flexibility to recover once the operation is disrupted.

Slack in different components of an airline schedule serves different purposes.

1) **Aircraft connection slack (ground time slack)** is additional ground time beyond the minimum turn time of each aircraft connection. The amount of aircraft connection slack in a schedule is a function of an aircraft routing. Aircraft connection slack can be used to absorb accumulated flight delays from prior flights along the aircraft route and thus reduce a likelihood of delay propagation to subsequent flights.

2) **Passenger connection slack** is additional time beyond the minimum connection time between two flight legs in a passenger's itinerary. It is a function of the arrival time of an inbound flight and the departure time of an outbound flight. Hence, the amount of passenger connection slack in a schedule depends on flight schedules and passenger itineraries allowed for booking. Passenger connection slack plays an important role in decreasing the chance of passenger misconnection.

3) **Block time slack** is additional time added to the expected block time of each flight. It is a function of a flight's departure and arrival times. Therefore, the amount of block time slack in a schedule depends on flight schedules.

Although both block time slack and aircraft connection slack can be used to absorb flight delays, they work differently. Block time slack provides greater flexibility compared to aircraft connection slack. It can absorb propagated delay from the preceding flight, taxi delay (at both departure and arrival airports), and airborne delay; while ground time slack can absorb only propagated delay from the preceding flight.

To illustrate the difference, consider a flight leg departing from a busy airport, and suppose we are allowed to change only the departure time of this flight. Moving the departure time earlier is equivalent to increasing the block time and decreasing the ground time preceding the flight. Thus, ground time slack is transformed into block time slack. Conversely, moving the departure time later is equivalent to transforming block time slack into ground time slack. Because in this particular case the aircraft is expected to spend a reasonable amount of time on the runway awaiting the departure queue, an airline is better off moving the arrival time earlier in order to queue up for the departure slot as soon as possible and letting the block time slack absorb the delay instead of using the ground time slack, waiting at the gate and then incurring the delay in the departure queue. In this latter case, even though the ground slack can absorb all the propagated delay from the prior flights, and the flight departs on time, the aircraft still has to spend a long time in the queue and ends up arriving late at the destination.

Despite the advantages of slack in a schedule, it is, on the other hand, considered a waste of resources from an airline perspective, which focuses mainly on cost minimization. Airlines have made numerous efforts to increase the utilization of all resources in airline operations and consequently reduce slack in a schedule.

The role of slack in the trade-off between aircraft and crew productivity and saving on costs due to disruptions incurred during a day of operation is depicted in Figure 2-4. There is a cost associated with slack in an airline schedule. In an extreme case, a schedule with an abundant amount of slack in a schedule may require more aircraft and crews to operate the schedule. As the amount of slack in a schedule increases, the planned costs associated with the schedule increase; the recovery costs, however, decrease because a schedule with more slack is likely to be more robust and results in fewer delays and disruptions. Finally, it is important to note that after a certain amount of slack is added into a schedule, the saving gained from additional slack in

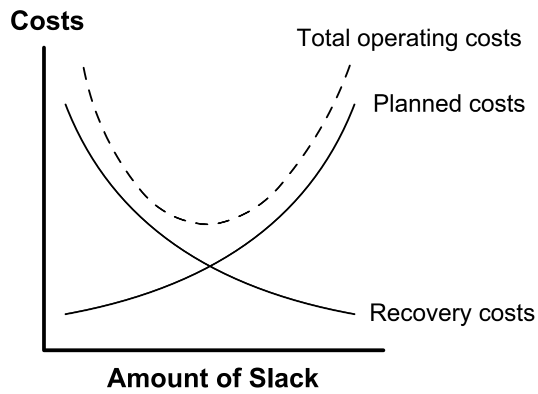


Figure 2-4: Trade-off between amount of slack and recovery costs

a schedule may not make up for the increase in planned costs because most of the time, slack does not get fully utilized.

Therefore, the recent trend in robust airline scheduling is to re-allocate, rather than simply increase, the existing slack in the schedules such that the resulting distribution of slack is more effective in absorbing delays and minimizing disruptions. We summarize here three major schemes of slack re-allocation.

2.4.1 Aircraft Re-Routing

In an aircraft re-routing problem, a flight schedule and fleet assignment are fixed, i.e., arrival and departure times of every flight remain the same as the original schedule, but the aircraft tail assignment of each flight can be changed. As a result, the modified routing yields a different distribution of aircraft connection slack.

2.4.2 Flight Schedule Re-timing

In a flight schedule re-timing problem, an aircraft routing and fleet assignment are fixed, but the departure time of each flight is allowed to change within a small time window. An arrival time of each flight must change by the same amount as the departure time, i.e., the block time of each flight is fixed. When a flight is moved earlier, slack in the aircraft connection preceding the flight decreases; while slack in the aircraft connection succeeding the flight increases. Because a flight schedule is allowed to change, it does not only affect aircraft connection slack, but also passenger connection slack.

2.4.3 Block Time Adjustment

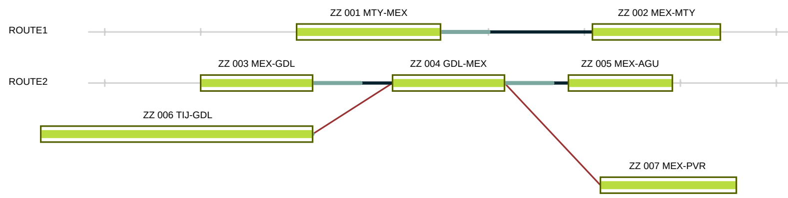
In a block time adjustment problem, an aircraft routing and fleet assignment are again fixed, but *both* departure and arrival times of each flight are allowed to change independently. Therefore, in addition to aircraft connection slack and passenger connection slack, it also affects block time slack. In particular, ground time slack can be transformed into block time slack.

2.4.4 Slack Re-Allocation Example

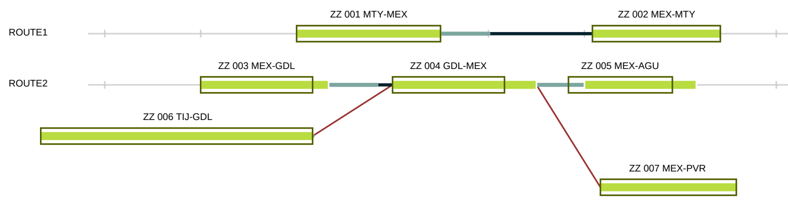
The following example illustrates how each slack reallocation scheme works. This schedule contains seven flights operated by four aircraft (see Table 2.3). A minimum aircraft turn time and a minimum passenger connection time are assumed to be 30 minutes. There are two passenger connections, ZZ 006-ZZ 004 and ZZ 004-ZZ 007, indicated by thin red lines in Figure 2-5a. Suppose that the expected independent delays of flight ZZ 003, ZZ 004, and ZZ 005 are 10, 20 and 5 minutes, respectively. Figure 2-5b suggests that the slack in the aircraft connection between ZZ 004 and ZZ 005 is not sufficient to absorb the delay from ZZ 004, and ZZ 005 is thus expected to experience the delay propagated from ZZ 004.

Aircraft re-routing

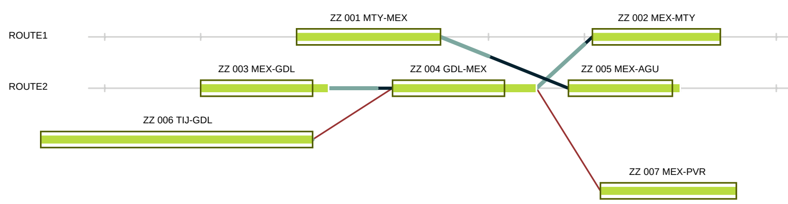
In Figure 2-5c, we modify the original aircraft routing such that the first aircraft flies ZZ 001 and ZZ 005; and the second aircraft flies ZZ 003, ZZ 004, and ZZ 002. Because ZZ 004 is now followed by ZZ 002, which departs 15 minutes later than ZZ 005, the arrival delay of ZZ 004 does not propagate anymore. In particular, the ground time slack in the schedule is re-allocated as follows:



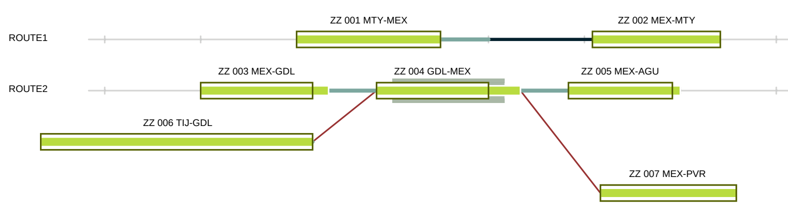
(a) Original schedule



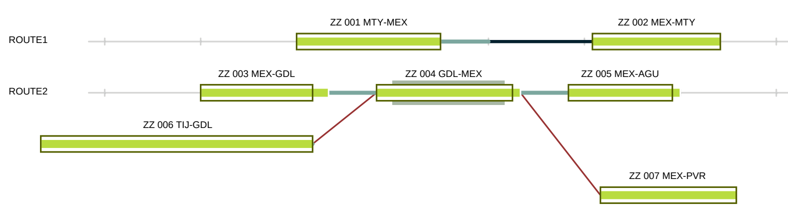
(b) Original schedule with expected delays



(c) Re-routing



(d) Re-timing



(e) Block time adjustment

Figure 2-5: Slack re-allocation example

Route	Flight	Departure Airport	Arrival Airport	Departure Time	Arrival Time	Block Time	Connection Time
1	ZZ 001	MTY	MEX	09:00	10:30	90	-
1	ZZ 002	MEX	MTY	12:05	13:25	80	95
2	ZZ 003	MEX	GDL	08:00	09:10	70	-
2	ZZ 004	GDL	MEX	10:00	11:10	70	50
2	ZZ 005	MEX	AGU	11:50	12:55	65	40
3	ZZ 006	TIJ	GDL	06:20	09:10	170	-
4	ZZ 007	MEX	PVR	12:10	13:35	85	-

Table 2.3: Slack re-allocation example

	Connections	Connection times	Ground time slack
Original	ZZ 001 - ZZ 002	95	65
	ZZ 004 - ZZ 005	40	**10**
Modified	ZZ 001 - ZZ 005	80	50
	ZZ 004 - ZZ 002	55	**25**

In the modified schedule, although the total amount of slack remains the same as in the original schedule, more slack is allocated to the connection following flight ZZ 004, which is expected to have a long arrival delay. As a result, the ground time slack can absorb the arrival delay from flight ZZ 004. Note that because the flight schedule is fixed in the aircraft re-routing problem, passenger connection slack is unaffected.

Flight Re-timing

In Figure 2-5d, flight ZZ 004 is shifted 10 minutes earlier. The distribution of ground slack changes as follows:

	Connections	Connection times	Ground time slack
Original	ZZ 003 - ZZ 004	50	20
	ZZ 004 - ZZ 005	40	**10**
Modified	ZZ 003 - ZZ 004	40	10
	ZZ 004 - ZZ 005	50	**20**

This simply moves 10 minutes of ground time slack from the connection preceding ZZ 004 to the connection succeeding ZZ 004. As a result, there is sufficient ground time slack to absorb the arrival delay from flight ZZ 004, and flight ZZ 005 can depart on time. Note that if we moved the departure time of ZZ 004 back further, there would not be adequate slack to absorb the arrival delay of ZZ 003, and the delay would start propagating to ZZ 004.

Because the flight schedule is changed, the passenger connection times, and hence passenger connection slack, are affected. In particular, the passenger connection time for ZZ 006-ZZ 004 becomes shorter; while it becomes longer for ZZ 004-ZZ 007.

Block Time Adjustment

In Figure 2-5e, the departure time of flight ZZ 004 is moved 10 minutes earlier, and the arrival time is moved 5 minutes later. Therefore, the block time increases by 15 minutes. The resulting distribution of slack can be summarized as follows:

	Connections	Connection times	Ground time slack	Block time slack (ZZ 004)
Original	ZZ 003 - ZZ 004	50	20	-
	ZZ 004 - ZZ 005	40	10	
Modified	ZZ 003 - ZZ 004	40	10	**+15**
	ZZ 004 - ZZ 005	35	5	

Moving the departure time of ZZ 004 10 minutes earlier yields the same effect as in the flight re-timing case, and hence no delay propagates from ZZ 004 to ZZ 005. Moving the arrival time of ZZ 004 5 minutes later, however, converts 15 minutes of ground time slack into block time slack. Consequently, the increased block time slack absorbs most of the independent delay of ZZ 004, and only 5 minutes of ground time slack in the ZZ 004-ZZ 005 connection is needed to absorb the rest of the delay. Although flight ZZ 004 is expected to arrive at the same time as in the flight re-timing case, the arrival delay of ZZ 004 is as small as 5 minutes. This illustrates how schedule padding helps airlines improve their on-time performance.

Lastly, because flight ZZ 004 is scheduled to arrive later than the original schedule, the scheduled passenger connection time for ZZ 004-ZZ 007 becomes smaller.

In the next chapter, we will show how each slack re-allocation scheme can be formulated as an optimization model to minimize expected delays.

Chapter 3

Optimization Models for Slack Re-allocation

3.1 Robust Aircraft Re-routing Model

The aircraft re-routing model presented in this section is a modification of the formulation developed by Lan, Clarke, and Barnhart (2006) [24].

3.1.1 Underlying idea

As discussed in Section 2.2.2, delay propagation can exacerbate the impact of delays in an airline network. Because propagated delay is a function of aircraft routes, we can reduce delay propagation in the system by cleverly re-routing aircraft as demonstrated in Section 2.4.4. In particular, we seek to re-route aircraft such that ground time slack is optimally allocated to the connections that historically cause delay propagation, and the resulting aircraft routes minimize the expected total propagated delay with respect to a given set of historical data.

3.1.2 Formulation

We first introduce the notations used in this formulation:

S : set of feasible strings

F : set of flight legs

M^+ : set of starting airports

M^- : set of ending airports

S_{m^+} : set of strings $s \in S$ starting at airport $m^+ \in M^+$

S_{m^-} : set of strings $s \in S$ ending at airport $m^- \in M^-$

pd_{ij}^s = propagated delay from flight leg $i \in F$ to the succeeding flight leg $j \in F$ in string $s \in S$

N_{m^+} = number of aircraft starting at airport $m^+ \in M^+$ in the original aircraft routes

N_{m^-} = number of aircraft ending at airport $m^- \in M^-$ in the original aircraft routes

$$a_{is} = \begin{cases} 1 & \text{if flight leg } i \in F \text{ is in string } s \in S \\ 0 & \text{otherwise} \end{cases}$$

Note that the departure airport of the first flight in each original aircraft route is called a *starting airport*, and the arrival airport of the last flight in each original aircraft route is called an *ending airport*.

Because the flight schedules in our dataset do not repeat daily for most flights, the airline treats the aircraft routing problem as a dated problem, as opposed to a cyclic problem which can be repeated over a certain planning horizon. Therefore, the original formulation by Lan, Clarke, and Barnhart (2006) is not applicable here.

Rather than formulating an aircraft maintenance routing problem, we focus on *re-routing* aircraft on a given day of operation, assuming that the resulting aircraft routes do not violate maintenance feasibility. In particular, we want to fly the same set of flight legs using the aircraft that are ready at each starting airport. Note that this will automatically ensure that the number of aircraft at each ending airport is the same as in the original schedule, and hence there will be sufficient aircraft to operate the next day's schedule at each ending airport. In addition, we assume further that each aircraft that is scheduled to fly on that day is ready at the beginning of the day and also available until the end of the day. Because the fleet assignment is fixed, we can solve this problem separately for each fleet type.

Because delays propagate along aircraft routes, it is more convenient to model the robust aircraft re-routing problem using a string-based formulation. In this case, a string s is defined as a sequence of flight legs that begins at some starting airport $m^+ \in M^+$ and ends at some ending airport $m^- \in M^-$. Each string s represents a set of flight legs that are operated by a single aircraft on a given day of opeartion. Let x_s be a binary decision variable which takes value 1 if string $s \in S$ is included in the optimal solution, and 0 otherwise. The robust aircraft re-routing problem (AR) can be formulated as follows:

$$\text{Minimize} \quad \mathbb{E}\left[\sum_{s\in S}\left(x_s \times \sum_{(i,j)\in S} pd^s_{ij}\right)\right] \tag{AR-1}$$

$$\text{subject to} \quad \sum_{s\in S} a_{is}x_s = 1 \qquad \forall i \in F \tag{AR-2}$$

$$\sum_{s\in S_{m^+}} x_s = N_{m^+} \qquad \forall m^+ \in M^+ \tag{AR-3}$$

$$x_s \in \{0,1\} \qquad \forall s \in S \tag{AR-4}$$

The objective function (AR-1) is to minimize the expected total propagated delay of all strings in the solution. Constraints (AR-2) ensure that each flight leg in the schedule is covered by exactly one string in the solution. Constraints (AR-3) guarantee that the number of strings beginning at each starting airport in the solution is equal to the number of aircraft available at that airport at the beginning of the day.

Given the constraints (AR-2) and (AR-3), it is guaranteed that the number of strings ending at each ending airport in the solution is equal to the number of aircraft required at that airport for the next day's schedule, mathematically, $\sum_{s \in S_{m^-}} x_s = N_{m^-}, \forall m^- \in M^-$.

Because the random variable pd_{ij}^s only appears in the objective function, we can rewrite the objective function as:

$$\mathbb{E}\left[\sum_{s \in S}\left(x_s \times \sum_{(i,j) \in S} pd_{ij}^s\right)\right] = \sum_{s \in S}\left(x_s \times \mathbb{E}\left[\sum_{(i,j) \in S} pd_{ij}^s\right]\right) \tag{AR-5}$$

In words, the expected total propagated delay can be simply computed as the total expected propagated delay. The cost coefficient, $\mathbb{E}\left[\sum_{(i,j) \in S} pd_{ij}^s\right]$, associated with each string s can be calculated offline. Therefore, the AR model is a deterministic mixed integer program.

3.1.3 Computing Objective Function Coefficients for Feasible Strings

Although, for any feasible string that has been operated, the total arrival delay of each flight leg and the propagated delay on each connection in the string can be computed directly from historical data, this is not applicable to the feasible strings that have never been operated before. We, however, can still compute the independent arrival delay for each flight leg from historical data because it is not a function of aircraft routing. This overcomes the difficulty in modeling total arrival delays and propagated delays in a string-based formulation.

Algorithm 1 describes how we can determine, for any feasible string, the total arrival delay of each flight leg and the propagated delay on each connection.

Let Ω be a set of possible delay scenarios. We assume Ω has finite cardinality, and each $\omega \in \Omega$ occurs with probability p_ω. Each objective function coefficient can be rewritten as:

$$\mathbb{E}\left[\sum_{(i,j) \in S} pd_{ij}^s\right] = \sum_{\omega \in \Omega} p_\omega \sum_{(i,j) \in S} pd_{ij}^s(\omega) \tag{AR-6}$$

Algorithm 1 Compute total arrival delays and propagated delays for a feasible string

1) For each pair of consecutive flight legs i and j flown by the same aircraft in historical data, the propagated delay from flight i to flight j is given by $PD_{ij} = \text{Max}(TAD_i - Slack_{ij}, 0)$.

2) For each flight leg j,

    - if j is the first flight of a string, $IAD_j = TAD_j$;
    - otherwise, $IAD_j = TAD_j - PD_{ij}$.

3) Using the independent arrival delays computed in the previous step, we can determine the total arrival delay of each flight leg and the propagated delay on each connection for any feasible string s' as follows:

    - For the first flight leg j in the string, we have $TAD'_j = \text{Max}(IAD_j, 0)$, assuming that the first flight of each string has zero propagated delay.
    - For each subsequent flight j following flight i in the string, we have $PD'_{ij} = \text{Max}(TAD'_i - Slack_{ij}, 0)$ and $TAD'_j = \text{Max}(IAD_j + PD'_{ij}, 0)$.

---

where $pd^s_{ij}(\omega)$ is the propagated delay from flight leg i to the succeeding flight leg j in string s for a given delay scenario ω.

Note that this calculation of the objective function coefficients is different from the original work by Lan, Clarke, and Barnhart. They first model total arrival delays using a lognormal distribution whose parameters, for a given set of historical data, can be estimated using Maximum Likelihood Estimation (MLE). Then, the expected propagated delay for each pair of consecutive flights is given by $(1 - \Phi(\frac{ln(\frac{-\theta}{m})}{\sigma}))(\theta + me^{\frac{1}{2}\sigma^2})$, where $\Phi(\cdot)$ is the cumulation distribution function of the standard normal distribution; and σ, θ, and m are the shape parameter, the location parameter, and the scale parameter of the total arrival delay distribution of the first flight. The detailed calculation is described in Lan (2003) [23].

3.1.4 Alternative Objective Functions

Maximizing the total expected effective aircraft connection slack

Recall that the propagated delay from flight i to flight j is defined as $PD_{ij} = \text{Max}(TAD_i - Slack_{ij}, 0)$. It takes a positive value only when the total arrival delay of flight i exceeds the planned slack in that connection as depicted in Figure 3-1. As a result, the propagated delay PD_{ij} obscures the arrival delay of flight i when $TAD_i < Slack_{ij}$. In particular, given two solutions with $PD_{ij} = 0$ but different TAD_i, the model that minimizes the total propagated delay cannot distinguish between the two, even though the one with smaller TAD_i is more desirable.

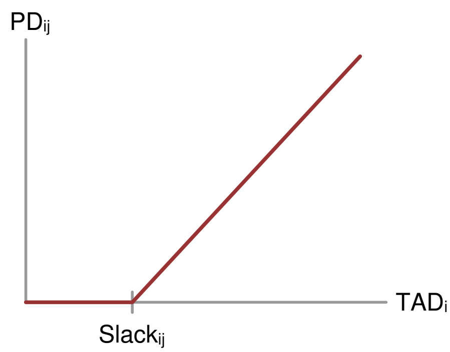


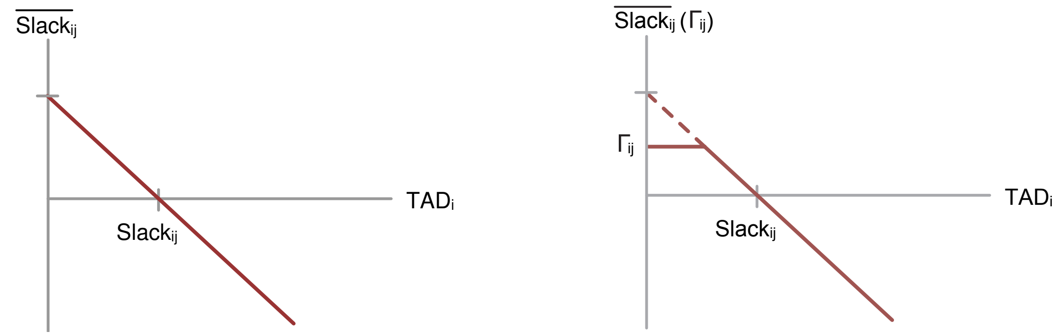
Figure 3-1: Propagated delay from flight i to j versus total arrival delay of flight i

To overcome this difficulty, we introduce the notion of *effective slack*. Let i and j be two consecutive flights in the same string. We define the effective slack in the connection between flights i and j ($\overline{Slack}_{ij}$) as:

$$\overline{Slack}_{ij} = Slack_{ij} - TAD_i. \quad \text{(AR-7)}$$

In other words, the effective slack in each connection represents the remaining slack after accounting for the arrival delay of the inbound flight (see Figure 3-2a). Note that effective slack may take a negative value. In this case, the arrival delay of the inbound flight will propagate to the outbound flight.

The proposed objective function is to maximize the total expected effective slack. To ensure that the model has no incentive to add more slack to the connections



(a) Effective slack

(b) Effective slack capped at Γ_{ij}

Figure 3-2: Effective slack versus total arrival delay

that already have a reasonable amount of slack, we introduce another parameter– a nonnegative *cap* Γ_{ij} for each aircraft connection from flight i to flight j. We then redefine the effective slack to the minimum of Γ_{ij} and the difference of the planned slack and the total arrival delay associated with the aircraft connection from flight i to flight j. Specifically,

$$\overline{Slack}_{ij}(\Gamma_{ij}) = \text{Min}(Slack_{ij} - TAD_i, \Gamma_{ij}). \tag{AR-8}$$

Therefore, any aircraft connection from flight i to flight j with effective slack more than the protection level Γ_{ij} contributes only Γ_{ij} minutes to the objective function. This results in the allocation of more slack to connections for which the expected effective slack is smaller than the corresponding cap. In addition, because caps are specific to aircraft connections, we can set them to different levels for different fleet types, connection airports, and so forth.

In summary, the proposed objective function is given by:

$$\begin{aligned} \text{Maximize} \qquad & \mathbb{E}\left[\sum_{s\in S}\left(x_s \times \sum_{(i,j)\in S} \overline{Slack}^{s}_{ij}(\Gamma_{ij})\right)\right] \\ & = \sum_{s\in S}\left(x_s \times \mathbb{E}\left[\sum_{(i,j)\in S} \overline{Slack}^{s}_{ij}(\Gamma_{ij})\right]\right) \end{aligned} \tag{AR-9}$$

where $\overline{Slack}^{s}_{ij}(\Gamma_{ij})$ is the effective slack capped at Γ_{ij} in the connection between flights i and j in a string s.

Each objective function coefficient $\mathbb{E}\left[\sum_{(i,j)\in S} \overline{Slack}^s_{ij}(\Gamma_{ij})\right]$ associated with x_s can be computed using the same approach outlined in Section 3.1.3.

In fact, the objective function of minimizing total expected propagated delay in (AR-1) is a special case of the proposed objective function in (AR-9) with caps set equal to zero for all aircraft connections. In particular, for any aircraft connection from flight i to flight j, we have

$$\begin{aligned}\overline{Slack}_{ij}(\Gamma_{ij}=0) &= \text{Min}(Slack_{ij} - TAD_i, 0)\\ &= -\text{Max}(TAD_i - Slack_{ij}, 0)\\ &= -PD_{ij}\end{aligned}$$

Minimizing the total expected arrival delay

Even though, at first glance, minimizing the total expected propagated delay may seem equivalent to minimizing the total expected arrival delay, this is true only when independent arrival delays (IAD) of every flight is nonnegative. Recall that we define a propagated delay (PD) and a total arrival delay (TAD) of each flight such that they only take nonnegative values (see (2.5) and (2.7)), but an IAD may take a negative value if an airline expedites the ground process, flies a flight faster, or pads the schedule by increasing the block time to account for potential delays. Consequently, it is possible that a flight with a negative IAD experiences some propagated delay and still arrives on-time or earlier. This leads to the discrepancy between the two objectives. When delays in the system are unavoidable, it might be desirable to propagate delay to those flights with negative IADs and let their block time slack help absorb the delays.

Figure 3-3 illustrates the idea. In the figure, a number associated with each flight indicates the independent arrival delay of that flight. The amount of slack and corresponding propagated delay in each connection are as follows:

Feasible solution	Connections	Ground time slack	Propagated delay
I	ZZ 004 - ZZ 005	5	15
	ZZ 008 - ZZ 009	20	0
II	ZZ 004 - ZZ 009	9	11
	ZZ 008 - ZZ 005	16	0

The feasible solution II minimizes total propagated delay in this case. However, because the IAD of ZZ 005 is -10 minutes (equivalently, ZZ 005 has block time slack of 10 minutes), the total arrival delay for the feasible solution I is only 5 minutes, whereas in the feasible solution II, the total arrival delay is 11 minutes, which is the propagated delay from ZZ 004 to ZZ 009. Therefore, the feasible solution I minimizes the total arrival delay, although it has a higher total propagated delay.

(a) Feasible solution I (b) Feasible solution II

Figure 3-3: Minimizing total arrival delay versus minimizing total propagated delay

In conclusion, if the airline under consideration extensively pads its schedule, it might be more appropriate to minimize the total expected arrival delay instead of the total expected propagated delay, as propagated delay is to a large extent absorbed without impact to downstream flights.

The objective function of maximizing the total expected arrival delay is given by

$$\mathbb{E}\left[\sum_{i\in S} tad_i^s\right] = \sum_{\omega\in\Omega} p_\omega \sum_{i\in S} tad_i^s(\omega). \tag{AR-10}$$

We replace $pd_{ij}^s(\omega)$ in (AR-6) with $tad_{ij}^s(\omega)$, the total arrival delay from flight i to the succeeding flight j in string s for a given delay scenario ω, and the values of $tad_{ij}^s(\omega)$ can be obtained from Algorithm 1, presented in Section 3.1.3.

(a) Feasible solution I

(b) Feasible solution II

Figure 3-4: Multiple optimal solutions in the robust aircraft re-routing problem

Each figure above shows the aircraft routes corresponding to a feasible solution. The numbers indicate the expected independent delays of each flight. We can see that both feasible solutions have zero total expected propagated delay and are thus optimal with respect to the objective function (AR-1). The feasible solution II, however, may be preferable to the feasible solution I in the sense that the latter contains a tight connection from ZZ 004 to ZZ 005.

3.1.5 Multiple Optimal Solutions

Solving the AR problem with either the original objective function (AR-1) or the proposed objective function (AR-9) typically yields multiple optimal solutions. The number of optimal solutions depends mainly on 1) re-routing opportunities in the airline schedule which are impacted by fleet homogeneity, connecting banks at hubs, etc.; and 2) the extent of delays in the given set of historical data. Specifically, smaller level of delay in the system leads to a larger number of optimal solutions. Figure 3-4 illustrates the idea for the original objective function (AR-1). If the delay, however, in the system is extensive, some aircraft connections are required in the optimal solution in order to minimize the total expected delay. Therefore, the number of optimal solutions is decreased.

Although these optimal solutions give the same objection function value, they might not be equally effective with respect to other performance metrics or objective functions. Therefore, we desire to select, among the optimal solutions to the initial objective function, the solution that is optimal with respect to another objective function.

Suppose z_f is the optimal objective function value of the AR problem with an objective function of minimizing $f(\mathbf{x})$, where $\mathbf{x}$ is a vector of binary decision variables x_s associated with each string $s \in S$. We add a constraint of the form $f(\mathbf{x}) \leq z_f$ to the original AR formulation and replace the objective function f with another objective function. If there still exist multiple optimal solutions, we can repeat this process sequentially for other objective functions that are of interest.

We emphasize that this sequential approach is aimed at finding the "best" solution among the optimal solutions to the initial objective function, rather than finding the solution that performs best with respect to all objective functions overall. Otherwise, multi-criteria optimization is more applicable.

Because the original aircraft routing is presumably optimized with respect to some objectives that are of interest to the airline, given a set of optimal solutions to our objective function, we may want to select the solution that is closest to the original one to preserve the good features of the original aircraft routing. One possible way to minimize the difference between the original and the optimal aircraft routing is to maximize the number of aircraft connections in the original routing that are included in the optimal solution.

Let C be the set of aircraft connections in the original routing and b_{cs} take value 1 if a string $s \in S$ contains an aircraft connection $c \in C$, and 0 otherwise. The objective function of maximizing the number of aircraft connections in the original routing that are included in the optimal solution is given by:

$$\text{Maximize} \qquad \sum_{s \in S} \left(x_s \times \sum_{c \in C} b_{cs} \right). \tag{AR-11}$$

3.2 Robust Flight Schedule Re-timing Model

3.2.1 Underlying Idea

As discussed in Section 2.4, re-timing flight departure times can affect both aircraft connection slack and passenger connection slack. Moving a flight departure time later increases the amount of slack in the connection preceding the flight and simultaneously decreases the amount of slack in the connection succeeding the flight. Therefore, given a set of historical data, we want to re-time flight departure times optimally such that the resulting schedule minimizes some proxy of expected delays or disruptions as demonstrated in Section 2.4.4. In particular, we want to re-allocate the existing slack in the schedule to the connections that historically often cause delays.

In the next section, we present the robust flight schedule re-timing model that minimizes the total expected propagated delay. Then, we propose alternative objective functions in subsequent sections.

3.2.2 Formulation

We first introduce the notations used in this formulation:

F : set of flight legs

F_0 : set of first flight leg in each aircraft route, for all aircraft routes

A : set of aircraft connections

P : set of passenger connections

Ω : set of possible delay scenarios

p_ω : probability that a delay scenario $\omega \in \Omega$ occurs

pd_{ij} : propagated delay from flight leg $i \in F$ to the succeeding flight leg $j \in F$

$aSlack_{ij}$: original planned aircraft connection slack in an aircraft connection $(i, j) \in A$

$pSlack_{ij}$: original planned passenger connection slack in a passenger connection $(i, j) \in P$

IAD_i^ω : independent arrival delay of flight $i \in F$ for a given delay scenario $\omega \in \Omega$

pd_{ij}^ω : propagated delay from flight leg $i \in F$ to the succeeding flight leg $j \in F$ for a given delay scenario $\omega \in \Omega$

tad_i^ω : total arrival delay of flight leg $i \in F$ for a given delay scenario $\omega \in \Omega$

$aSlack'_{ij}$: resulting planned aircraft connection slack in an aircraft connection $(i,j) \in A$ after re-timing

$pSlack'_{ij}$: resulting planned passenger connection slack in a passenger connection $(i,j) \in P$ after re-timing

Let x_i be the difference between the new and the original departure time of flight $i \in F$. Hence, x_i takes a negative value if the departure time is moved earlier and takes a positive value if the departure time is moved later. We limit the change in the departure time of each flight i within a small time window $[l_i, u_i]$, and assume that the demand for flight i remains the same within this range.

Using the notions of delay propagation described in Section 2.2.2, the robust flight re-timing problem (`FR`) is given by:

$$\text{Minimize} \qquad \sum_{(i,j)\in A} \mathbb{E}\left[pd_{ij}\right] = \sum_{(i,j)\in A} \left(\sum_{\omega\in\Omega} p_\omega pd_{ij}^\omega \right) \tag{FR-1}$$

$$\text{subject to} \qquad aSlack'_{ij} = aSlack_{ij} - x_i + x_j \qquad \forall (i,j) \in A \tag{FR-2}$$

$$aSlack'_{ij} \geq 0 \qquad \forall (i,j) \in A \tag{FR-3}$$

$$pSlack'_{ij} = pSlack_{ij} - x_i + x_j \qquad \forall (i,j) \in P \tag{FR-4}$$

$$pSlack'_{ij} \geq 0 \qquad \forall (i,j) \in P \tag{FR-5}$$

$$pd_{ij}^\omega \geq tad_i^\omega - aSlack'_{ij} \qquad \forall (i,j) \in A, \forall \omega \in \Omega \tag{FR-6}$$

$$pd_{ij}^\omega \geq 0 \qquad \forall (i,j) \in A, \forall \omega \in \Omega \tag{FR-7}$$

$$tad_i^\omega \geq IAD_i^\omega \qquad \forall i \in F_0, \forall \omega \in \Omega \tag{FR-8}$$

$$tad_j^\omega \geq pd_{ij}^\omega + IAD_j^\omega \qquad \forall (i,j) \in A, \forall \omega \in \Omega \tag{FR-9}$$

$$tad_i^\omega \geq 0 \qquad \forall i \in F, \forall \omega \in \Omega \tag{FR-10}$$

$$l_i \leq x_i \leq u_i \qquad \forall i \in F \tag{FR-11}$$

$$x_i \in \mathbb{Z}^n \qquad \forall i \in F \tag{FR-12}$$

The objective function (FR-1) is to minimize the total expected propagated delay over all aircraft connections. Again, we assume Ω has finite cardinality.

Constraint set (FR-2) is the resulting planned aircraft connection slack after re-timing for each aircraft connection. Specifically, for an aircraft connection $(i, j) \in A$, the resulting planned aircraft connection slack ($aSlack'_{ij}$) increases when the departure time of flight i is moved earlier, i.e., when x_i takes a negative value. Also, it increases when the departure time of flight j is moved later, i.e., when x_j takes a positive value. The non-negativity constraints of the resulting planned aircraft connection slack, (FR-3), ensure that every aircraft connection is longer than its corresponding minimum aircraft turn time, and thus the current aircraft routing remains feasible.

Similarly, constraint set (FR-4) is the resulting planned passenger connection slack after re-timing for each passenger connection. The non-negativity constraints of the resulting planned passenger connection slack, (FR-5), ensure that every passenger connection is longer than its corresponding minimum passenger connection time, and thus every itinerary remains feasible.

Given a re-timed flight schedule, constraints (FR-6) and (FR-7) are the propagated delays for each aircraft connection under different delay scenarios; constraints (FR-8)-(FR-10) determine the total arrival delays for each flight leg under different delay scenarios, assuming that the first flight of each string has zero propagated delay. We will shortly discuss the correctness of this calculation of total arrival delays and propagated delays.

Lastly, constraints (FR-11) limit the change in the departure time of each flight i within a specific time window $[l_i, u_i]$.

Remarks:

- Constraints (FR-2) and (FR-4) are indeed redundant, and we can replace $aSlack'_{ij}$ and $pSlack'_{ij}$ everywhere with the right-hand sides of (FR-2) and (FR-4). We intentionally write them this way to facilitate understanding of the model and keep the formulation clean.

- An independent arrival delay of flight i for a given delay scenario ω, IAD_i^ω can be computed using the first two steps of Algorithm 1 in Section 3.1.3.
- The calculation of total arrival delays and propagated delays in this formulation is similar to the one presented in Section 2.2.2.
- Because passenger connections may involve two flights flown by aircraft in different fleet types, the `FR` problem must be solved for all fleet types simultaneously. (Recall that in the string-based formulation for the `AR` problem, we can solve the problem separately for each fleet type.)

To show the correctness of the formulation, we prove the following claim.

Claim: In the optimal solution, every variable pd_{ij}^ω correctly models the propagated delay of an aircraft connection $(i, j) \in A$ under a delay scenario $\omega \in \Omega$ (as defined in Section 2.2.2), and each variable tad_i^ω represents the total arrival delay of flight i if pd_{ij}^ω is positive, i.e., the total arrival delay of flight i propagates to the succeeding flight j.

Proof: Consider a delay scenario $\omega \in \Omega$ and an aircraft route $r = (i_0, i_1, ..., i_n)$ where $i_k \in F$ for $k = 0, ..., n$ and $(i_{k-1}, i_k) \in A$ for $k = 1, ..., n$. We will prove the claim by induction on the index k.

For the base case $k = 0$, we will show that the propagated delay of an aircraft connection (i_0, i_1) is given by pd_{i_0,i_1}^ω, and $tad_{i_0}^\omega$ represents the total arrival delay of flight i_0 if $pd_{i_0,i_1}^\omega > 0$.

From (FR-8) and (FR-10), we have $tad_{i_0}^\omega \geq \text{Max}(IAD_{i_0}^\omega, 0)$. Note that in order to minimize the objective function $\sum_{(i,j)\in A} \left(\sum_{\omega\in\Omega} p_\omega pd_{ij}^\omega\right)$, every pd_{ij}^ω must be minimized. In particular, increasing the value of some pd_{ij}^ω cannot decrease the values of other pd_{ij}^ω, and hence the objective function. As a result, pd_{i_0,i_1}^ω must attain the value of $\text{Max}(tad_{i_0}^\omega - aSlack'_{i_0,i_1}, 0)$ by (FR-6) and (FR-7). Consider the following two cases:

- If $\text{Max}(IAD^{\omega}_{i_0}, 0) > aSlack'_{i_0,i_1}$, then for given x_{i_0} and x_{i_1}, we have

$$\begin{aligned} 0 &< \text{Max}(IAD^{\omega}_{i_0}, 0) - aSlack'_{i_0,i_1} \\ &\leq tad^{\omega}_{i_0} - aSlack'_{i_0,i_1} \\ &= tad^{\omega}_{i_0} - aSlack_{i_0,i_1} + x_{i_0} - x_{i_1} \\ &= pd^{\omega}_{i_0,i_1}. \end{aligned}$$

  In order to minimize $pd^{\omega}_{i_0,i_1} = \text{Max}(IAD^{\omega}_{i_0}, 0) - aSlack'_{i_0,i_1} = \text{Max}(tad^{\omega}_{i_0} - aSlack'_{i_0,i_1}, 0)$, $tad^{\omega}_{i_0}$ must be minimized, and thus we have $tad^{\omega}_{i_0} = \text{Max}(IAD^{\omega}_{i_0}, 0)$, which correctly represents the total arrival delay of flight i_0.

- If $\text{Max}(IAD^{\omega}_{i_0}, 0) \leq aSlack'_{i_0,i_1}$ then for given x_{i_0} and x_{i_1}, we have that any value of $tad^{\omega}_{i_0}$ between $\text{Max}(IAD^{\omega}_{i_0}, 0)$ and $aSlack'_{i_0,i_1} = aSlack_{i_0,i_1} - x_{i_0} + x_{i_1}$ is feasible and yields $pd^{\omega}_{i_0,i_1} = 0 = \text{Max}(tad^{\omega}_{i_0} - aSlack'_{i_0,i_1}, 0)$. In particular, $tad^{\omega}_{i_0}$ does not necessarily represent the total arrival delay of flight i_0.

We established the claim for $k = 0$. Now suppose the claim is true for all indices less than or equal to k. To prove the claim for $tad^{\omega}_{i_{k+1}}$ and $pd^{\omega}_{i_{k+1},i_{k+2}}$, we can use a similar argument together with the induction hypothesis that $pd^{\omega}_{i_k,i_{k+1}}$ correctly models the propagated delay of an aircraft connection (i_k, i_{k+1}). $\square$

Theorem 1: *The polyhedron formed by constraints* (FR-2)-(FR-11) *is integral, given that all data and parameters in those constraints are integral.*

Proof: We first state some characterizations of totally unimodular matrices. Let $\mathbf{A}$ be a matrix with -1, 0, or +1 entries, and $\mathbf{A_j}$ be the j-th column of $\mathbf{A}$. The following statements are equivalent:

i) $\mathbf{A}$ is totally unimodular.

ii) (Hoffman and Kruskal) The polyhedron $\{\mathbf{x} \mid \mathbf{a} \leq \mathbf{Ax} \leq \mathbf{b}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}\}$ is integral for all integral vectors $\mathbf{a}, \mathbf{b}, \mathbf{l}$, and $\mathbf{u}$.

iii) (Ghouila-Houri) Each collection C of columns of $\mathbf{A}$ can be partitioned into two sets, C_1 and C_2 such that $\sum_{j\in C_1} \mathbf{A_j} - \sum_{j\in C_2} \mathbf{A_j}$ is a vector with -1, 0, or +1 entries.

From the equivalence of statements i) and ii), in order to prove the claim, it is sufficient to show that the coefficient matrix $\mathbf{A}$ corresponding to the constraints that are not bound constraints is totally unimodular. We start off by rewriting the coefficient matrix $\mathbf{A}$ in terms of tad_i^ω and x_i.

Consider an aircraft route $r = (i_0, i_1, ..., i_n)$ where $i_k \in F$ for $k = 0, ..., n$ and $(i_{k-1}, i_k) \in A$ for $k = 1, ..., n$. For a given delay scenario $\omega \in \Omega$, the total arrival delay of each flight i_k in the aircraft route r is given by:

$$
\begin{aligned}
tad_{i_k}^\omega &\geq pd_{i_{k-1},i_k}^\omega + IAD_{i_k}^\omega && \text{from (FR-9)}\\
&\geq (tad_{i_{k-1}}^\omega - aSlack'_{i_{k-1},i_k}) + IAD_{i_k}^\omega && \text{from (FR-6)}\\
&= tad_{i_{k-1}}^\omega - (aSlack_{i_{k-1},i_k} - x_{i_{k-1}} + x_{i_k}) + IAD_{i_k}^\omega && \text{from (FR-2)}\\
&\geq (pd_{i_{k-2},i_{k-1}}^\omega + IAD_{i_{k-1}}^\omega) - aSlack_{i_{k-1},i_k} + x_{i_{k-1}} - x_{i_k} + IAD_{i_k}^\omega && \text{from (FR-9)}\\
&\geq (tad_{i_{k-2}}^\omega - aSlack'_{i_{k-2},i_{k-1}}) && \text{from (FR-6)}\\
&\quad - aSlack_{i_{k-1},i_k} + x_{i_{k-1}} - x_{i_k} + \sum_{j=k-1}^{k} IAD_{i_j}^\omega\\
&= tad_{i_{k-2}}^\omega - (aSlack_{i_{k-2},i_{k-1}} - x_{i_{k-2}} + \cancel{x_{i_{k-1}}}) && \text{from (FR-2)}\\
&\quad - aSlack_{i_{k-1},i_k} + \cancel{x_{i_{k-1}}} - x_{i_k} + \sum_{j=k-1}^{k} IAD_{i_j}^\omega\\
&= tad_{i_{k-2}}^\omega + x_{i_{k-2}} - x_{i_k} - \sum_{j=k-1}^{k} aSlack_{i_{j-1},i_j} + \sum_{j=k-1}^{k} IAD_{i_j}^\omega\\
&\ \vdots\\
&\geq tad_{i_0}^\omega + x_{i_0} - x_{i_k} - \sum_{j=1}^{k} aSlack_{i_{j-1},i_j} + \sum_{j=1}^{k} IAD_{i_j}^\omega\\
&\geq x_{i_0} - x_{i_k} - \sum_{j=1}^{k} aSlack_{i_{j-1},i_j} + \sum_{j=0}^{k} IAD_{i_j}^\omega && \text{from (FR-8)}
\end{aligned}
$$

Thus, for each flight $i_k \in F$ and $\omega \in \Omega$, we have

$$tad^{\omega}_{i_k} - x_{i_0} + x_{i_k} \geq \sum_{j=0}^{k} IAD^{\omega}_{i_j} - \sum_{j=1}^{k} aSlack_{i_{j-1},i_j} \tag{FR-13}$$

where i_0 is the first flight in the aircraft route containing flight i_k.

In the derivation above, we have made use of all decision variables but $pSlack'_{ij}$. From (FR-4) and (FR-5), we have

$$x_j - x_i \geq -pSlack_{ij} \qquad \forall (i,j) \in P \tag{FR-14}$$

Now we will use Ghouila-Houri's characterization to show that the coefficient matrix $\mathbf{A}$ corresponding to the constraints (FR-13) and (FR-14) is totally unimodular. For **any** collection C of columns of $\mathbf{A}$, let C_1 be a set of the columns in C associated with the decision variables tad^{ω}_i, and C_2 be a set of columns in C associated with the decision variables x_i.

According to (FR-13) and (FR-14), all entries a_{ij} in the coefficient matrix $\mathbf{A}$ are -1, 0, or +1. For each row i, the sum of the coefficients in C_1 ($\sum_{j \in C_1} a_{ij}$) is either 0 or +1, and the sum of the coefficients in C_2 ($\sum_{j \in C_2} a_{ij}$) is -1, 0, or +1, depending on the collection C. If there exists a row i such that $\sum_{j \in C_1} a_{ij} = 1$ and $\sum_{j \in C_2} a_{ij} = -1$, then $\sum_{j \in C_1} a_{ij} - \sum_{j \in C_2} a_{ij} = 2 \notin \{-1, 0, +1\}$. We, however, can modify the partitions by moving the column $j \in C_1$ contributing 1 in $\sum_{j \in C_1} a_{ij}$ to set C_2. The resulting partitions C'_1 and C'_2 yield $\sum_{j \in C'_1} a_{ij} - \sum_{j \in C'_2} a_{ij} = 0$. Note that because the variable tad^{ω}_{ij} appears in exactly one row, the modification only affects row i. We can repeat this process and obtain partitions C^*_1 and C^*_2 such that $| \sum_{j \in C^*_1} \mathbf{A_j} - \sum_{j \in C^*_2} \mathbf{A_j} | \leq 1$. Because this is true for any collection C of columns of $\mathbf{A}$, the coefficient matrix $\mathbf{A}$ is totally unimodular by Ghouila-Houri's characterization.

Because the constraints (FR-13) and (FR-14) are equivalent to the constraints (FR-2)-(FR-11), we established that, for all integral data ($aSlack_{ij}, pSlack_{ij}, IAD^{\omega}_i$) and parameters (l_i, u_i), the polyhedron formed by the constraints (FR-2)-(FR-11) is integral by Hoffman and Kruskal's Theorem. □

As a result, we can relax the integrality constraint (FR-12) and solve the FR problem as a linear optimization problem, instead of an integer optimization problem. This allows us to solve this problem more efficiently, especially when the size of the network is very large.

This model can be considered a variation of the flight schedule re-timing model proposed by AhmadBeygi, Cohn, and Lapp (2008) [2]. In particular, the decision variables – the changes in the departure times – are modeled in a similar manner, but the calculations of the total propagated delay are different.

Our model considers only the delay propagation due to aircraft arriving late, whereas their model considers the delay propagation due to aircraft and cockpit crews. As a result, our calculation of the total propagated delay can be accomplished simply through aircraft routes, using the notion of propagated delay introduced in Section 2.2.2. AhmadBeygi, Cohn, and Lapp, on the other hand, propose the notion of a *propagation tree* in order to capture delay propagation from multiple resources. As mentioned in their paper, their model using propagation trees does not accurately take into account simultaneous delays from different propagation trees. Our model, however, can capture this correctly considering only delay propagation due to aircraft arriving late.

Another difference is that we allow independent arrival delays to take negative values to reflect overestimated block times of some flights or overestimated minimum turn times of some aircraft connections according to the historical operating data. Lastly, we also enforce the feasibility of every existing passenger itinerary.

3.2.3 Alternative Objective Functions

In Section 3.1.4, we introduce the objective function of maximizing the total expected *effective slack*, which can potentially lead to more robust solutions. We can modify the formulation in the previous section to incorporate this idea as well.

Maximizing the total expected effective aircraft connection slack

Let $\overline{aSlack}_{ij}^{\omega}$ be the effective aircraft connection slack associated with a connection $(i,j) \in A$ capped at a nonnegative level Γ_{ij} for a given delay scenario $\omega \in \Omega$. Because the resulting planned aircraft connection slack of the aircraft connection (i,j) after re-timing is given by $aSlack'_{ij}$, we have

$$\overline{aSlack}_{ij}^{\omega} = \mathrm{Min}(aSlack'_{ij} - tad_i^{\omega}, \Gamma_{ij}). \tag{FR-15}$$

Therefore, the flight schedule re-timing model maximizing the total expected effective aircraft connection slack is given by:

$$\text{Maximize} \qquad \sum_{(i,j)\in A} \left(\sum_{\omega\in\Omega} p_{\omega} \overline{aSlack}_{ij}^{\omega} \right) \tag{FR-16}$$

$$\text{subject to} \qquad \overline{aSlack}_{ij}^{\omega} \leq aSlack'_{ij} - tad_i^{\omega} \qquad \forall (i,j) \in A \tag{FR-17}$$

$$\overline{aSlack}_{ij}^{\omega} \leq \Gamma_{ij} \qquad \forall (i,j) \in A \tag{FR-18}$$

(FR-2) – (FR-12)

Maximizing the total expected effective passenger connection slack

As discussed in Section 2.2.3, misconnecting passengers typically represent a significant portion of total passenger delay. One possible way to reduce the likelihood of passenger misconnection is to provide sufficient slack in passenger connections. Unlike the string-based formulation for the `AR` problem, the formulation for the `FR` problem presented in the previous section can also capture the information regarding passenger connections. Therefore, we can apply the notion of *effective slack* to passenger connections as well.

In particular, $\overline{pSlack}_{ij}^{\omega}$, the effective passenger connection slack associated with a connection $(i,j) \in P$ capped at a nonnegative level Γ_{ij} for a given delay scenario

$\omega \in \Omega$, is defined as

$$\overline{pSlack}_{ij}^{\omega} = \text{Min}(pSlack'_{ij} - tad_i^{\omega}, \Gamma_{ij}). \tag{FR-19}$$

Thus, the flight schedule re-timing model maximizing the total expected effective passenger connection slack is given by:

$$\text{Maximize} \qquad \sum_{(i,j)\in P} \left(\sum_{\omega\in\Omega} p_{\omega} \overline{pSlack}_{ij}^{\omega} \right) \tag{FR-20}$$

$$\text{subject to} \qquad \overline{pSlack}_{ij}^{\omega} \leq pSlack'_{ij} - tad_i^{\omega} \qquad \forall (i,j) \in P \tag{FR-21}$$

$$\overline{pSlack}_{ij}^{\omega} \leq \Gamma_{ij} \qquad \forall (i,j) \in P \tag{FR-22}$$

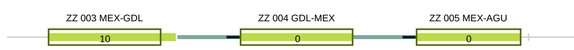
$$\text{(FR-2)} - \text{(FR-12)}$$

3.2.4 Multiple Optimal Solutions

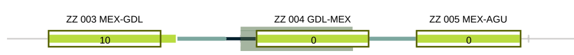
The issue of multiple optimal solutions also arises in the flight schedule re-timing problem, regardless of the objective function. Similar to the robust aircraft re-routing problem, the number of optimal solutions depends typically on 1) re-timing opportunities in the airline schedule which are determined by allowable time windows for re-timing, the number of passenger connections, planned aircraft connection times in the given aircraft routes, etc.; and 2) the extent of delays in the given set of historical data.

The example in Figure 3-5 illustrates multiple optimal solutions in the robust flight schedule re-timing problem. The number associated with each flight indicates the independent delay. The first feasible solution represents the original schedule. The departure time of flight ZZ 004 is moved later in the second feasible solution and moved earlier in the third feasible solution. Suppose the objective function is to minimize the total expected propagated delay, then all three feasible solutions result in the same objective function value of zero, i.e., no delay propagation. These feasible solutions, however, may not be equal with respect to other performance metrics or objective functions. For example,

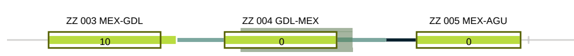
- The feasible solution I might be preferable because flight schedule consistency is maintained from one planning period to the next.
- The feasible solution II might be preferable because more ground slack is allocated to the aircraft connection between ZZ 003 and ZZ 004 for which delay propagation is more likely to occur.
- The feasible solution III might be preferable because most of the passengers on flight ZZ 004 might have to make connections to other flights, and moving the departure time of flight ZZ 004 earlier could potentially reduce the likelihood of passenger misconnections. Note that this is usually the case for a flight operating from a spoke into a hub.



(a) Feasible solution I



(b) Feasible solution II



(c) Feasible solution III

Figure 3-5: Multiple optimal solutions in the robust flight schedule re-timing problem

As discussed in Section 3.1.5, we can solve the flight schedule re-timing problem sequentially using different objective functions in order to select, among the optimal solutions to the initial objective, a solution that is optimal with respect to other objectives.

One advantage of the `FR` formulation over the `AR` formulation is its ability to capture information regarding passenger connections. Thus, one possible sequence of objective functions is to maximize the total expected effective aircraft connection slack followed by maximizing the total expected effective passenger connection slack, or the opposite order, if our primary goal is to improve passenger delays.

Another possible objective is to minimize the difference between the original and the optimal flight schedule because an airline typically prefers a degree of consistency across planning periods. One way to achieve this objective is to simply minimize the total change in the departure times, which is given by $\sum_{i \in F} |x_i|$. Some flights, however, might be affected more by re-timing, e.g., short-haul high-frequency flights, flights in business market, and flights during peak hours. We can specify different weights w_i for each flight $i \in F$ to reflect its importance, and minimize $\sum_{i \in F} w_i |x_i|$ instead. Additionally, we can introduce a piecewise linear convex function $w_i(x_i)$ that gives a larger penalty to a larger change. This enables the model to distinguish between two flights that are moved by 0 and 15 minutes and two flights that are moved by 7 and 8 minutes.

3.3 Robust Block Time Adjustment Model

3.3.1 Underlying Idea

Among the three slack re-allocation schemes introduced in Section 2.4, block time adjustment provides the greatest flexibility as it affects not only aircraft connection slack and passenger connection slack but also block time slack. Block time slack can absorb independent delay such as taxi delay and airborne delay, which increases with increased air traffic. To improve on-time performance, many airlines *pad* their schedules by increasing their block times to account for potential delays. The usefulness of block time slack, however, comes at a cost. Longer block times can lead to longer crew duties and less productivity of aircraft.

Given a set of historical data, we want to adjust flight block times optimally such that the resulting schedule minimizes some proxy of expected delays or disruptions, as demonstrated in Section 2.4.4. In particular, we want to re-allocate the existing slack in the schedule to where it is most needed historically.

In the next section, we present the robust block time adjustment model to minimize the total expected arrival delay. Then, we propose alternative objective functions in subsequent sections.

3.3.2 Formulation

The formulation we present in this section is an extension of the robust flight schedule re-timing model introduced in Section 3.2.2. In particular, we add another decision variable to capture the change in the arrival time of each flight, rather than assuming it equals the change in the departure time as in the `FR` formulation. This enables us to move the departure and arrival times of each flight independently. As discussed earlier, an increase in block times can help absorb independent delays; while a decrease in block times can result in larger arrival delays. Therefore, the total arrival delay of each flight changes with changes in block time.

For each flight $i \in F$, let x_i be the difference between the new and the original departure time, and y_i be the difference between the new and the original arrival time. The variable $x_i(y_i)$ takes a negative value if the departure(arrival) time is moved earlier and takes a positive value if the departure(arrival) time is moved later. We limit the change in the departure time of each flight i to a small time window $[l_{x_i}, u_{x_i}]$, and the arrival time to $[l_{y_i}, u_{y_i}]$. Note that in our block time adjustment model, we allow for block time reduction of some flights as well.

Using the notations from Section 3.2.2, the robust block time adjustment problem (`BA`) is given by:

$$\text{Minimize} \qquad \sum_{i \in F} \mathbb{E}\left[tad_i\right] = \sum_{i \in F} \left(\sum_{\omega \in \Omega} p_\omega tad_i^\omega \right) \tag{BA-1}$$

$$\text{subject to} \qquad aSlack'_{ij} = aSlack_{ij} - y_i + x_j \qquad \forall (i,j) \in A \tag{BA-2}$$

$$aSlack'_{ij} \geq 0 \qquad \forall (i,j) \in A \tag{BA-3}$$

$$pSlack'_{ij} = pSlack_{ij} - y_i + x_j \qquad \forall (i,j) \in P \tag{BA-4}$$

$$pSlack'_{ij} \geq 0 \qquad \forall (i,j) \in P \quad \text{(BA-5)}$$

$$pd^{\omega}_{ij} \geq tad^{\omega}_{i} - aSlack'_{ij} \qquad \forall (i,j) \in A, \forall \omega \in \Omega \quad \text{(BA-6)}$$

$$pd^{\omega}_{ij} \geq 0 \qquad \forall (i,j) \in A, \forall \omega \in \Omega \quad \text{(BA-7)}$$

$$tad^{\omega}_{i} \geq IAD^{\omega}_{i} + x_i - y_i \qquad \forall i \in F_0, \forall \omega \in \Omega \quad \text{(BA-8)}$$

$$tad^{\omega}_{j} \geq pd^{\omega}_{ij} + IAD^{\omega}_{j} + x_j - y_j \qquad \forall (i,j) \in A, \forall \omega \in \Omega \quad \text{(BA-9)}$$

$$tad^{\omega}_{i} \geq 0 \qquad \forall i \in F, \forall \omega \in \Omega \text{ (BA-10)}$$

$$l_i \leq y_i - x_i \leq u_i \qquad \forall i \in F \text{ (BA-11)}$$

$$l_{x_i} \leq x_i \leq u_{x_i} \qquad \forall i \in F \text{ (BA-12)}$$

$$l_{y_i} \leq y_i \leq u_{y_i} \qquad \forall i \in F \text{ (BA-13)}$$

$$x_i, y_i \in \mathbb{Z}^n \qquad \forall i \in F \text{ (BA-14)}$$

The objective function (BA-1) is to minimize the total expected arrival delay over all flights. Again, we assume Ω has finite cardinality.

Constraint set (BA-2) is the resulting planned slack of each aircraft connection $(i,j) \in A$ after moving the departure and arrival times of flights i and j. The resulting planned slack ($aSlack'_{ij}$) increases when the arrival time of flight i is moved earlier, i.e., when y_i takes a negative value. Also, it increases when the departure time of flight j is moved later, i.e., when x_j takes a positive value. The non-negativity constraints of the resulting planned aircraft connection slack, (BA-3), ensure that every aircraft connection is longer than its corresponding minimum aircraft turn time, and thus the current aircraft routing remains feasible.

Similarly, constraint set (BA-4) is the resulting planned passenger connection slack after re-timing for each passenger connection. The non-negativity constraints of the resulting planned passenger connection slack, (BA-5), ensure that every passenger connection is longer than its corresponding minimum passenger connection time, and thus every itinerary remains feasible.

Given a re-timed flight schedule, constraints (BA-6) and (BA-7) are the propagated delays for each aircraft connection under different delay scenarios; constraints (BA-8)-(BA-10) determine the total arrival delays for each flight leg under different delay scenarios, assuming that the first flight of each string has zero propagated delay. Note that one can think of the term $IAD^{\omega}_{j} + x_j - y_j$ in (BA-8) and (BA-9) as the resulting independent arrival delay of flight j after changing the block time.

Constraints (BA-11) restrict the total change in a block time of each flight within the range $[l_i, u_i]$. Lastly, constraints (BA-12) and (BA-13) limit the change in the departure and arrival times of each flight i within specific time windows $[l_{x_i}, u_{x_i}]$ and $[l_{y_i}, u_{y_i}]$.

This formulation is an extension of the robust flight schedule re-timing formulation introduced in Section 3.2.2, and it inherits the property that the polyhedron formed by the constraints in the formulation is integral, given integral data and parameters.

Theorem 2: *The polyhedron formed by constraints* (BA-2)-(BA-13) *is integral, given that all data and parameters in those constraints are integral.*

Proof: This proof is similar to the one presented in Section 3.2.2, except that we have an additional set of decision variables y_i and constraints (BA-11).

In order to prove the claim, it is sufficient to show that the coefficient matrix $\mathbf{A}$ corresponding to the constraints that are not bound constraints is totally unimodular. We start off by rewriting the coefficient matrix $\mathbf{A}$ in terms of tad^{ω}_{i},x_i, and y_i.

Consider an aircraft route $r = (i_0, i_1, ..., i_n)$ where $i_k \in F$ for $k = 0, ..., n$ and $(i_{k-1}, i_k) \in A$ for $k = 1, ..., n$. For a given delay scenario $\omega \in \Omega$, the total arrival delay of each flight i_k in the aircraft route r is given by:

$$
\begin{aligned}
tad^{\omega}_{i_k} &\geq pd^{\omega}_{i_{k-1},i_k} + IAD^{\omega}_{i_k} + x_{i_k} - y_{i_k} && \text{from (BA-9)}\\
&\geq (tad^{\omega}_{i_{k-1}} - aSlack'_{i_{k-1},i_k}) + IAD^{\omega}_{i_k} + x_{i_k} - y_{i_k} && \text{from (BA-6)}\\
&= tad^{\omega}_{i_{k-1}} - (aSlack_{i_{k-1},i_k} - y_{i_{k-1}} + \cancel{x_{i_k}}) + IAD^{\omega}_{i_k} + \cancel{x_{i_k}} - y_{i_k} && \text{from (BA-2)}\\
&\geq (pd^{\omega}_{i_{k-2},i_{k-1}} + IAD^{\omega}_{i_{k-1}} + x_{i_{k-1}} - \cancel{y_{i_{k-1}}}) && \text{from (BA-9)}\\
&\quad - aSlack_{i_{k-1},i_k} + \cancel{y_{i_{k-1}}} - y_{i_k} + IAD^{\omega}_{i_k}
\end{aligned}
$$

$$
\begin{aligned}
&\geq (tad^{\omega}_{i_{k-2}} - aSlack'_{i_{k-2},i_{k-1}}) && \text{from (BA-6)}\\
&\quad - aSlack_{i_{k-1},i_k} + x_{i_{k-1}} - y_{i_k} + \sum_{j=k-1}^{k} IAD^{\omega}_{i_j} \\
&= tad^{\omega}_{i_{k-2}} - (aSlack_{i_{k-2},i_{k-1}} - y_{i_{k-2}} + \cancel{x_{i_{k-1}}}) && \text{from (BA-2)}\\
&\quad - aSlack_{i_{k-1},i_k} + \cancel{x_{i_{k-1}}} - y_{i_k} + \sum_{j=k-1}^{k} IAD^{\omega}_{i_j} \\
&= tad^{\omega}_{i_{k-2}} + y_{i_{k-2}} - y_{i_k} - \sum_{j=k-1}^{k} aSlack_{i_{j-1},i_j} + \sum_{j=k-1}^{k} IAD^{\omega}_{i_j} \\
&\ \vdots \\
&\geq tad^{\omega}_{i_0} + y_{i_0} - y_{i_k} - \sum_{j=1}^{k} aSlack_{i_{j-1},i_j} + \sum_{j=1}^{k} IAD^{\omega}_{i_j} \\
&\geq (IAD^{\omega}_{i_0} + x_{i_0} - \cancel{y_{i_0}}) + \cancel{y_{i_0}} - y_{i_k} - \sum_{j=1}^{k} aSlack_{i_{j-1},i_j} + \sum_{j=1}^{k} IAD^{\omega}_{i_j} && \text{from (BA-8)}\\
&= x_{i_0} - y_{i_k} - \sum_{j=1}^{k} aSlack_{i_{j-1},i_j} + \sum_{j=0}^{k} IAD^{\omega}_{i_j}
\end{aligned}
$$

Thus, for all flight $i_k \in F$ and $\omega \in \Omega$, we have

$$
tad^{\omega}_{i_k} - x_{i_0} + y_{i_k} \geq \sum_{j=0}^{k} IAD^{\omega}_{i_j} - \sum_{j=1}^{k} aSlack_{i_{j-1},i_j} \tag{BA-15}
$$

where i_0 is the first flight in the aircraft route containing flight i_k.

Analogously to the inequalities (FR-14) in the proof for `FR` formulation, we have

$$
x_j - x_i \geq -pSlack_{ij} \qquad \forall (i,j) \in P \tag{BA-16}
$$

from (BA-4) and (BA-5).

We also have additional constraints that have not been considered:

$$
l_i \leq y_i - x_i \leq u_i \qquad \forall i \in F \tag{BA-17}
$$

For **any** collection C of columns of $\mathbf{A}$, let C_1 be a set of the columns in C associated with the decision variables tad_i^ω, and C_2 be a set of columns in C associated with the decision variables x_i and y_i.

According to (BA-15)-(BA-17), all entries a_{ij} in the coefficient matrix $\mathbf{A}$ are -1, 0, or +1. For each row i, the sum of the coefficients in C_1 ($\sum_{j \in C_1} a_{ij}$) is either 0 or +1, and the sum of the coefficients in C_2 ($\sum_{j \in C_2} a_{ij}$) is -1, 0, or +1, depending on the collection C. If there exists a row i such that $\sum_{j \in C_1} a_{ij} = 1$ and $\sum_{j \in C_2} a_{ij} = -1$, then $\sum_{j \in C_1} a_{ij} - \sum_{j \in C_2} a_{ij} = 2 \notin \{-1, 0, +1\}$. We, however, can modify the partitions by moving the column $j \in C_1$ contributing 1 in $\sum_{j \in C_1} a_{ij}$ to set C_2. The resulting partitions C_1' and C_2' yield $\sum_{j \in C_1'} a_{ij} - \sum_{j \in C_2'} a_{ij} = 0$. Note that because the variable tad_{ij}^ω appears in exactly one row, the modification only affects row i. We can repeat this process and obtain partitions C_1^* and C_2^* such that $| \sum_{j \in C_1^*} \mathbf{A_j} - \sum_{j \in C_2^*} \mathbf{A_j} | \leq 1$. Because this is true for any collection C of columns of $\mathbf{A}$, the coefficient matrix $\mathbf{A}$ is totally unimodular by Ghouila-Houri's characterization.

Because the constraints (BA-15)-(BA-17) are equivalent to the constraints (BA-2)-(BA-13), by Hoffman and Kruskal's Theorem, we establish that, for all integral data ($aSlack_{ij}$, $pSlack_{ij}$, IAD_i^ω) and parameters ($l_{x_i}, u_{x_i}, l_{y_i}, u_{y_i}$), the polyhedron formed by the constraints (BA-2)-(BA-13) is integral. $\square$

Therefore, we can relax the integrality constraint (BA-14) and solve the `BA` problem as a linear optimization problem, instead of an integer optimization problem.

3.3.3 Alternative Objective Functions

Minimizing the total expected propagated delay

At this point, one question that may arise is why the objective function of the `BA` model presented in the previous section is to minimize the total expected arrival delay, rather than minimizing the total expected propagated delay. Unlike in the `AR` or `FR` models, minimizing the total expected propagated delay in the `BA` formulation is in fact not a good proxy for minimizing the total expected arrival delay.

More specifically, recall that a total arrival delay for flight j is given by $tad_j =$

$\text{Max}(pd_{ij} + IAD_j, 0)$ (see (2.7)), where flight i is immediately preceding flight j in the aircraft route. In the AR and FR models, because independent arrival delays are fixed, decreasing a propagated delay from flight i to flight j, pd_{ij} will result in a smaller total arrival delay for flight j as long as the term $pd_{ij} + IAD_j$ remains positive. This, however, is not the case for the BA model. In particular, an independent arrival delay of flight j is now a function of decision variables x_j and y_j and equal to $IAD_j + x_j - y_j$, instead of a constant IAD_j. Thus, one can possibly decrease pd_{ij} and simultaneously increase the resulting independent arrival delay, $IAD_j + x_j - y_j$ as well as the total arrival delay tad_j. For instance, consider a flight j that experiences a nonzero propagated delay from flight i. We can simply avoid this delay propagation by moving the departure time of flight j to when the aircraft is ready. Without changing the arrival time of flight j, the total arrival delay will remain the same, in spite of the reduction in propagated delay. In fact, if we also move the arrival time of flight j earlier, the total arrival delay will increase.

Moreover, when there exist multiple optimal solutions, the objective function of minimizing the total expected propagated delay will be indifferent to the optimal solution that also minimizes the total expected arrival delay. From (BA-6) and (BA-2), we have that $pd_{ij} \geq tad_i - aSlack_{ij} + y_i - x_j$. Suppose $i \notin F_0$. If $tad_i > 0$, it follows from (BA-9) that

$$pd_{ij} \geq (pd_{i-1,i} + IAD_i + x_i - \cancel{y_i}) - aSlack_{ij} + \cancel{y_i} - x_j,$$

where flight $i - 1$ is immediately preceding flight i in the aircraft route. In other words, the scheduled arrival time of flight i, y_i, does not affect the propagated delay from flight i to j, and thus we can set it to any allowable value, as long as the total arrival delay of flight i remains positive (see Figure 3-6). As a result, given x_i and x_j that minimize pd_{ij}, the objective function of minimizing the total expected propagated delay has no incentive to set y_i such that it minimizes tad_i.

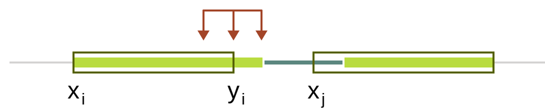


Figure 3-6: Minimizing propagated delay in the robust block time adjustment model

Maximizing the total expected effective aircraft connection slack

We can argue similarly that maximizing the total expected effective aircraft connection slack is also not a good proxy for minimizing the total expected arrival delay. In particular, it is possible that the total arrival delay remains the same, in spite of the increase in effective slack. Additionally, the scheduled arrival time of flight i, y_i does not affect the effective slack of the aircraft connection from flight i to j, and we can set it to any allowable value, as long as the total arrival delay of flight i remains positive. Therefore, given x_i and x_j that maximize the effective slack in the aircraft connection from flight i to j, the objective function of maximizing the total expected effective aircraft connection slack has no incentive to set y_i such that it minimizes tad_i.

Maximizing the total expected effective passenger connection slack

Because the `BA` formulation is capable of modifying passenger connection slack, we can apply the notion of effective passenger connection slack to the `BA` model as well. This can be done in the exact same way as we did for the robust flight schedule re-timing model.

3.3.4 Multiple Optimal Solutions

The issue of multiple optimal solutions, again, arises in the robust block time adjustment formulation, introduced in Section 3.3.2. Consider two consecutive flights i and j in the same aircraft route. Suppose $i, j \notin F_0$ and $tad_j > 0$. Then, from (BA-9), we have that $tad_j \geq pd_{ij} + IAD_j + x_j - y_j$. If $pd_{ij} > 0$, it follows that $tad_i > 0$ as well, and thus we have

$$
\begin{aligned}
tad_j &\geq pd_{ij} + IAD_j + x_j - y_j \\
&\geq (tad_i - aSlack_{ij} + y_i - \cancel{x_j}) + IAD_j + \cancel{x_j} - y_j && \text{(From (BA-6))} \\
&\geq (pd_{i-1,i} + IAD_i + x_i - \cancel{y_i}) - aSlack_{ij} + \cancel{y_i} + IAD_j - y_j && \text{(From (BA-9))}
\end{aligned}
$$

where flight $i-1$ is immediately preceding flight i in the aircraft route. In other words, the scheduled departure time of flight j, x_j does not affect the total arrival delay of flight j, and hence we can set it to any allowable value, as long as $tad_j > 0$ and $pd_{ij} > 0$ (see Figure 3-7). Therefore, given x_i and y_j that minimizes tad_j, the objective function of minimizing the total expected arrival delay has no incentive to set x_j such that the departure delay of flight j is minimized, or the slack in passenger connections to flight j is maximized.

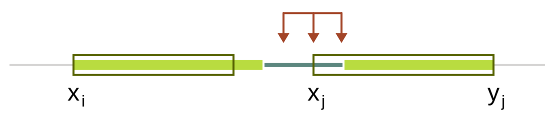


Figure 3-7: Multiple optimal solutions in the robust block time adjustment problem

As we did in the previous models, we can solve the robust block time adjustment sequentially using different objective functions in order to select, among the optimal solutions to the initial objective, a solution that is optimal with respect to other objectives. Some possible secondary objectives are

- minimizing the total expected propagated delay;
- maximizing the total expected effective aircraft connection slack;
- maximizing the total expected effective passenger connection slack; and
- minimizing the difference between the original and the optimal flight schedule.

For the first three objectives, this can be done in a way similar to that which we used in the `FR` models. In order to measure the difference between the original and the optimal flight schedule, we calculate the *total absolute block time change*, defined as $\sum_{i \in F} |y_i - x_i|$. Note that this metric is intended to measure the *difference* and does not distinguish block time reduction from block time increase.

Chapter 4

Proof of Concepts

4.1 Data and Evaluation process

We obtained three months of historical operations data (from January 1st to March 25th 2008) from an international carrier (denoted as "Airline A" in this thesis). The dataset contains

- flight information: origin, destination, planned departure and arrival times, actual departure and arrival times, flight status (diverted, canceled, arrived), and aircraft type. (Note that we do not have information about taxi-out and taxi-in times.);
- planned and actual aircraft routings;
- minimum turn times currently used by Airline A; and
- number of passengers booked on each itinerary (for March data only).

Because airlines use historical data to build schedules for future operations, we divide our dataset into two disjoint subsets representing historical data and future operations.

January and February data are used as historical data. In particular, we consider each day of operation in January and February as one instance of delay scenario ω. Hence, the set of delay scenarios Ω has cardinality $31 + 29 = 60$. We also assume that each disruption scenario is equally likely. To obtain our planned schedule for March, we solve the proposed robust slack re-allocation models over each day of operation. The size of a problem on each day of operation is summarized in Table 4.1. Note that for the robust aircraft re-routing problem, we can decompose the problem further into subproblems for different aircraft types.

Next, we use the actual delay information from March data to evaluate performance of the robust schedules. Specifically, we apply the formulas given in Section 2.2.2 to compute the actual independent delay of each flight for each day of operation in March. Given the actual independent delay of each flight, we simulate the actual departure and arrival times of each flight with respect to the new schedules, assuming no flight cancellations and aircraft swaps during the day of operation. Finally, we compute the performance evaluation statistics, presented in Section 2.2. The data flow and evaluation process are depicted in Figure 4-1.

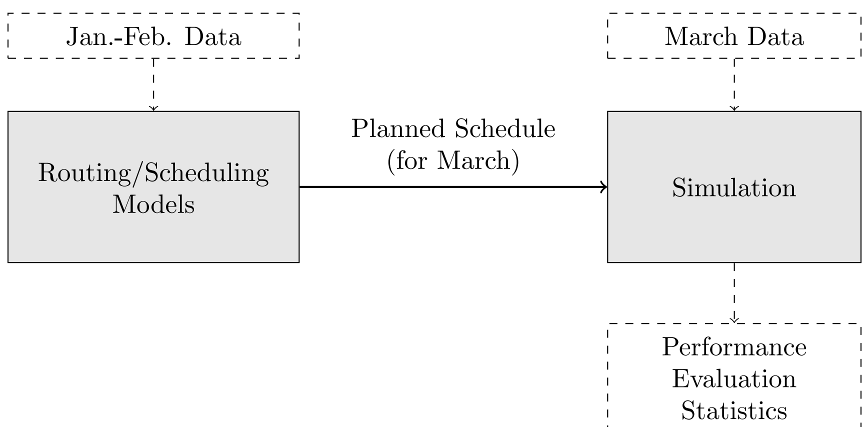


Figure 4-1: Data Flow and Evaluation Process

4.1.1 Passenger Delay Calculation

As discussed in Section 2.2, understanding the extent of passenger delays is crucial in evaluating schedule performance. In contrast to other flight delay metrics, passenger delays cannot be obtained directly from actual flight operation data because we need to consider re-accommodation of *disrupted passengers*– those who missed their

Day of Operation (March 2008)	Number of Flights	Number of Aircraft	Number of Passenger Connections[a]	Number of Passengers
1	231	59	469	12840
2	243	61	544	16027
3	250	58	520	13889
4	236	57	449	12455
5	240	59	514	14187
6	260	59	560	15288
7	263	61	595	17416
8	242	61	506	14620
9	243	60	566	16485
10	246	59	464	14634
11	226	55	299	8725
12	232	57	461	16258
13	268	61	561	19891
14	260	62	606	22906
15	242	61	470	19871
16	246	61	523	18040
17	247	61	511	15192
18	229	57	409	12918
19	230	59	496	14718
20	231	60	500	13993
21	211	58	447	11409
22	249	62	430	16192
23	244	61	557	20648
24	252	61	583	19226
25	229	56	453	15201

Table 4.1: Problem sizes

[a] The number of passenger connections indicates the number of *distinct* passenger connections in the passenger booking data. Note, this is different from the number of distinct itineraries.

connections or whose flights were canceled. In this work, we calculate passenger delays using the Passenger Delay Calculator algorithm (PDC) developed by Barnhart and Bratu (2005) [10].

The inputs to PDC consist of a flight schedule, actual flight operation data (i.e., actual departure/arrival times and flight cancellations), and passenger booking data. The algorithm first determines itineraries that are *disrupted* according to the actual flight operation data. Recall that an itinerary is *disrupted* if one or more flights in the itinerary are canceled, or some connecting time between consecutive flights becomes less than the minimum connecting time required. Non-disrupted passengers are assigned to their original itinerary, whereas disrupted passengers are inserted into the *recovery queue*, which is processed in some specific order. Next, the PDC algorithm finds each disrupted passenger the best available recovery itinerary– the one that arrives earliest at his or her final destination. Note that a recovery itinerary is available only if it is feasible, that is, every connection satisfies the minimum connection time requirement, and every flight in the itinerary has at least one unassigned seat.

The passenger delay of a non-disrupted passenger is simply given by the arrival delay of the last flight in his or her itinerary. For each disrupted passenger, the passenger delay is the difference between the planned arrival time of the last flight in his or her original itinerary and the actual arrival time of the last flight in his or her recovery itinerary. Finally, if the passenger delay associated with a disrupted passenger exceeds the *maximum passenger delay* threshold, it is assumed that the passenger is re-accommodated on another airline. Because flight schedules of other airlines as well as seat availability information are not available to PDC, we cannot accurately compute passenger delays for disrupted passengers that are re-accommodated on other airlines. Instead, we assign delay equal to the maximum passenger delay threshold for each disrupted passenger.

In this work, we re-accommodate disrupted passengers on a *first-disrupted-first-recovered* basis. In particular, the recovery queue is sorted in increasing order of *disruption times*. The *disruption time* associated with each itinerary is determined as follows: 1) If an itinerary is disrupted due to flight cancellation, its disruption time is given by the planned departure time of the flight; 2) If an itinerary is disrupted due to misconnection, its disruption time is given by the actual arrival time of the last flight before the missed connection. Moreover, we assume that the minimum connection time for every connection is 30 minutes, and the maximum passenger delay is 12 hours. For simplicity, we also consider only itineraries with at most two flight legs, that is, at most one connection. According to the dataset, such itineraries represent an average of almost 99 percent of the passenger itineraries.

Note that the PDC algorithm only provides approximations of passenger delays. We summarize here the assumptions of the PDC algorithm that may limit the accuracy of these approximations:

- Because the actual flight information of every flight is known to PDC, disrupted passengers are re-accommodated as if perfect information on future operations is known. In reality, a disrupted passenger might be re-booked on a flight that is later canceled, and thus additional re-booking is necessary. As a result, actual passenger delay might be much larger than the passenger delay obtained from PDC.

- PDC always gives priority to non-disrupted passengers. Specifically, if a flight is already full, it cannot be used in a recover itinerary. However, it might be preferable to ask a volunteer to give up his or her seat to a disrupted passenger who is severely delayed and/or needs to connect to another flight.

- The disruption time of each disrupted passenger can be different from our definition. For example, a misconnecting passenger might be aware of his itinerary disruption before his first flight departed, and he could have been re-accommodated with a better itinerary before he began his trip. As a result, PDC may overestimate passenger delays. Also, a passenger might be aware of his flight cancellation after the planned departure time of the flight, for example, after waiting an hour for an aircraft repair attempt. In this case, PDC can underestimate passenger delays.

	Original	AR_minPD		AR_maxEffACSlack15	
Schedule Statistics					
Total A/C Connection Slack (mins)	6676.76	7459.16	+11.72%	7494.36	+12.25%
% of Original A/C Connections Included	-	67.92%		68.60%	
Flight Delay Statistics					
Total Propagated Delay (mins)	1009.60	818.60	-18.92%	781.40	-22.60%
% of Flights with PD > 0	17.74%	14.86%		14.41%	
Total Arrival Delay (mins)	3141.16	2965.56	-5.59%	2929.16	-6.75%
15-min On-Time Performance	76.53%	77.82%		78.02%	
60-min On-Time Performance	96.89%	97.26%		97.36%	
Passenger Delay Statistics					
Total Pax Delay (mins)	260565	250325	-3.93%	246903	-5.24%
Total Disrupted Pax (pax)	47.56	45.16	-5.05%	44.72	-5.97%

Table 4.2: Average performance evaluation statistics over 25 days (March 1-25, 2008) for the `AR` models

For details of the PDC algorithm and discussion on the validity of the assumptions underlying it, readers are referred to Bratu (2003) [9] and Bratu and Barnhart (2005) [10].

4.2 Robust Aircraft Re-routing Model

In this section, we present the results obtained from the robust aircraft re-routing models introduced in Section 3.1, with different objective functions and parameters. Let `Original` denote the airline's original schedule, and define the solutions as follows.

`AR_minPD` the solution to the `AR` model that minimizes the total expected propagated delay (see (AR-6))

`AR_maxEffACSlack<Γ>` the solution to the `AR` model that maximizes the total expected effective aircraft connection slack with caps set equal to Γ minutes for every aircraft connection (see (AR-9))

4.2.1 Computational Results

Table 4.2 summarizes average statistics of `Original`, `AR_minPD`, and `AR_maxEff-ACSlack15` solutions over the period of March 1-25, 2008. To obtain these statistics, we evaluate the performance of the solutions over each day of operation in March using the approach discussed in Section 4.1, and compute the average statistics over 25 days. We divide the statistics into three categories: 1) schedule statistics, 2) flight delay statistics, and 3) passenger delay statistics.

Schedule Statistics

According to Table 4.2, both `AR_minPD` and `AR_maxEffACSlack15` solutions have more total slack than the original schedule. There are two factors contributing to the increase:

1) In the `AR` model, we assume that every aircraft is ready at the beginning of the day of operation and also available until the end of the day. Therefore, it is possible that some aircraft have longer elapsed times between the start and end of flying in a given day, as compared to the original schedule. Figure 4-2 exemplifies the situation. Flights in the original schedule that are operated by a single aircraft are drawn along a horizontal line.

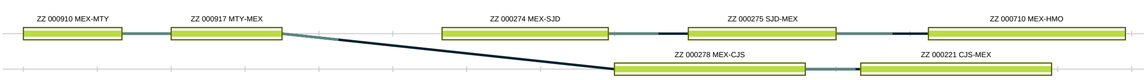


Figure 4-2: Slack increase in the `AR` model

In this example, we can see that the new aircraft connection (ZZ 917 - ZZ 278) adds almost three hours of slack into the schedule. It decreases the elapsed time for the aircraft flying the sequence of flights in the top horizontal line and increases that for the aircraft operating the flights in the lower horizontal line.

2) For this airline, a minimum turn time required for each aircraft connection is determined not only by the aircraft type and the connection airport, but also by the departure airport of the inbound flight and arrival airport of the outbound flight. As a result, it is possible that the set of included connections requires less turn time and hence increases slack in the resulting schedule. The contribution of this factor to the slack increase is, however, not as significant as the former.

The percentage of original aircraft connections (e.g., aircraft connections in the original schedule) included in each solution reflects the degree of difference between the resulting schedule and the original schedule. We report this statistic because airlines may prefer schedules that require minimal modifications to the original schedule. In addition, the assumptions that we make regarding maintenance opportunities and crew feasibility, are more likely to hold with fewer schedule changes.

In both solutions, to the robust `AR` models, approximately 30% of the original aircraft connections are modified. Figure 4-3 depicts the `AR_maxEffACSlack15` schedule for March 1, 2008. Again, flights in the original schedule that are operated by the same aircraft are drawn along the same horizontal line. Therefore, any aircraft connections in the figure that are not horizontal indicate changes to the original aircraft routes.

Flight Delay Analysis

On average, both `AR_minPD` and `AR_maxEffACSlack15` solutions improve all flight delays statistics under consideration. Moreover, the `AR_maxEffACSlack15` solution shows larger expected improvements relative to the `AR_minPD` solution.

In Figures 4-4, we show evaluated propagated delays on March 1-25, 2008. We can see that both `AR_minPD` and `AR_maxEffACSlack15` schedules perform better than the `Original` schedule in all but one case (`AR_minPD` on March 16). The decreases in delay propagation are very significant on bad days– those with high level of delays, such as March 8 (23%), 9 (24%), and 14 (31%); while only small improvements are achieved on typical days.

In 10 out of 25 days, the `AR_minPD` and `AR_maxEffACSlack15` solutions have exactly the same evaluated propagated delays. (Note that the solutions are not necessarily the same.) There are also some cases where `AR_maxEffACSlack15` performs a little worse than `AR_minPD`. However, on bad days, the performance of `AR_maxEffACSlack15` is much better. For example, on March 23, the propagated delay reduction achieved by `AR_maxEffACSlack15` is about 8% more than `AR_minPD`. Indeed, on March 16, the `AR_minPD` solution has even higher propagated delay than the original solution. This

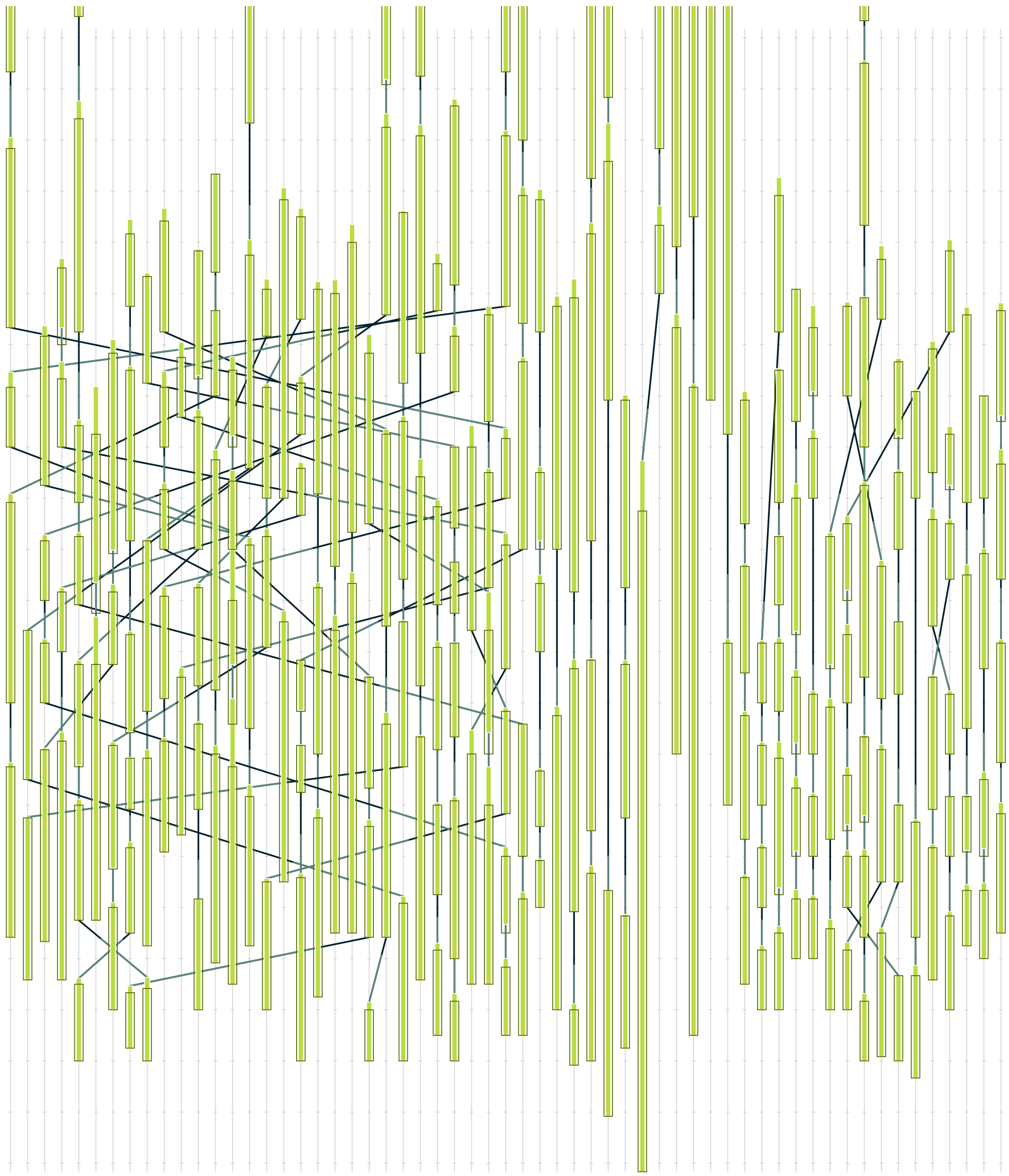


Figure 4-3: The `AR_maxEffACSlack15` solution for March 1, 2008

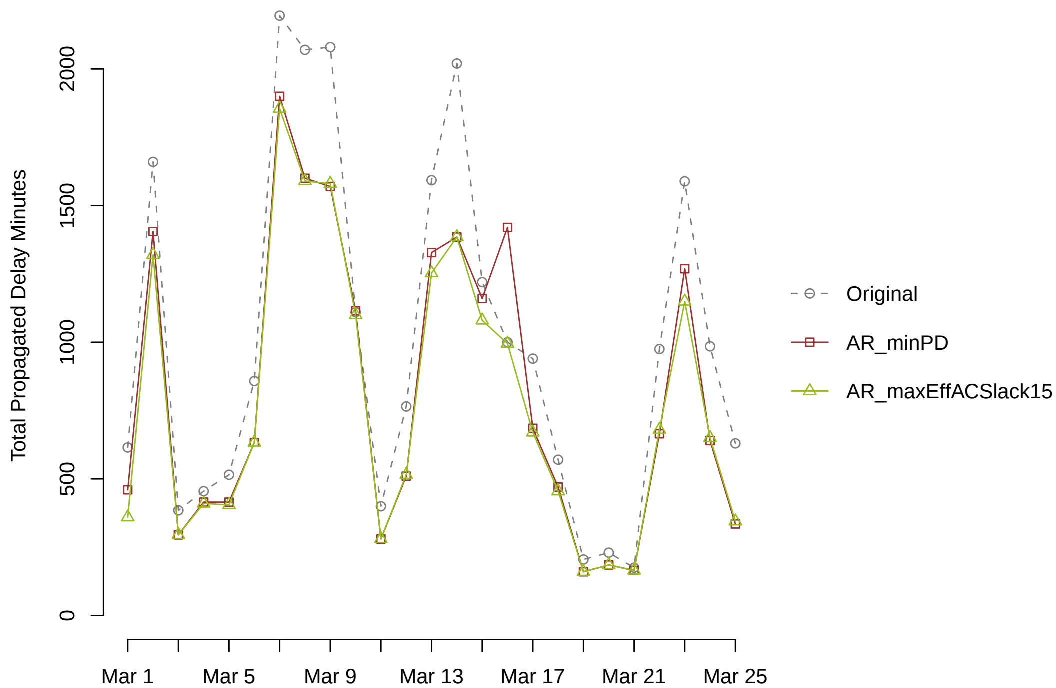


Figure 4-4: Evaluated total propagated delays on March 1-25, 2008 for the `AR` models

result suggests that the objective function of maximizing the total expected effective slack (with an appropriate limit) is more robust than minimizing the total expected propagated delay.

Tables 4.3 and 4.4 summarize the distributions of propagated delays and total arrival delays over 25 days. In `AR_minPD` and `AR_maxEffACSlack15` solutions, propagated delays are reduced every positive range. Consequently, the percentage of flights with positive propagated delays is reduced as shown in Table 4.2. In addition, the distribution of total arrival delays is shifted towards smaller delays. This results in reductions in total arrival delay and improvements in 15-minute and 60-minute on-time performance values. More importantly, the reduction of flights with long delays can potentially reduce the number of misconnecting passengers.

One observation we want to mention is that minimizing total propagated delay does not necessarily maximize 15-minute on-time performance. This is because a long delay propagated along one aircraft route may be split into smaller delays in several routes after re-routing. Figure 4-5 exemplifies the situation. The number associated

Propagated Delay (mins)	0	(0,30]	(30,60]	(60,90]	(90,120]	>120
`Original` (%)	82.26	14.89	1.65	0.55	0.26	0.38
`AR_minPD` (%)	85.14	12.61	1.34	0.40	0.18	0.33
`AR_maxEffACSlack15` (%)	85.60	12.31	1.26	0.33	0.18	0.31

Table 4.3: Distributions of propagated delays for the `AR` models

Total Arrival Delay (mins)	0	(0,15]	(15,60]	(60,120]	>120
`Original` (%)	44.33	32.20	20.36	2.17	0.94
`AR_minPD` (%)	45.27	32.55	19.44	1.85	0.89
`AR_maxEffACSlack15` (%)	45.47	32.56	19.32	1.75	0.89

Table 4.4: Distributions of total arrival delays for the `AR` models

with each flight indicates the independent delay. In feasible solution I, there is a large delay propagation on the connection from ZZ 8 to ZZ 9, and only flights ZZ 8 and ZZ 9 are delayed longer than 15 minutes. Alternatively, in feasible solution II, both aircraft connections experience some propagated delay. Although the total propagated delay is smaller than feasible solution I, flight ZZ 5 is also delayed longer than 15 minutes in addition to ZZ 8 and ZZ 9.

This situation, however, does not happen frequently. Suppose, in the example, there is a sequence of flights following flight ZZ 9, and the propagated delay is sufficiently large. It is very likely that the delay will propagate to the subsequent flights after ZZ 9, and many more flights will be delayed longer than 15 minutes. Consequently, on average, 15-minute on-time performance is improved when propagated delay is minimized, as presented in Table 4.2.

(a) Feasible solution I

(b) Feasible solution II

Figure 4-5: An example illustrates inconsistency in minimizing the total propagated delay and maximizing 15-minute on-time performance.

Passenger Delay Analysis

In order to keep the model tractable, we do not explicitly take into account passenger delays in the objective functions. We, however, anticipate that minimizing a proxy of flight delay should also result in decreasing passenger delay. The results in Table 4.2 confirm our expectation. In both `AR_minPD` and `AR_maxEffACSlack15` solutions, total passenger delay, as well as the number of disrupted passengers, are decreased.

Note that although the solutions to the `AR` models reduce the number of disrupted passengers by two or three passengers on average, the total passenger delay decreases by more than 10,000 minutes. This suggests that the decrease in total passenger delay is mainly achieved through smaller flight delays for the majority of passengers. Figure 4-6 shows the evaluated total passenger delays on March 1-25, 2008. We can see that these results exhibit similar trends as those for propagated delays, shown in Figure 4-4. In particular, the decreases in passenger delay are significant on bad days, while only small improvements are achieved on typical days.

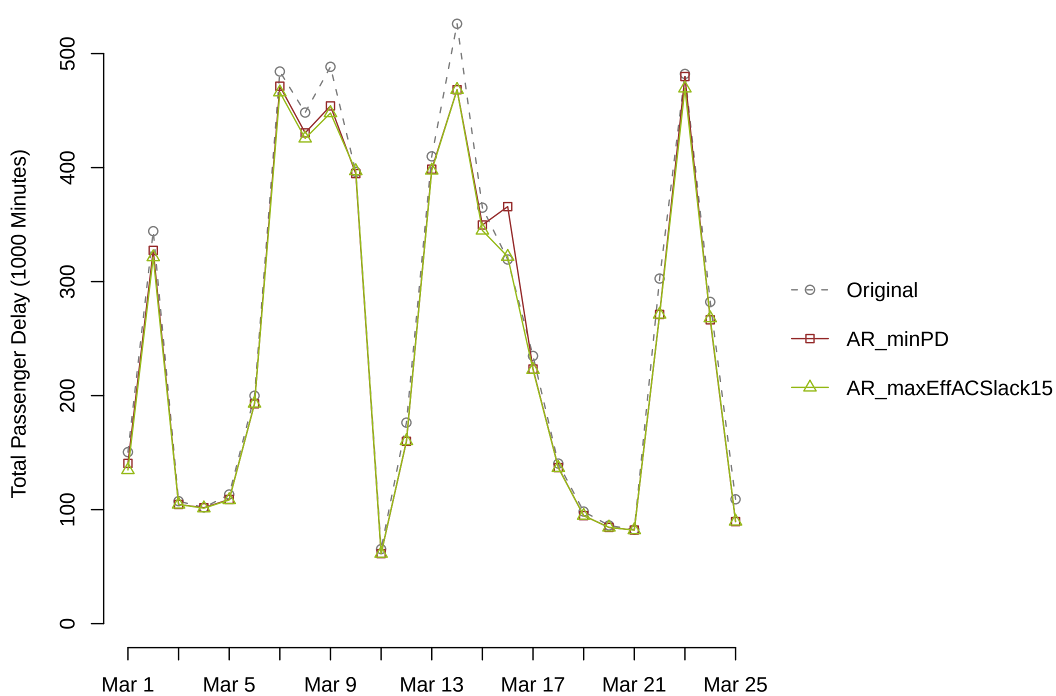


Figure 4-6: Evaluated total passenger delays on March 1-25, 2008 for the `AR` models

Nonetheless, it is important to point out that improvement in all the passenger-centric metrics is not assured. In fact, re-routing aircraft may result in more disrupted passengers. The evaluated number of disrupted passengers, shown in Figure 4-7, demonstrates the point. We can see that there are many cases where our generated schedules cause more disrupted passengers than for the original case. The large reduction in the number of disrupted passengers on March 2 and 22, however, result in fewer disrupted passengers on average for the `AR` model solutions.

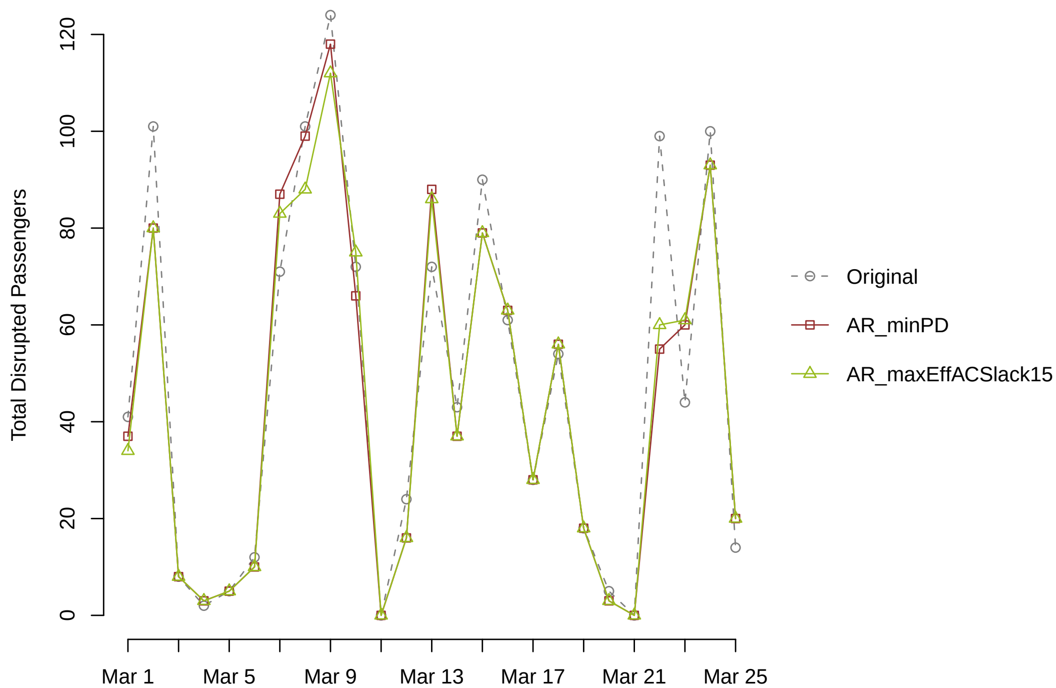
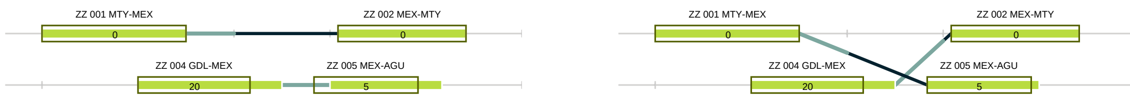


Figure 4-7: Evaluated number of disrupted passengers on March 1-25, 2008 for the `AR` models

To see how re-routing aircraft can result in disrupted passengers, consider the example in Figure 4-8. Again, the numbers indicate the independent delays of each flight. Suppose there are connecting passengers from ZZ 4 to ZZ 5. In the original schedule, the connecting passengers will always make their connections because flights ZZ 4 and ZZ 5 are flown by the same aircraft. In contrast, assume that minimizing the total expected propagated delay or maximizing the total expected effective slack yields a solution allowing ZZ 5 to depart on time. In this case, the connecting passengers from ZZ 4 to ZZ 5 miss their connections.



(a) Original (b) Robust Solution

Figure 4-8: Passenger disruption due to aircraft re-routing

Generally speaking, for each passenger connection, although the inbound flight is delayed, a passenger can still make his or her connection if the outbound flight is delayed as well. In terms of passenger misconnections, *relative* flight delays for each connection are more important than *absolute* flight delays. Therefore, it is possible that the original schedule may have more delayed flights, but fewer disrupted passengers. Readers are referred to [10] and [9] for the detailed discussion about the discrepancy between flight and passenger delays.

If, however, there are connecting passengers from ZZ 5 to other flights, ZZ 5 departing on time can decrease the likelihood of misconnection for those passengers. Consequently, the improvement in passenger delay statistics depends on the number passengers connecting to and from flight ZZ 5.

Moreover, total passenger delay is affected by the number of disrupted passengers and by which itineraries are disrupted. Some disrupted itineraries include a flight that operates only daily, and the disrupted passengers might have to wait for a full twenty four hours before they are re-accommodated. Others might involve flights that are operated hourly, and the disrupted passengers are then re-accommodated within an hour or so. Therefore, given two different solutions that yield exactly the same number of misconnecting passengers, the corresponding total passenger delays can be very different.

In summary, minimizing the total expected propagated delay or maximizing the total expected effective slack do seem, however, to reduce passenger delays, even though neither passenger delays nor disruptions are explicitly modeled.

	AR_minPD	AR_minPD _maxOrgConn	AR_maxEffACSlack15	AR_maxEffACSlack15 _maxOrgConn
Schedule Statistics				
Total A/C Connection Slack (mins)	7459.16	7454.16	7494.36	7504.96
% of Original A/C Connections Included	67.92%	69.74%	68.60%	69.96%
Flight Delay Statistics				
Total Propagated Delay (mins)	818.60	818.60	781.40	781.80
% of Flights with PD > 0	14.86%	14.83%	14.41%	14.40%
Total Arrival Delay (mins)	2965.56	2965.56	2929.16	2929.56
15-min On-Time Performance	77.82%	77.83%	78.02%	78.03%
60-min On-Time Performance	97.26%	97.27%	97.36%	97.37%
Passenger Delay Statistics				
Total Pax Delay (mins)	250325	250249	246903	246820
Total Disrupted Pax (pax)	45.16	45.00	44.72	44.56

Table 4.5: Average performance evaluation statistics over 25 days of alternative optimal solutions to AR_minPD and AR_maxEffACSlack15

4.2.2 Discussion on Models

The discussion in the previous section focuses mainly on the performance of the two solutions, AR_minPD and AR_maxEffACSlack15, compared to the Original schedule. In this section, we will address some attributes of the robust aircraft re-routing model including the objective functions, model parameters, and solution quality.

Multiple Optimal Solutions

As discussed in section 3.1.5, the AR model typically yields multiple optimal solutions, which might not be equally effective with respect to other performance metrics or objective functions. We select, among the optimal AR_minPD and AR_maxEffACSlack15 solutions, the solutions that maximize the number of aircraft connections in the original routing included in the solutions. In particular, we solve the AR models again using (AR-11) as an objective function with an additional constraint ensuring optimality of the solution with respect to the initial objectives. We denote the AR_minPD and AR_maxEffACSlack15 solutions that maximize the number of original aircraft connections as AR_minPD_maxOrgConn and AR_maxEffACSlack15_maxOrgConn, respectively. The average performance evaluation statistics are summarized in Table 4.5.

The underlying idea of this selection is to potentially preserve good features of the original aircraft routing. With about 1 to 2 percent increase in the number of original aircraft connections, the AR_minPD_maxOrgConn and AR_maxEffACSlack15_maxOrgConn solutions show slight improvements over the AR_minPD and AR_maxEffACSlack15 solutions in almost all performance metrics. Even though the results here show limited improvements, we still expect significant improvements when we solve the AR models for a longer planning period or a larger airline with more re-routing opportunities. Note that most re-routing opportunities exist at hubs, and this particular airline, Airline A, has only one major hub.

Lastly, because solving the AR models can be computationally expensive, a decision to optimize a secondary objective depends largely on the potential improvement, and hence the airline's underlying network. For Airline A, it might not be worth solving the second-stage optimization problem.

Increasing Slack versus Performance Improvement

Recall that the AR models assume that every aircraft that is scheduled to operate on a particular day is ready at the beginning of the day and also available until the end of the day. This results in a solution with a significant increase of total aircraft connection slack as discussed in the previous section. It is questionable that the improvements shown earlier are mainly attributed to the increasing slack in the solutions.

We noticed that a large increase in slack occurs when the *first* flight of some aircraft route in the original routing is *preceded* by some other flights in the solution to the AR models, as illustrated in Figure 4-2, or similarly when the *last* flight of some aircraft route in the original routing is *followed* by some other flights in the solution to the AR models. Therefore, to forbid solutions with this property, we re-define a string as a sequence of flights such that (i) the first flight of the string must originally be the first flight of some aircraft route; and (ii) the last flight of the string must originally be the last flight of some aircraft route. Note that this does not completely rule out the possibility that some aircraft have longer elapsed times between the start and end of flying in a given day.

	Original	AR_minPD'		AR_maxEffACSlack15'	
Schedule Statistics					
Total A/C Connection Slack (mins)	6676.76	6683.36	+0.10%	6690.76	+0.21%
% of Original A/C Connections Included	-	72.39%		72.70%	
Flight Delay Statistics					
Total Propagated Delay (mins)	1009.60	854.80	-15.33%	839.00	-16.90%
% of Flights with PD > 0	17.74%	15.36%		15.12%	
Total Arrival Delay (mins)	3141.16	2999.56	-4.51%	2983.96	-5.00%
15-min On-Time Performance	76.53%	77.70%		77.80%	
60-min On-Time Performance	96.89%	97.17%		97.26%	
Passenger Delay Statistics					
Total Pax Delay (mins)	260565	249277	-4.33%	248111	-4.78%
Total Disrupted Pax (pax)	47.56	42.76	-10.09%	41.84	-12.03%

Table 4.6: Average performance evaluation statistics over 25 days of solutions to the AR models with the new definition of a flight string

Given this definition of a flight string, the count constraints (AR-3) are automatically satisfied, and thus can be removed from the formulation. Let AR_minPD' and AR_maxEffACSlack15' be the corresponding solutions with this new definition of a flight string. The performance of the solutions is summarized in Table 4.6.

As we expected, both AR_minPD' and AR_maxEffACSlack15' solutions now have almost the same amount of slack as the Original schedule. They still, however, improve every performance evaluation metric. In particular, except for total number of disrupted passengers, the improvements are only slightly inferior to the AR_minPD and AR_maxEffACSlack15 solutions, as shown in Table 4.2. For example, the reduction in total propagated delay drops by only 4 and 6 percent for AR_minPD' and AR_maxEffACSlack15' solutions, respectively.

More importantly, the reduction in disrupted passengers achieved by the AR_minPD' and AR_maxEffACSlack15' solutions is more significant, compared to the AR_minPD and AR_maxEffACSlack15 solutions. Because, in this case, more original aircraft connections are included in the solutions, the results support our earlier discussion that re-routing aircraft can result in disrupted passengers.

Solution Quality

In order to exhibit the quality of a solution to the AR model, we compare its performance with the other two solutions, denoted by the suffixes _expected and _perfect-

Info. The only difference among these solutions is the approach of using historical data. Recall that so far we consider each day of operation in January and February as one instance of delay scenario ω, and thus the set of delay scenarios Ω has cardinality 60. We also assume that each disruption scenario is equally likely.

For an _expected solution, the set of delay scenarios Ω has a single element representing the *average* independent delays of every flight, obtained from January and February data. This simple approach of using historical data can be useful when 1) daily flight operations data are not accessible or not very complete, but estimates of average flight delays are available; or 2) it is computationally too expensive to solve the model with many delay scenarios.

The drawback of this solution is that it ignores the stochastic nature of delays and the correlations of delays among different flights, which can partially be captured by using many different delay scenarios. Furthermore, it can be shown that the function of total propagated delay and the function of total effective slack are convex and concave, respectively. By Jensen's inequality, for a convex(concave) function, the expected value of functions is no smaller(larger) than the function of the expected value. Therefore, it follows that an _expected solution underestimates the total expected propagated delay and overestimates the total expected effective slack, as compared to the base solutions, AR_minPD and AR_maxEffACSlack15. Intuitively, for total propagated delay, it is possible that an *average* independent arrival delay of a flight is so small that it does not propagate in the _expected solution, although for some instances of delay scenarios, the delay could be very large and potentially propagate to subsequent flights.

For a _perfectInfo solution, the set of delay scenarios Ω has a single element representing the *actual* independent delays of every flight. Note that this is the same set of independent delays we use to evaluate the performance of schedules. In other words, we solve the AR models as if we have perfect information of the future operations. This solution provides a bound on an improvement we can possibly achieve through a particular model.

Table 4.7 summarizes the performance of each solution. The results show that even the _expected solutions, which are computationally less expensive, are reasonably better than the Original schedule. Moreover, the superior performance of AR_minPD and AR_maxEffACSlack15 solutions, compared to the corresponding _expected solutions, indicates that it is beneficial to use many different delay scenarios from historical data to capture the stochasticity of delays. For instance, the average total propagated delay of the AR_minPD solution is more than 6% smaller than the AR_minPD_expected solution.

Remarkably, the performance of the AR_maxEffACSlack15_expected solution is somewhat comparable to the AR_minPD solution, which is more difficult to obtain. This again emphasizes the strength of the objective of maximizing the total expected effective aircraft connection slack.

The performance of the _perfectInfo solutions indicate that the AR models, with the objective of minimizing total expected propagated delay or maximizing total expected effective slack, can reduce propagated delay by at most 30% and passenger delay by at most 8%, on average. We will again compare these bounds with the flight schedule re-timing models and the block time adjustment models.

Figure 4-9 shows the evaluated propagated delays on March 1-25, 2008 for different approaches of using historical data. The performance gaps between each solution are clearly illustrated here. We can see that all solutions actually perform relatively the same on typical days, and hence the difference in average performance of each solution is mainly driven by their performance on bad days.

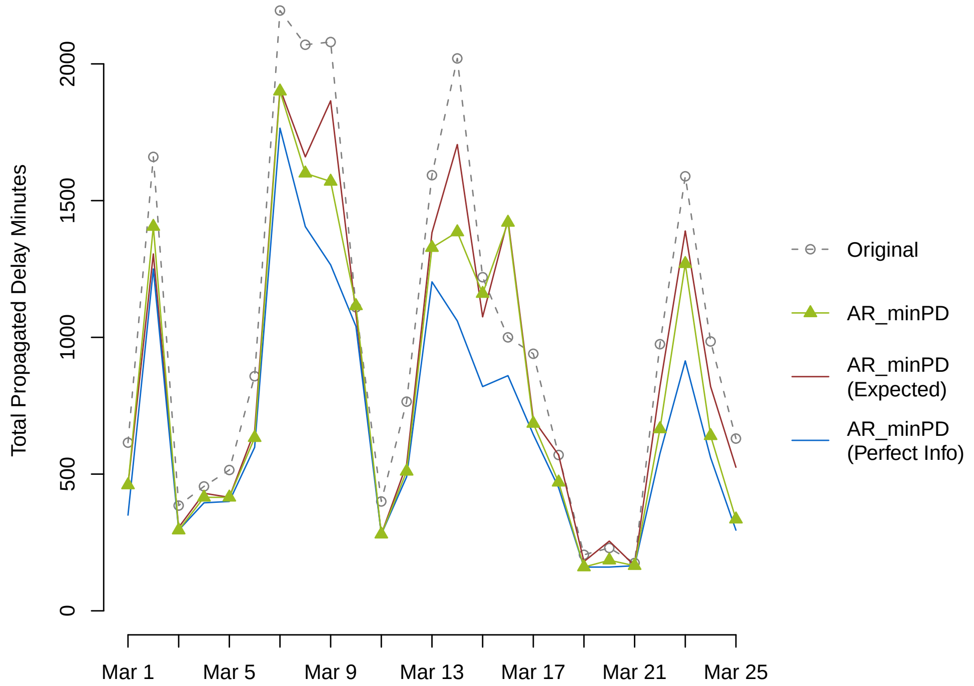
Cap Values

In the AR model that maximizes the total expected effective slack, we need to specify a cap Γ_{ij} for each aircraft connection from flight i to flight j. So far, we only show the results for the AR_maxEffACSlack15 solution for which the caps are set to 15 minutes for every aircraft connection.

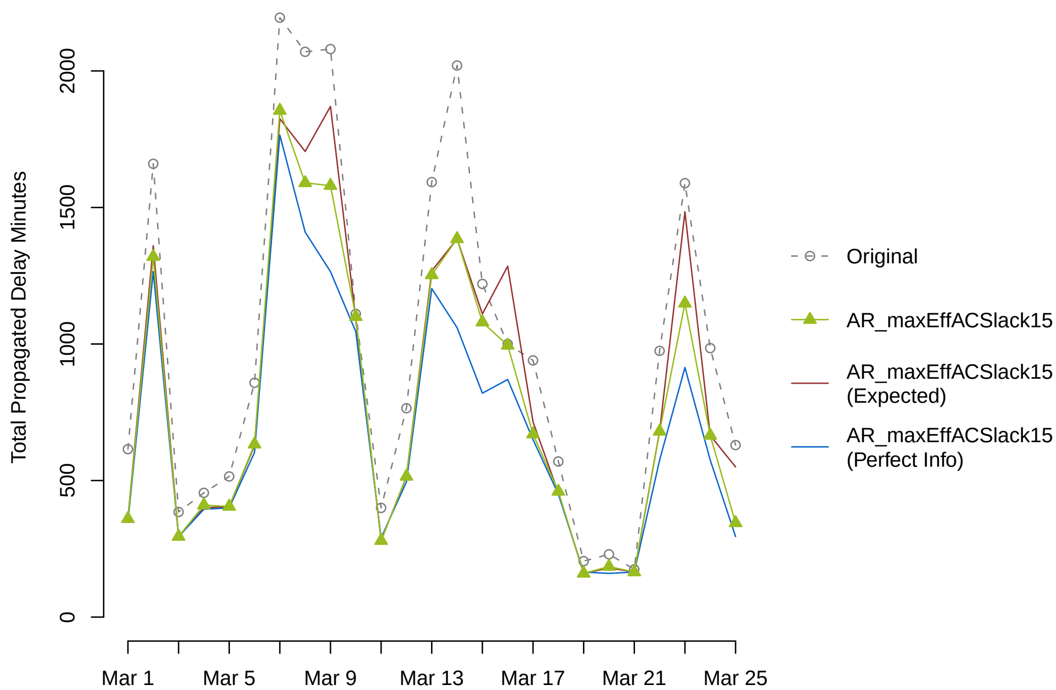
It is not clear a priori how we should set the values of caps. On one extreme, we can set all the caps to zero, which is equivalent to minimizing propagated delay, as proved in Section 3.1.4. In this case, the model focuses mainly on the connections

	Original	AR_minPD	AR_minPD _expected	AR_minPD _perfectInfo	AR_maxEffACSlack15	AR_maxEffACSlack15 _expected	AR_maxEffACSlack15 _perfectInfo
Schedule Statistics							
Total A/C Connection Slack (mins)	6676.76	7459.16	7241.56	7125.76	7494.36	7365.96	7333.76
% of Original A/C Connections Included	-	67.92%	69.35%	68.91%	68.60%	68.51%	67.77%
Flight Delay Statistics							
Total Propagated Delay (mins)	1009.60	818.60	877.00	696.00	781.40	833.40	699.20
% of Flights with PD > 0	17.74%	14.86%	15.50%	13.62%	14.41%	14.56%	13.62%
Total Arrival Delay (mins)	3141.16	2965.56	3022.36	2842.76	2929.16	2981.16	2845.76
15-min On-Time Performance	76.53%	77.82%	77.59%	78.40%	78.02%	77.80%	78.35%
60-min On-Time Performance	96.89%	97.26%	97.07%	97.59%	97.36%	97.16%	97.59%
Passenger Delay Statistics							
Total Pax Delay (mins)	260565	250325	252951	238665	246903	249863	238026
Total Disrupted Pax (pax)	47.56	45.16	44.12	44.20	44.72	43.56	42.56

Table 4.7: Average performance evaluation statistics over 25 days of solutions to the AR models with different approaches of using historical data



(a) `AR_minPD`



(b) `AR_maxEffACSlack15`

Figure 4-9: Evaluated total propagated delays on March 1-25, 2008 for different approaches of using historical data that have negative effective slack (or, equivalently, positive propagated delay) and ignores other connections for which effective slack might be close to zero. On the other extreme, we can set all the caps to infinity. In this case, any amount of slack in an aircraft connection fully contributes to the objective value, and more slack might be unnecessarily added to the connections that already have a reasonable amount of slack. Ideally, we wish to find the cap values that balance trade-offs between the two cases and yield the best result.

Figure 4-10 depicts some key performance statistics of the solutions with different cap values. The results confirm that the performance of a solution is maximized when the caps are not too large or too small. In this case, setting all the caps to 20 minutes yields the best result. Interestingly, the plots in Figure 4-10 are relatively flat around the minimum points, that is, if we pick other values of caps around 20 minutes, the improvements will drop only modestly. More importantly, setting caps to anything between 5 and 30 minutes, rather than zero, results in the better performance evaluation statistics. This again emphasizes the advantage of maximizing total expected effective slack over minimizing total expected propagated delay.

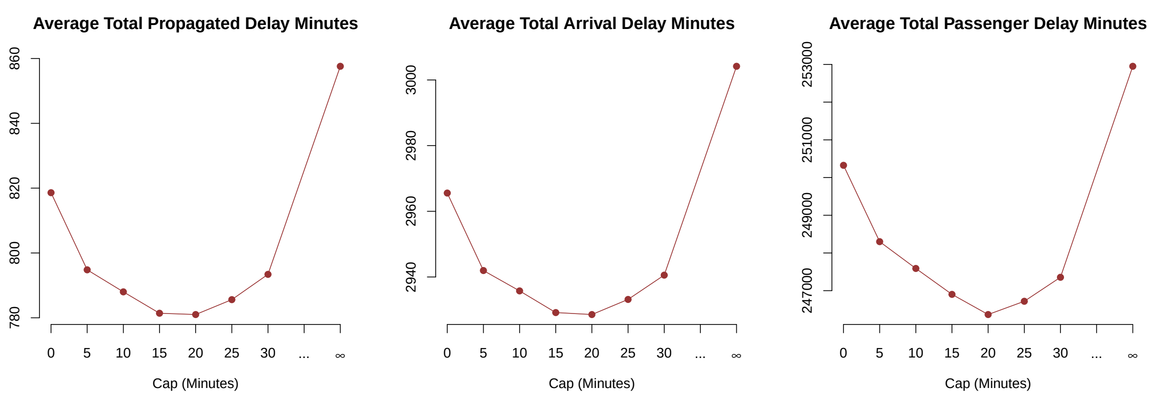


Figure 4-10: Average performance of solutions with different cap values

The evaluated total propagated delays on March 1-25, 2008 for different values of caps are shown in Figure 4-11. Although the performance of each solution does not differ significantly on typical days, on March 16, the `AR_maxEffACSlack0` and `AR_maxEffACSlack`∞ solutions perform much worse than the other solutions, including the `Original` schedule.

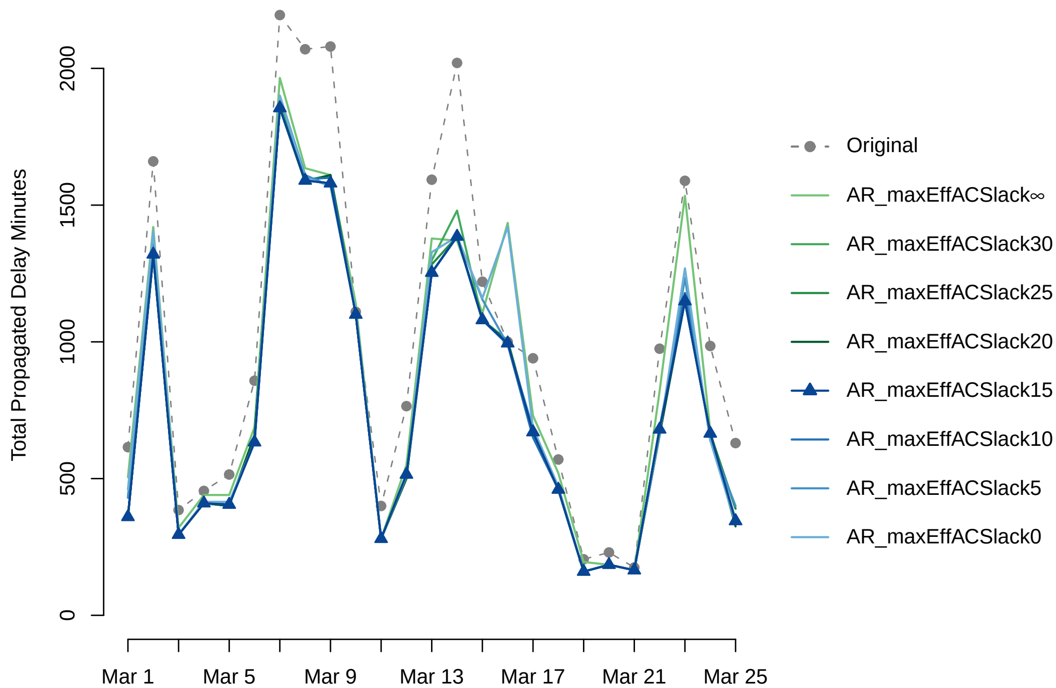


Figure 4-11: Evaluated total propagated delays on March 1-25, 2008 of solutions with different cap values

Recall that the caps need not be the same for every aircraft connection. An airline can use additional information to set values of caps to different levels for different fleet types, connection airports, and so forth.

Minimizing the Expected Total Arrival Delay

As discussed in Section 3.1.4, when an airline extensively pads its schedule to account for potential delays, it might be more appropriate to minimize the total expected arrival delay, rather than minimize the total expected propagated delay. This is, however, not the case for this particular airline. Only about 6% of flights have negative average independent arrival delays. Comparing the total propagated delay and the total arrival delay of the `Original` schedule in Table 4.2, we can see that propagated delay contributes only about one third of total arrival delay, that is, another two thirds of total arrival delay is due to independent arrival delays of each flight.

	AR_minPD	AR_minTAD
Schedule Statistics		
Total A/C Connection Slack (mins)	7459.16	7464.56
% of Original A/C Connections Included	67.92%	68.43%
Flight Delay Statistics		
Total Propagated Delay (mins)	818.60	818.00
% of Flights with PD > 0	14.86%	14.93%
Total Arrival Delay (mins)	2965.56	2964.76
15-min On-Time Performance	77.82%	77.80%
60-min On-Time Performance	97.26%	97.27%
Passenger Delay Statistics		
Total Pax Delay (mins)	250325	250589
Total Disrupted Pax (pax)	45.16	45.12

Table 4.8: Average performance evaluation statistics over 25 days (March 1-25, 2008) for the AR_minPD and AR_minTAD solutions

Let AR_minTAD denote the solution to the AR model that minimizes the total expected arrival delay (see (AR-10)). Table 4.8 compares the performance of the AR_minPD and AR_minTAD solutions. The results show that there is almost no difference between the two solutions. This is consistent with our earlier observation that only a small number of flights have negative average independent arrival delays, and hence minimizing the total expected propagated delay is almost equivalent to minimizing the total expected arrival delay.

Finally, according to Table 4.2, the AR_maxEffACSlack15 solution also outperforms the AR_minTAD solution in every performance evaluation metric.

4.3 Robust Flight Schedule Re-timing Model

In this section, we present the computational results obtained from the robust flight schedule re-timing models introduced in Section 3.2, with different objective functions and parameters. Again, let Original denote the airline's original schedule, and define the other solutions as follows.

FR_minPD	the solution to the FR model that minimizes the total expected propagated delay (see (FR-1))
FR_maxEffACSlack<Γ>	the solution to the FR model that maximizes the total expected effective *aircraft connection* slack with caps set equal to Γ minutes for every aircraft connection (see (FR-16))
FR_maxEffPaxSlack<Γ>	the solution to the FR model that maximizes the total expected effective *passenger connection* slack with caps set equal to Γ minutes for every passenger connection (see (FR-20))

Recall that, in the FR models, we need to specify a time window $[l_i, u_i]$ for which the departure time of flight i is allowed to change. Throughout this section, we assume the following, unless stated otherwise. For a flight leg i,

- if i is the first flight of some flight string, then $l_i = 0$ and $u_i = 15$;
- if i is the last flight of some flight string, then $l_i = -15$ and $u_i = 0$;
- otherwise, $l_i = -15$ and $u_i = 15$.

In other words, each flight is allowed to move at most 15 minutes earlier or later. Additionally, the first and last flights of each string are not allowed to move earlier and later, respectively. Consequently, the elapsed time between the start and end of flying for each aircraft remains the same.

4.3.1 Computational Results

Table 4.9 summarizes average statistics of Original, FR_minPD, FR_maxEffACSlack15, and FR_maxEffPaxSlack15 solutions over the period of March 1-25, 2008.

Schedule Statistics

First of all, we note that every solution has the same amount of total aircraft connection slack, because an aircraft routing is fixed in the FR models, and we assume

	Original	FR_minPD		FR_maxEffACSlack15		FR_maxEffPaxSlack15	
Schedule Statistics							
Total A/C Connection Slack (mins)	6676.76	6676.76		6676.76		6676.76	
Total Re-timing	-	1258.80		1223.32		1230.80	
Flight Delay Statistics							
Total Propagated Delay (mins)	1009.60	756.24	-25.10%	741.48	-26.56%	1076.40	+6.62%
% of Flights with PD > 0	17.74%	11.97%		11.34%		20.12%	
Total Arrival Delay (mins)	3141.16	2967.40	-5.53%	2952.52	-6.01%	3219.36	+2.49%
15-min On-Time Performance	76.53%	78.26%		78.48%		76.08%	
60-min On-Time Performance	96.89%	97.01%		97.04%		96.86%	
Passenger Delay Statistics							
Total Pax Delay (mins)	260565	256540	-1.54%	256211	-1.67%	260854	+0.11%
Total Disrupted Pax (pax)	47.56	64.52	+35.66%	67.16	+41.21%	36.80	-22.62%

Table 4.9: Average performance evaluation statistics over 25 days (March 1-25, 2008) for the FR models

that the first and last flights of each string are not allowed to move earlier and later, respectively.

Recall that for the AR models, we report the number of original aircraft connections included as a measure of difference between the resulting schedule and the original schedule. Analogously, for the FR models, we compute the total change in the departure times, denoted as "Total Re-timing" in Table 4.9. In particular, a total re-timing statistic is given by $\sum_{i \in F} |x_i|$, where x_i is the difference between the new and the original departure time of flight i in the set F of all flight legs. According to Table 4.9, all solutions to the FR model are different from the Original schedule by approximately 1200 minutes of re-timing. To illustrate the extent of difference, the FR_maxEffACSlack15 solution for March 1, 2008 is depicted in Figure 4-12. The visualization notation for flight re-timing is given in Figure 2-2.

Flight Delay Analysis

The results show that, on average, the FR_minPD and FR_maxEffACSlack15 solutions perform better than the Original schedule with respect to every flight delay statistic. Also, the FR_maxEffACSlack15 solution shows slightly larger improvements relative to the FR_minPD solution. In contrast, the FR_maxEffPaxSlack15 solution performs worse than the Original schedule in every flight delay statistic. This is simply because the objective function of maximizing the total expected effective passenger

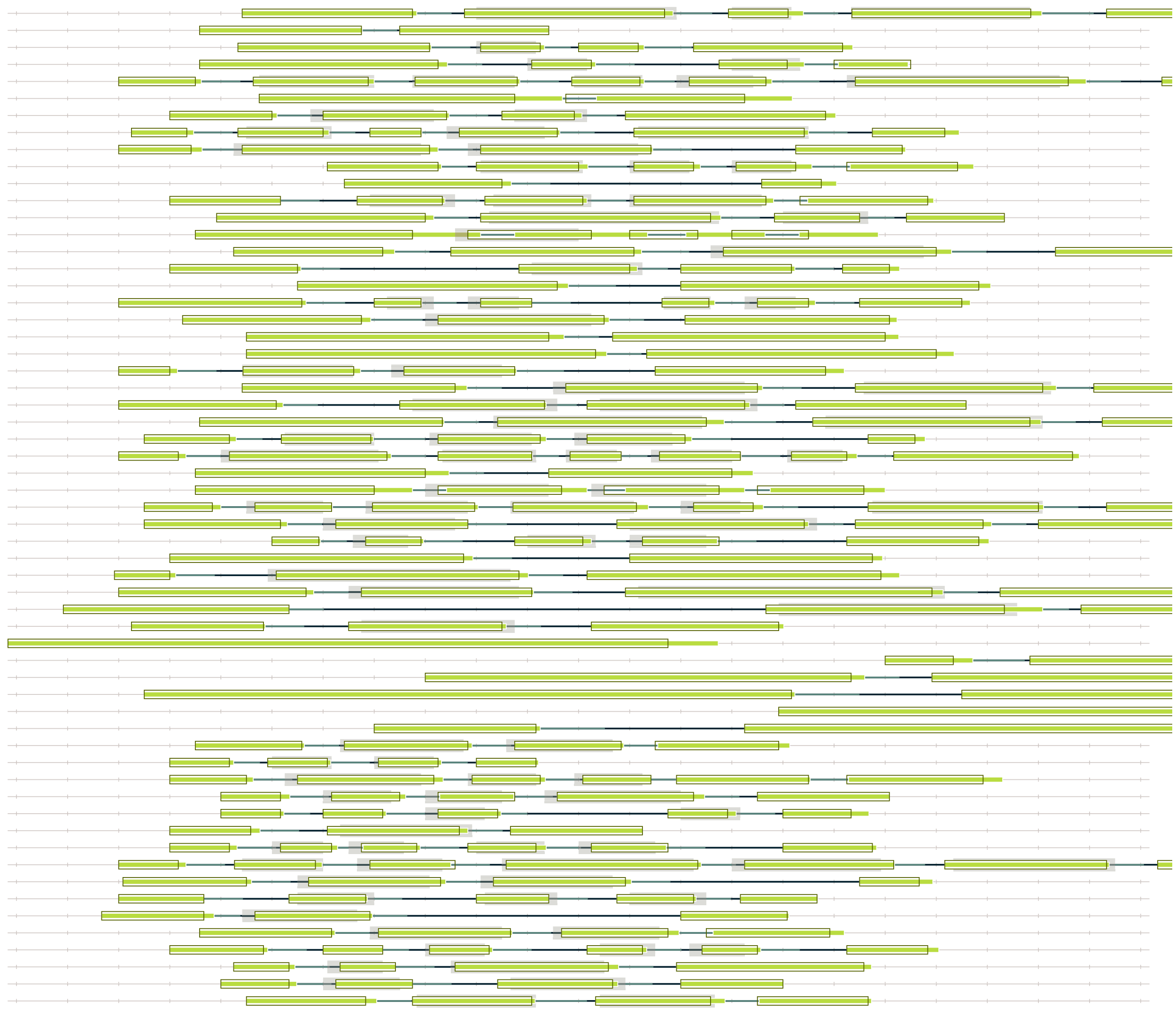


Figure 4-12: The FR_maxEffACSlack15 solution for March 1, 2008

connection slack has no direct link to flight delay improvements.

Comparing the performance of the `FR_minPD` and `FR_maxEffACSlack15` to the `AR_minPD` and `AR_maxEffACSlack15` solutions, we can see that even though the `FR_minPD` and `FR_maxEffACSlack15` solutions yield larger reductions in propagated delay and percentage of flights with positive propagated delay, the reduction in total arrival delay is slightly smaller than in the `AR_minPD` and `AR_maxEffACSlack15` solutions. This happens because some proportion of the decreases in propagated delay goes to the flights with negative independent arrival delays. Mathematically, because $TAD_j = \text{Max}(PD_{ij} + IAD_j, 0)$, a decrease in the propagated delay from flight i to flight j only affects the total arrival delay of flight j when $PD_{ij} + IAD_j$ is nonpositive.

The distributions of propagated delays and total arrival delays over 25 days for the `FR_minPD` and `FR_maxEffACSlack15` solutions are summarized in Table 4.10 and 4.11. In both solutions, total arrival delays are reduced in every positive range. Propagated delays, however, slightly increase in the (90,120] range. The results show that this slight increase is due to the mismatch between the expected delays in historical data and the actual delays in the day of operation. An example in Figure 4-13 illustrates the situation.

(a) `Original`

(b) `FR_maxEffACSlack15`

Figure 4-13: The mismatch between the expected delays in historical data and the actual delays in the day of operation

Figure 4-13 shows the expected actual operations of the `Original` schedule and the `FR_maxEffACSlack15` solution. According to the historical data, the average independent arrival delay (IAD) of flight ZZ 6513 is about 32 minutes, and the largest IAD is as long as 1 hour. On the other hand, the average IAD of flight ZZ 444 is only 12 minutes, and 60% of the historical IADs are zeroes. As a result, in the

FR_maxEffACSlack15 solution, ZZ 444 is scheduled to depart later in anticipation of delay propagation from ZZ 6513. However, the actual IADs of ZZ 6513 and ZZ 444 are 5 and 25 minutes, respectively. Consequently, flight ZZ 444 departs earlier in the Original schedule, and a smaller delay propagates from ZZ 444 to ZZ 278 and subsequently to ZZ 221.

This situation, however, does not occur frequently, and the FR_minPD and FR_maxEffACSlack15 solutions still, on average, improve the total propagated delay statistic. Because in the AR_minPD and AR_maxEffACSlack15 solutions, propagated delays are reduced in every positive range, this suggests that the robust flight schedule re-timing models are more sensitive to the discrepancy between historical delays and actual delays than the robust aircraft re-routing models.

Because in the FR models, we assume that the aircraft routing is fixed, and each flight is allowed to move at most 15 minutes, their solutions are not capable of reducing large delay propagation. In particular, we can increase an aircraft connection slack by at most 30 minutes by moving the inbound flight 15 minutes earlier and the outbound flight 15 minutes later, given that the corresponding aircraft and passenger connections are feasible. In contrast, it is possible for the AR models to increase an aircraft connection slack by more than 30 minutes, given that there exists such a re-routing opportunity. As a result, the percentage of flights with propagated delays in the ranges larger than 30 minutes is reduced more in the AR_minPD and AR_maxEffACSlack15 solutions than in the FR_minPD and FR_maxEffACSlack15 solutions.

On the other hand, the FR models are more flexible than the AR models in the sense that they allow finer re-allocation of slack. In particular, in the FR models, slack can be increased or decreased by any amounts from 0 to 30 minutes, given that the corresponding aircraft and passenger connections are feasible. On the contrary, in the AR models, an amount of increase or decrease in slack hinges on the available re-routing opportunities. Furthermore, some connections, especially those departing from spokes, may have no re-routing opportunities. Therefore, the AR models are not capable of reducing delay propagation in those connections. This flexibility advantage of the FR models results in larger reductions of propagated delays in the (0,30] range in the FR_minPD and FR_maxEffACSlack15 solutions, as compared to the AR_minPD and AR_maxEffACSlack15 solutions.

Propagated Delay (mins)	0	(0,30]	(30,60]	(60,90]	(90,120]	>120
Original (%)	82.26%	14.89%	1.65%	0.55%	0.26%	0.38%
FR_minPD (%)	88.03%	9.55%	1.36%	0.40%	0.30%	0.36%
FR_maxEffACSlack15 (%)	88.66%	8.93%	1.39%	0.36%	0.33%	0.33%

Table 4.10: Distributions of propagated delays for the FR models

Total Arrival Delay (mins)	0	(0,15]	(15,60]	(60,120]	>120
Original (%)	44.33%	32.20%	20.36%	2.17%	0.94%
FR_minPD (%)	46.69%	31.57%	18.74%	2.07%	0.93%
FR_maxEffACSlack15 (%)	46.55%	31.93%	18.56%	2.10%	0.86%

Table 4.11: Distributions of total arrival delays for the FR models

Figure 4-14 shows evaluated propagated delays on March 1-25, 2008. Comparing this figure with the similar plot for the AR models (Figure 4-4), we can see that, on good days (e.g., March 3-6), the solutions to the FR models can achieve larger reduction in total propagated delay because of the flexibility advantage of the FR models; while on bad days (e.g. 8, 9, and 14 March), the solutions to the AR models can do better because of its capability to mitigate impacts of large delays.

Passenger Delay Analysis

Although the feasibility of passenger connections are imposed in the FR formulation, the FR_minPD and FR_maxEffACSlack15 solutions, which focus only on minimizing flight delay metrics, result in the increases in average number of disrupted passengers from the Original schedule by 36% and 41%, respectively. However, the average total passenger delay still slightly decreases in both solutions.

Figure 4-15 illustrates how the FR models that minimize flight delay metrics cause passenger disruption. The figure shows the expected actual operation of the FR_max-EffACSlack15 solution on March 6, 2008. The departure times of flights ZZ 108 and ZZ 105 are pushed back in anticipation of delay propagation from flight ZZ 225.

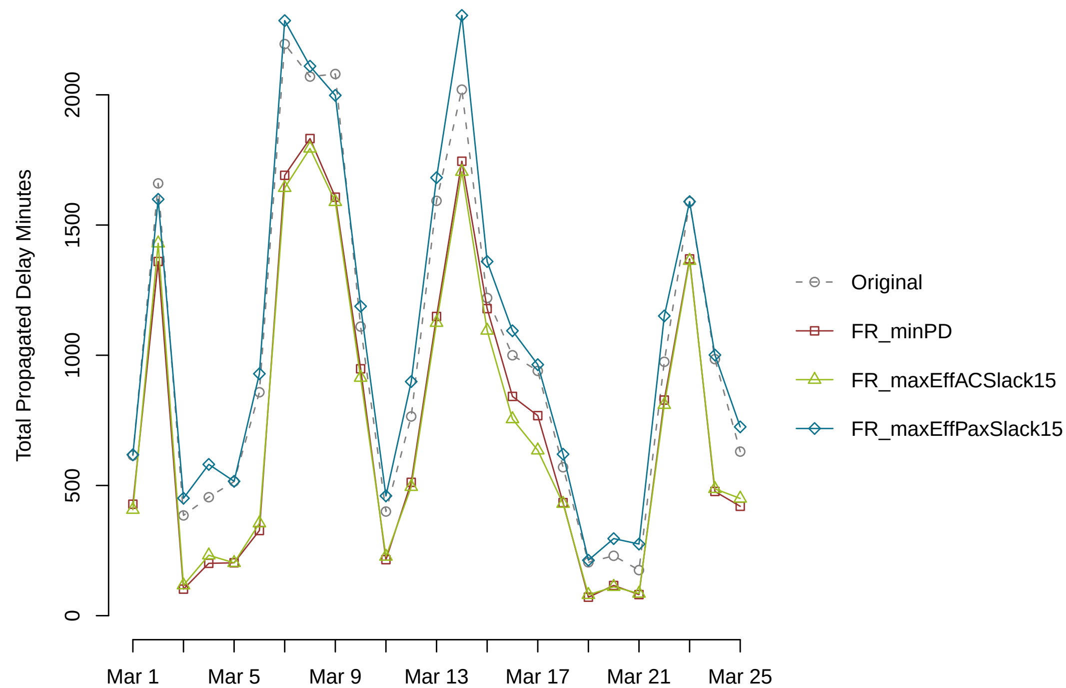


Figure 4-14: Evaluated total propagated delays on March 1-25, 2008 for the FR models

Also, the departure time of flight ZZ 222 is shifted earlier in anticipation of delay propagation to flight ZZ 217. Even though the passenger connection from ZZ 105 to ZZ 222 is feasible, it becomes much tighter. In particular, the passenger connection slack is decreased from 25 minutes to 5 minutes. Because the average IAD of flight ZZ 105 is 4.85 minutes, and almost 60% of the historical IADs are zeroes, there seems sufficient slack in the resulting schedule. However, the actual IAD of ZZ 105 turns out to be 10 minutes, and therefore the passengers do not have sufficient connection time, and they miss the connection.

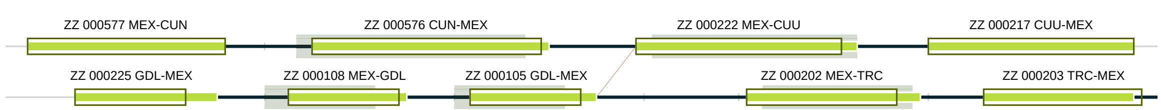


Figure 4-15: Flight delay versus passenger disruption

In contrast, the number of disrupted passengers is significantly reduced in the FR_maxEffPaxSlack15 solution. This indicates that our objective function of maximizing the total expected effective passenger connection slack serves as a good proxy for minimizing passenger misconnections. Despite the large reduction of disrupted passengers, the `FR_maxEffPaxSlack15` solution yields higher total passenger delay than the `Original` schedule.

Figure 4-16, 4-17, and 4-18 depicts the expected actual operations of the `Original`, `FR_maxEffACSlack15`, and `FR_maxEffPaxSlack15` schedules, respectively. A red line denotes a disrupted passenger connection. By comparing the three figures, we can easily see how each solution leads to fewer or more disrupted passengers.

At this point, two questions that may arise are

1) Why does total passenger delay decrease in the `FR_minPD` and `FR_maxEffACSlack15` solutions, in spite of the significant increase in disrupted passengers?; and

2) Why does total passenger delay increase in the `FR_maxEffPaxSlack15` solution, in spite of the significant reduction in disrupted passengers?

These happen because total passenger delay of this particular airline is mainly driven by flight delays, not passenger misconnections. According to the passenger booking data, almost 90% of passengers are local passengers– those that travel on a single flight leg. The passenger delay associated with these local passengers depends only on flight delays, given that their flights are not canceled. Moreover, we find that, on average, total *disrupted* passenger delay contributes less than 10% of total passenger delay. Total *disrupted* passenger delay depends not only on the number of disrupted passengers, but also on how fast an airline can re-accommodate disrupted passengers.

These findings suggest that for this particular airline, it might be more appropriate to focus on minimizing flight delays, rather than minimizing passenger misconnections. For other airlines with larger proportions of connecting passengers, however, the objective functions that minimize passenger misconnections might be more appropriate as the contribution of disrupted passenger delay to total passenger delay would be more significant.

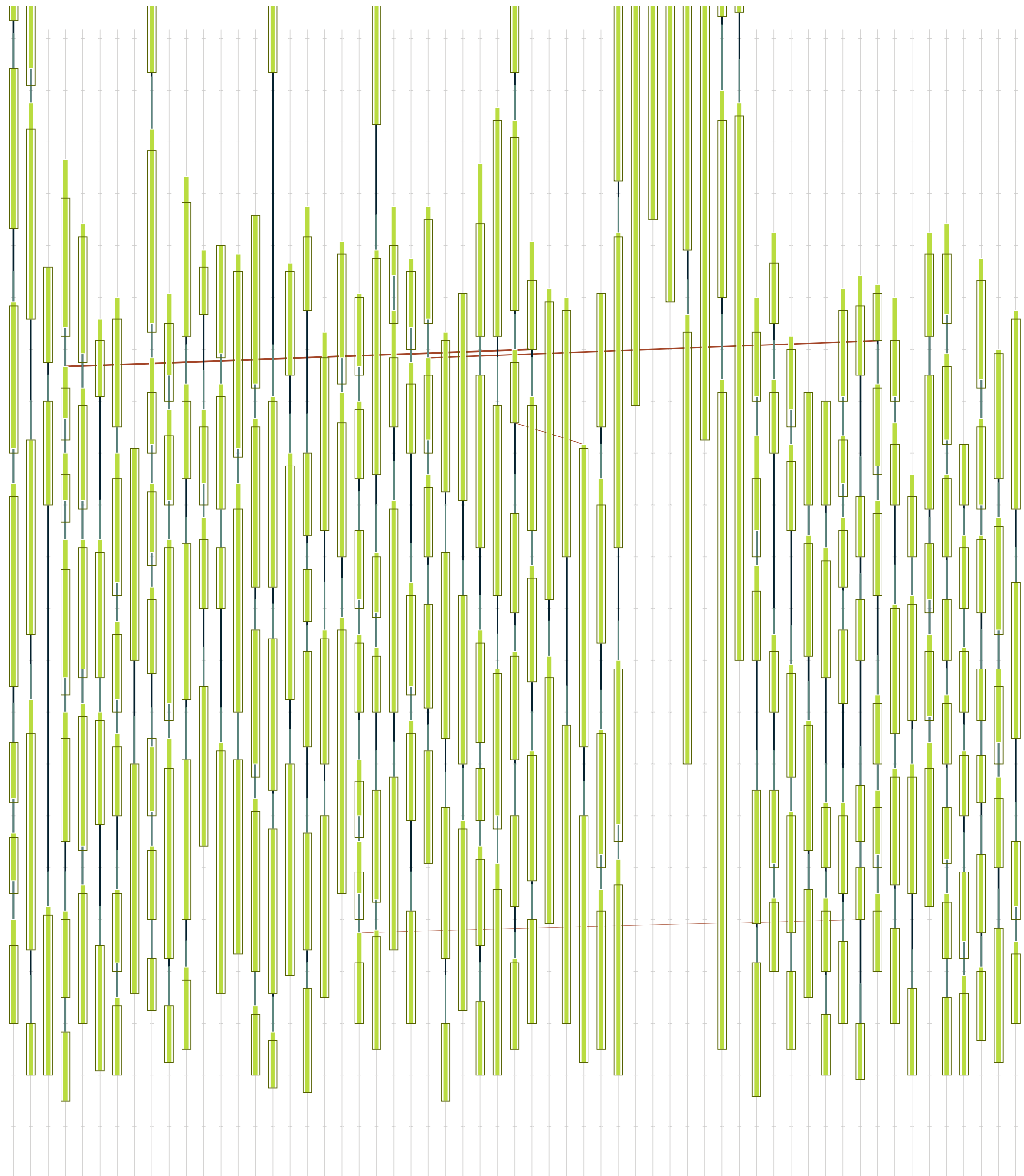


Figure 4-16: The `Original` solution for March 6, 2008

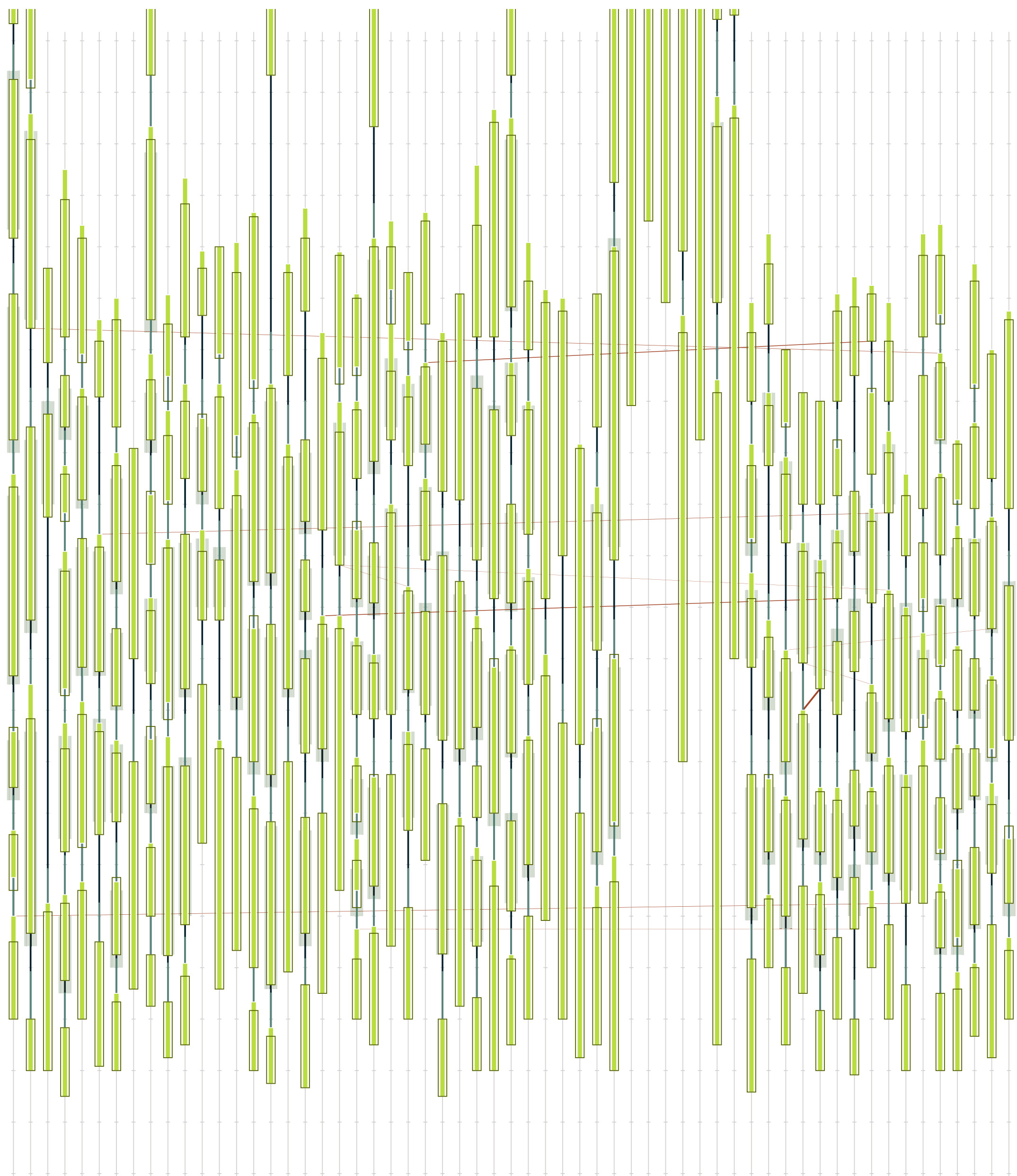


Figure 4-17: The `FR_maxEffACSlack15` solution for March 6, 2008

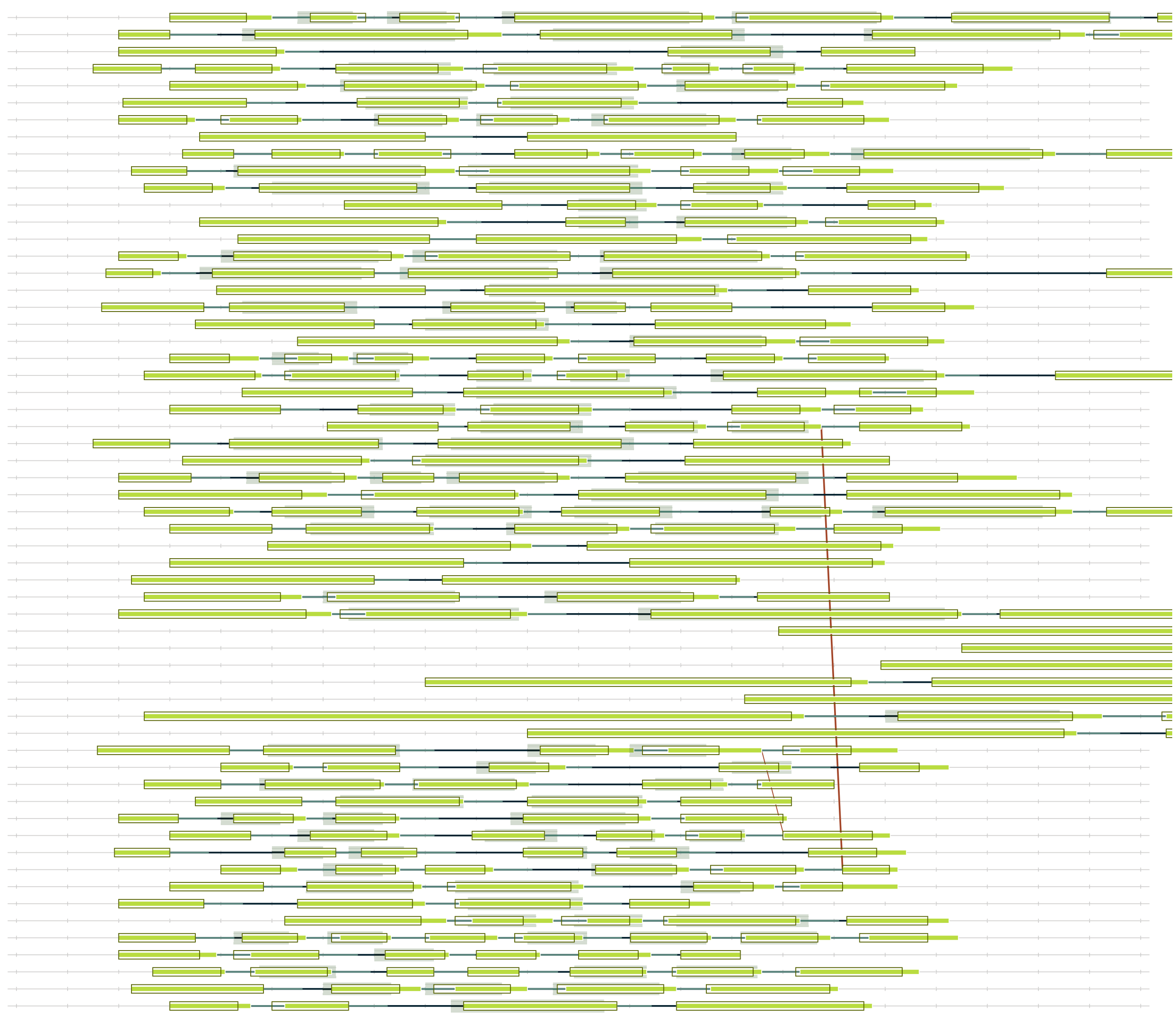


Figure 4-18: The FR_maxEffPaxSlack15 solution for March 6, 2008

To summarize, airline characteristics, such as a proportion of connecting passengers and contribution of disrupted passenger delay in total passenger delay, are crucial in determining which objective function is more appropriate for an airline to minimize the total passenger delay. The objective function that works well for one airline might lead to a bad solution for other airlines.

4.3.2 Discussion on Models

The Hybrid Objective Function

The `FR_maxEffACSlack15` and `FR_maxEffPaxSlack15` solutions, discussed in the previous section, represent two extreme solutions– one that focuses only on minimizing flight delays and one that focuses only on minimizing passenger misconnections. As a result, the number of disrupted passenger increases in the `FR_maxEffACSlack15` solution, and every flight delay statistic worsens in the `FR_maxEffPaxSlack15` solution. An airline might be interested in a solution that performs somewhere in between the two, i.e., a solution that results in small flight delays and at the same time does not cause too many passenger disruptions.

Instead of maximizing the total expected effective slack for only aircraft connections or only passenger connections, we now consider a convex combination of the two objective functions, (FR-16) and (FR-20), with a weight $\lambda \in [0,1]$ for aircraft connections. The weight λ should be set according to the sensitivity of total passenger delay to flight delays. For instance, if total passenger delay is mainly driven by flight delays, the λ should be set close to 1. Using the notation introduced in Section 3.2.2, the resulting formulation is given by:

$$\text{Minimize} \qquad \sum_{\omega\in\Omega} p_\omega \left[\lambda \sum_{(i,j)\in A} \overline{aSlack}_{ij}^{\omega} + (1-\lambda) \sum_{(i,j)\in P} \overline{pSlack}_{ij}^{\omega} \right]$$

$$\begin{aligned} \text{subject to} \qquad & \overline{aSlack}_{ij}^{\omega} \leq aSlack'_{ij} - tad_i^{\omega} && \forall (i,j) \in A \\ & \overline{aSlack}_{ij}^{\omega} \leq \Gamma_{ij} && \forall (i,j) \in A \end{aligned}$$

$$
\begin{aligned}
&\overline{pSlack}_{ij}^{\omega} \leq pSlack'_{ij} - tad_i^{\omega} && \forall (i,j) \in P \\
&\overline{pSlack}_{ij}^{\omega} \leq \Gamma_{ij} && \forall (i,j) \in P \\
&\text{(FR-2)} - \text{(FR-12)}
\end{aligned}
$$

We solve this hybrid model with caps Γ_{ij} set equal to 15 minutes for every aircraft and passenger connection for different values of λ. We denote each solution with a weight λ as FR_hybrid15_<λ>. The average total propagated delay and the average total disrupted passengers for solutions with different values of λ are plotted in Figure 4-19. The figure clearly illustrates the trade-off between maximizing the total expected effective slack for aircraft connections and passenger connections. As λ increases, i.e., more priority is given to effective aircraft connection slack, average total disrupted passengers increases, and average total propagated delay decreases. Figure 4-20 shows the averages total passenger delay for solutions with different values of λ. In this case, the FR_hybrid15_0.7 solution minimizes the total passenger delay.

When the value of λ increases from 0.0 to 0.1, the reduction in total propagated delay is very large; while the increase in disrupted passenger is relatively small. Similarly, when the value of λ decreases from 1.0 to 0.9 and 0.8, the reduction in the disrupted passengers metric is very large; while the increase in total propagated delay is relatively small. This suggests that, even when a modest weight is put on one part of the objective function, the hybrid model is of great help in balancing the benefits of the two extreme solutions.

Finally, we note that the FR_hybrid15_0.7 solution is still inferior to the AR_maxEff-ACSlack15 solution, presented in the previous section, with respect to most of the performance evaluation metrics, especially for the number of disrupted passengers.

Multiple Optimal Solutions

As discussed in Section 3.2.4, there are typically multiple optimal solutions to the FR models. In this section, we explore the following solutions:

1) the optimal FR_minPD, FR_maxEffACSlack15, and FR_maxEffPaxSlack15 solu-

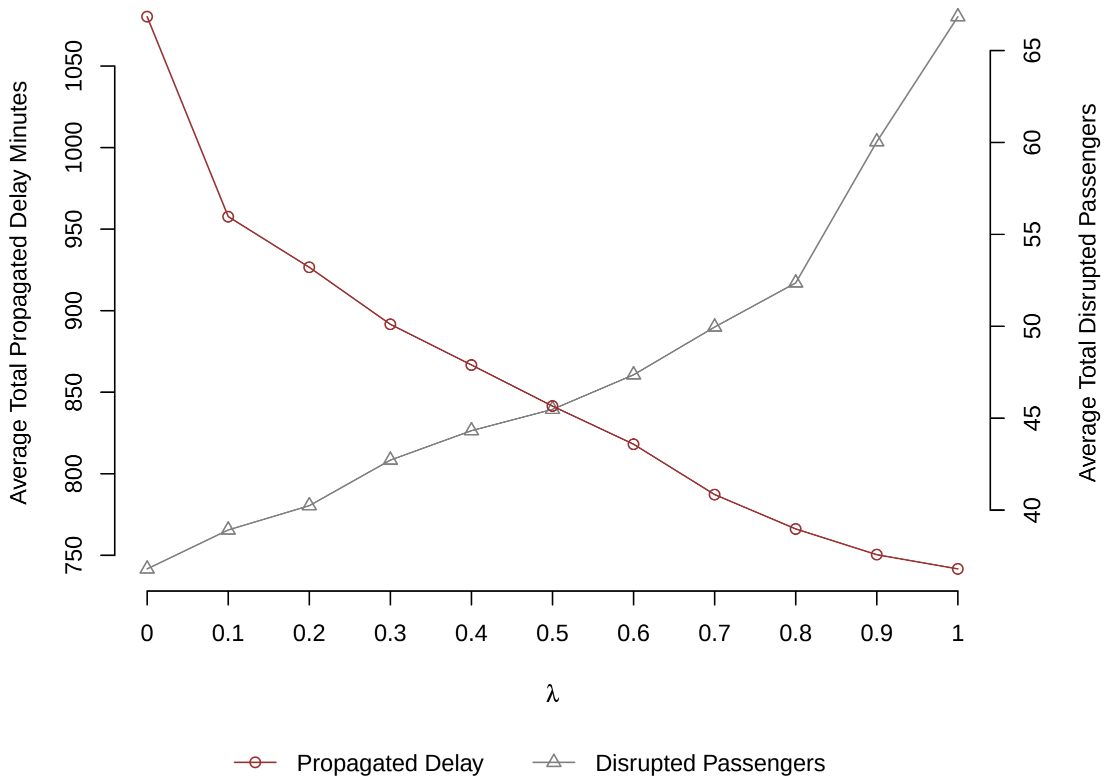


Figure 4-19: Average total propagated delays and average total disrupted passengers for solutions with different values of λ

tions that minimize the total change in the departure times. We denote these solutions with the suffix `_minRetime`.

2) the optimal `FR_maxEffACSlack15` solution that maximizes the total expected effective passenger connection slack, denoted as `FR_maxEffAC+PaxSlack15`. In particular, the primary objective function is to maximize the total expected effective *aircraft* connection slack.

3) the optimal `FR_maxEffPaxSlack15` solution that maximizes the total expected effective aircraft connection slack, denoted as `FR_maxEffPax+ACSlack15`. In particular, the primary objective function is to maximize the total expected effective *passenger* connection slack.

Table 4.12 summarizes the performance of the alternative optimal solutions to the `FR` models. The reductions in the total amount of re-timing range from 10 to more

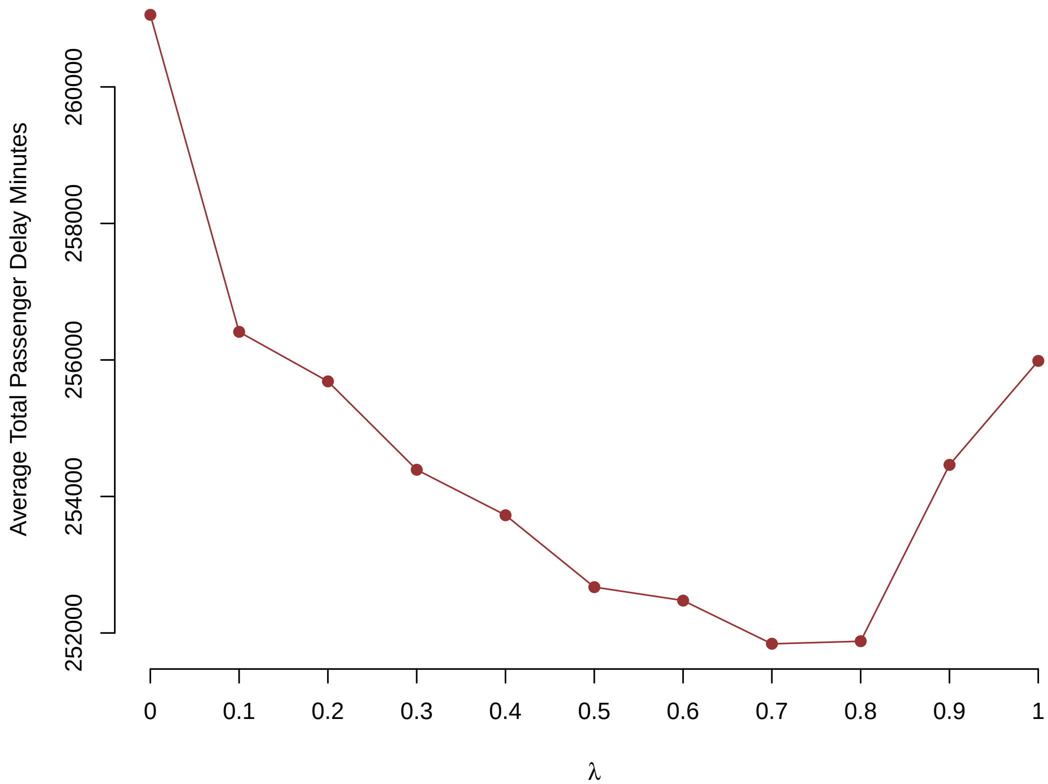


Figure 4-20: Average total passenger delay for solutions with different values of λ

than 200 minutes. In the `FR_minPD_minRetime`, `FR_maxEffACSlack15_minRetime`, and `FR_maxEffAC+PaxSlack15` solutions, the improvements in flight delay statistics drop modestly, while the number of disrupted passengers as well as total passenger delay are decreased. On the other hand, in the `FR_maxEffPaxSlack15_minRetime` and `FR_maxEffPax+ACSlack15` solutions, every performance evaluation metric is improved.

We also observe that the `FR_maxEffAC+PaxSlack15` solution performs somewhere between those of the `FR_hybrid15_0.9` and `FR_hybrid15_1.0` solutions, and similarly, the `FR_maxEffPax+ACSlack15` solution performs somewhere between those of the `FR_hybrid15_0.0` and `FR_hybrid15_0.1` solutions. Therefore, the `FR_maxEffAC+PaxSlack15` and `FR_maxEffPax+ACSlack15` solutions might not be as desirable as those solutions to the hybrid model with a weight λ between 0.1 and 0.9. In fact, it is also computationally less expensive to only solve the hybrid model once. Note that

	FR_minPD_minRetime	FR_maxEffACSlack15 _minRetime	FR_maxEffPaxSlack15 _minRetime
Schedule Statistics			
Total A/C Connection Slack (mins)	6676.76	6676.76	6676.76
Total Re-timing	1166.36	1179.36	1004.72
Flight Delay Statistics			
Total Propagated Delay (mins)	758.24	741.80	1042.20
% of Flights with PD > 0	12.03%	11.27%	19.17%
Total Arrival Delay (mins)	2969.28	2952.84	3184.24
15-min On-Time Performance	78.25%	78.50%	76.35%
60-min On-Time Performance	97.02%	97.04%	96.89%
Passenger Delay Statistics			
Total Pax Delay (mins)	256085	255740	258853
Total Disrupted Pax (pax)	63.12	66.00	36.76

	FR_maxEffACSlack15 _maxEffPaxSlack15	FR_maxEffPaxSlack15 _maxEffACSlack15
Schedule Statistics		
Total A/C Connection Slack (mins)	6676.76	6676.76
Total Re-timing	1212.40	1200.84
Flight Delay Statistics		
Total Propagated Delay (mins)	742.16	1000.52
% of Flights with PD > 0	11.37%	17.97%
Total Arrival Delay (mins)	2952.96	3153.68
15-min On-Time Performance	78.48%	76.55%
60-min On-Time Performance	97.06%	96.93%
Passenger Delay Statistics		
Total Pax Delay (mins)	255635	257229
Total Disrupted Pax (pax)	66.12	36.64

Table 4.12: Average performance evaluation statistics over 25 days (March 1-25, 2008) of the alternative optimal solutions to the `FR` models

the sets of constraints in the hybrid formulation and the formulation of the second-stage problem are exactly the same. In particular, they include the constraints that determine effective slack for both aircraft and passenger connections.

Given the limited improvements in the alternative optimal solutions, we conclude that solving a second-stage problem to obtain a "better" optimal solution with respect to another performance evaluation metric might not be worthwhile for this particular airline.

Solution Quality

Similarly to the analysis for the `AR` models, we compare the solutions to the `FR` models to their corresponding `_expected` and `_perfectInfo` solutions. The performance of these solutions is summarized in Table 4.13

First, we compare the base solutions with their corresponding _expected solutions. We can see that using many delay scenarios from historical data to capture the stochasticity of delays is helpful in improving the performance evaluation metrics that are positively correlated to the objective function. In particular, in the FR_minPD and FR_maxEffACSlack15 solutions, every flight delay statistic is better than in the corresponding _expected solutions, whereas in the FR_maxEffPaxSlack15 solution, there are fewer disrupted passenger, as compared to the FR_maxEffPaxSlack15_expected solution.

However, as we discussed earlier, a solution that minimizes flight delays tends to cause more disrupted passengers, and vice versa. As a result, the _expected solutions perform better than their base solutions with respect to other performance metrics that are not considered in the objective functions. Loosely speaking, the _expected solutions serve as "compromise" solutions. For instance, in the FR_minPD_expected solution, the number of disrupted passengers as well as the total passenger delay is smaller than in the FR_minPD solution, in spite of the increases in the total propagated delay and the total arrival delay. Therefore, it is not clear whether the _expected solutions are inferior to the base solutions.

In all _perfectInfo solutions, every performance evaluation metric is improved. More importantly, the improvements are more significant than in the _perfectInfo solutions to the AR models. This indicates that the FR models can potentially achieve larger improvements. However, such significant improvements can be obtained only when the historical data used as input to the model well reflect the actual delays in the day of operation. As we discussed earlier, despite the flexibility of the FR models in reducing small propagated delays, these FR models are very sensitive to the difference between historical delays and actual delays.

We note that the FR_minPD_perfectInfo solution performs better than the FR_maxEffACSlack15_perfectInfo solution. This is because the model that maximizes the total expected effective aircraft connection slack still tries to increase effective slack in those connections with delays close to propagating, which is not necessary, given the perfect information of the future operation.

Lastly, in the FR_maxEffPaxSlack15_perfectInfo solution, although the number of disrupted passengers is very low, the total propagated delay as well as the total passenger delay is still relatively large. This confirms that the objective function of maximizing the total expected effective passenger connection slack is not be appropriate for this particular airline.

Allowable changes in flight departure times

In the robust flight schedule re-timing models, an allowable change in the flight departure time of each flight is limited to ensure that demand for the flight remains the same. Clearly, increasing the time window width for allowable departure times of each flight will increase the flexibility of the FR model, and allow larger improvements in the resulting schedule.

So far we assume a time window width of 30 minutes (± 15 minutes) for every flight, except for the first and last flights of each flight string that are not allowed to moved earlier or later, respectively. To understand the effects of a time window width on the performance of solutions to the FR models, we solve for the FR_minPD, FR_maxEffACSlack15, and FR_maxEffPaxSlack15 solutions again with time window widths of 20 and 10 minutes. Table 4.14 summarizes the performance of the solutions.

We first look at the performance of the FR_minPD and FR_maxEffACSlack15 solutions. Even for the case of 10-minute time windows, the FR_minPD and FR_maxEffACSlack15 solutions improve, compared to the Original schedule, all but the number of disrupted passengers metric. As the time window width decreases, the total propagated delay in both solutions increases due to the limited flexibility. Limiting the changes to the flight schedule, however, also results in fewer disrupted passengers. It turns out that, despite having the largest total propagated delay, the FR_minPD and FR_maxEffACSlack15 solutions for the 10-minute, compared to the 30- and 20-minute, time window case yield the smallest total passenger delay

In contrast, for the FR_maxEffPaxSlack15 solutions, as the time window width decreases, the number of disrupted passenger increases; while flight delay statistics

	FR_minPD	FR_minPD _expected	FR_minPD _perfectInfo
Schedule Statistics			
Total A/C Connection Slack (mins)	6676.76	6676.76	6676.76
Total Re-timing	1258.80	1327.84	1420.48
Flight Delay Statistics			
Total Propagated Delay (mins)	756.24	850.04	563.76
% of Flights with PD > 0	11.97%	16.21%	6.86%
Total Arrival Delay (mins)	2967.40	3040.76	2793.76
15-min On-Time Performance	78.26%	77.49%	79.85%
60-min On-Time Performance	97.01%	96.98%	97.39%
Passenger Delay Statistics			
Total Pax Delay (mins)	256540	256338	240652
Total Disrupted Pax (pax)	64.52	55.20	51.40

(a) FR_minPD

	FR_maxEffACSlack15	FR_maxEffACSlack15 _expected	FR_maxEffACSlack15 _perfectInfo
Schedule Statistics			
Total A/C Connection Slack (mins)	6676.76	6676.76	6676.76
Total Re-timing	1223.32	1327.24	1355.84
Flight Delay Statistics			
Total Propagated Delay (mins)	741.48	767.88	602.72
% of Flights with PD > 0	11.34%	12.56%	8.08%
Total Arrival Delay (mins)	2952.52	2978.08	2814.52
15-min On-Time Performance	78.48%	78.00%	79.62%
60-min On-Time Performance	97.04%	97.01%	97.36%
Passenger Delay Statistics			
Total Pax Delay (mins)	256211	257583	243619
Total Disrupted Pax (pax)	67.16	66.40	58.16

(b) FR_maxEffACSlack15

	FR_maxEffPaxSlack15	FR_maxEffPaxSlack15 _expected	FR_maxEffPaxSlack15 _perfectInfo
Schedule Statistics			
Total A/C Connection Slack (mins)	6676.76	6676.76	6676.76
Total Re-timing	1230.80	1250.76	1314.28
Flight Delay Statistics			
Total Propagated Delay (mins)	1076.40	1016.76	904.72
% of Flights with PD > 0	20.12%	19.22%	16.64%
Total Arrival Delay (mins)	3219.36	3171.68	3046.68
15-min On-Time Performance	76.08%	76.28%	77.74%
60-min On-Time Performance	96.86%	96.88%	97.14%
Passenger Delay Statistics			
Total Pax Delay (mins)	260854	261073	249827
Total Disrupted Pax (pax)	36.80	43.72	36.00

(c) FR_maxEffPaxSlack15

Table 4.13: Average performance evaluation statistics over 25 days (March 1-25, 2008) of solutions to the FR models with different approaches of using historical data

	FR_minPD	FR_maxEffACSlack15	FR_maxEffPaxSlack15
Schedule Statistics			
Total A/C Connection Slack (mins)	6676.76	6676.76	6676.76
Total Re-timing	965.32	942.12	936.52
Flight Delay Statistics			
Total Propagated Delay (mins)	782.60	782.00	1047.84
% of Flights with PD > 0	12.36%	12.08%	19.19%
Total Arrival Delay (mins)	2984.28	2981.88	3189.60
15-min On-Time Performance	78.33%	78.35%	76.31%
60-min On-Time Performance	96.98%	97.01%	96.98%
Passenger Delay Statistics			
Total Pax Delay (mins)	256845	257242	259476
Total Disrupted Pax (pax)	60.28	61.84	38.20

(a) Time Window = 20 minutes (± 10 minutes)

	FR_minPD	FR_maxEffACSlack15	FR_maxEffPaxSlack15
Schedule Statistics			
Total A/C Connection Slack (mins)	6676.76	6676.76	6676.76
Total Re-timing	557.16	555.68	532.44
Flight Delay Statistics			
Total Propagated Delay (mins)	854.08	854.76	1020.52
% of Flights with PD > 0	14.03%	13.93%	18.21%
Total Arrival Delay (mins)	3030.84	3029.60	3158.44
15-min On-Time Performance	77.82%	77.74%	76.63%
60-min On-Time Performance	96.91%	96.98%	96.88%
Passenger Delay Statistics			
Total Pax Delay (mins)	255965	255348	259003
Total Disrupted Pax (pax)	53.00	52.52	40.64

(b) Time Window = 10 minutes (± 5 minutes)

Table 4.14: Average performance evaluation statistics over 25 days (March 1-25, 2008) for the FR models with different time window widths

are improved. Again, it turns out that the FR_maxEffPaxSlack15 solution for the 10-minute time window case yields the smallest total passenger delay, in spite of the largest number of disrupted passengers.

In summary, similar to the conclusion for the performance of _expected solutions discussed earlier, a solution for the case where a time window is small serves as a "compromise" solution. In particular, the improvements of performance evaluation metrics that are positively correlated with the objective function are limited, while the adverse effects of the solution on other metrics are moderated.

Alternatively, we fix the time window width at 30 minutes as before, but we allow to move the departure times of the first and last flight of each string at most 5 minutes

	Original	FR_minPD		FR_maxEffACSlack15		FR_maxEffPaxSlack15	
Schedule Statistics							
Total A/C Connection Slack (mins)	6676.76	7193.40	7.74%	7184.92	7.61%	6877.76	3.01%
Total Re-timing	0.00	1867.68		1819.36		2018.08	
Flight Delay Statistics							
Total Propagated Delay (mins)	1009.60	642.84	-36.33%	636.28	-36.98%	1042.84	3.29%
% of Flights with PD > 0	17.74%	9.65%		9.04%		19.72%	
Total Arrival Delay (mins)	3141.16	2871.40	-8.59%	2862.04	-8.89%	3187.72	1.48%
15-min On-Time Performance	76.53%	78.84%		79.16%		76.20%	
60-min On-Time Performance	96.89%	97.26%		97.26%		96.99%	
Passenger Delay Statistics							
Total Pax Delay (mins)	260565	247537	-5.00%	247828	-4.89%	256195	-1.68%
Total Disrupted Pax (pax)	47.56	57.64	21.19%	59.80	25.74%	33.24	-30.11%

Table 4.15: Average performance evaluation statistics over 25 days (March 1-25, 2008) for the FR models for which the first and last flights of each string are allowed to move earlier and later, respectively

earlier or later, respectively. Consequently, the elapsed time between the start and end of flying for each aircraft will increase at most 10 minutes. Table 4.15 summarizes the results.

The results, to begin with, show that the total aircraft connection slack increases almost 8% in the FR_minPD and FR_maxEffACSlack15 solutions and about 3% in the FR_maxEffPaxSlack15 solution. Also, the total amount of re-timing significantly increases in all solutions.

In the FR_minPD and FR_maxEffACSlack15 solutions, the reductions in total propagated delay are as large as 36% percent. These smaller delays also cause fewer disrupted passengers. As a result, the total passenger delays are considerably reduced in both FR_minPD and FR_maxEffACSlack15 solutions. In the FR_maxEffPaxSlack15 solution, the number of disrupted passengers is further decreased, and flight delay statistics are slightly improved, as compared to the case where the first and last flights of each string are not allowed to move earlier and later, respectively.

These resulting improvements suggest that by allowing each aircraft to operate a little longer on each day of operation, an airline can significantly improve its schedule performance.

4.4 Robust Block Time Adjusting Model

In this section, we present the computational results obtained from the robust block-time adjustment model introduced in Section 3.3, with different objective functions and parameters. Again, let `Original` denote the airline's original schedule, and define the other solutions as follows.

`BA_minTAD`	the solution to the `BA` model that minimizes the total expected arrival delay (see (BA-1))
`BA_minPD`	the solution to the `BA` model that minimizes the total expected propagated delay
`BA_maxEffACSlack<Γ>`	the solution to the `BA` model that maximizes the total expected effective *aircraft connection* slack with caps set equal to Γ minutes for every aircraft connection
`BA_maxEffPaxSlack<Γ>`	the solution to the `BA` model that maximizes the total expected effective *passenger connection* slack with caps set equal to Γ minutes for every passenger connection

Recall that, in the `BA` models, we need to specify time windows $[l_{x_i}, u_{x_i}]$ and $[l_{y_i}, u_{y_i}]$ within which the departure and arrival times of flight i are allowed to change, and a time window $[l_i, u_i]$ within which the total block time change of flight i, $y_i - x_i$ is allowed. Throughout this section, we assume the following, unless stated otherwise. For a flight leg i,

- if i is the first flight of some flight string, then $[l_{x_i}, u_{x_i}] = [0, 15]$ and $[l_{y_i}, u_{y_i}] = [-15, 15]$;
- if i is the last flight of some flight string, then $[l_{x_i}, u_{x_i}] = [-15, 15]$ and $[l_{y_i}, u_{y_i}] = [-15, 0]$;
- otherwise, $[l_{x_i}, u_{x_i}] = [l_{y_i}, u_{y_i}] = [-15, 15]$.

	`Original`	`BA_minTAD`		`BA_maxEffPaxSlack15`	
Schedule Statistics					
Total A/C Connection Slack (mins)	6676.76	4122.60	-38.25%	6654.80	-0.33%
Total Absolute Block Time Change (mins)	-	2627.12		2929.56	
Average Block Time Change (mins)		10.55		0.09	
Flight Delay Statistics					
Total Propagated Delay (mins)	1009.60	827.76	-18.01%	1174.84	+16.37%
% of Flights with PD > 0	17.74%	14.18%		20.79%	
Total Arrival Delay (mins)	3141.16	1873.48	-40.36%	4068.60	+29.53%
15-min On-Time Performance	76.53%	87.49%		67.07%	
60-min On-Time Performance	96.89%	97.69%		96.35%	
Passenger Delay Statistics					
Total Pax Delay (mins)	260565	178004	-31.69%	313990	+20.50%
Total Disrupted Pax (pax)	47.56	62.12	+30.61%	28.52	-40.03%

Table 4.16: Average performance evaluation statistics over 25 days (March 1-25, 2008) for the `BA` models

Also, we set $[l_i, u_i] = [-15, 15]$ for every flight. In other words, each flight's departure and arrival times are allowed to move at most 15 minutes earlier or later, and the maximum total change in block time is 15 minutes. For each flight string, the departure time of the first flight and the arrival time of the last flight are not allowed to move earlier and later, respectively.

4.4.1 Computational Results

The performances of the `BA_minTAD` and `BA_maxEffPaxSlack15` solutions over the period of March 1-25, 2008 are summarized in Table 4.16.

Schedule Statistics

As discussed in Section 2.4, the block time adjustment problem allows ground time slack to be transformed into block time slack. In the `BA_minTAD` solution, the total aircraft connection slack is decreased from the `Original` schedule by almost 40%.

To quantify the difference between the `Original` schedule and solutions to the `BA` models, we report "Total Absolute Block Time Change", as defined in Section 3.3.4. Because this metric does not distinguish block time reduction from block time increase, we also report "Average Block Time Change" to indicate overall direction of change in block times. Consistently with the decrease in aircraft connection slack, flight block times in the BA_minTAD solution increase about 10 minutes, on average. As we will discuss shortly, this is, in fact, the key to significant reduction in total arrival delay. To illustrate the extent of the difference, Figure 4-21 depicts the BA_minTAD solution for March 1, 2008.

In the BA_maxEffPaxSlack15 solution, although the total block time change is larger than in the BA_minTAD solution, the total amount of aircraft connection slack remains almost the same as in the Original schedule, and the average block time change suggests that the changes in block times occur equally in both directions.

Flight Delay Analysis

Similarly to the FR_maxEffPaxSlack15 solution, the BA_maxEffPaxSlack15 solution performs worse than the Original schedule in every flight delay statistic because the objective of maximizing the total expected effective passenger connection slack has no direct link to flight delay improvements.

Remarkably, the total arrival delay is reduced by more than 40 % in the BA_minTAD solution. Recall that even in the _perfectInfo solutions to AR and FR models, the total arrival delay is about 1,000 minutes larger than in the BA_minTAD solution. Additionally, other flight delay statistics are improved significantly, especially the 15-minute on-time performance metric. In fact, Figure 4-21, illustrating the expected actual operations of the BA_minTAD solution on March 1, 2008, shows that many flights now arrive earlier than the scheduled arrival times. Table 4.17 summarizes the total arrival delay distribution for the BA_minTAD solution. Total arrival delays are significantly reduced in every positive range, compared to distributions of the AR and FR solutions.

Total Arrival Delay (mins)	0	(0,15]	(15,60]	(60,120]	>120
Original (%)	44.33%	32.20%	20.36%	2.17%	0.94%
BA_minTAD (%)	69.90%	17.59%	10.20%	1.59%	0.73%

Table 4.17: Distributions of total arrival delays for the BA models

Why are solutions to the AR and FR models unable to reduce effectively flight

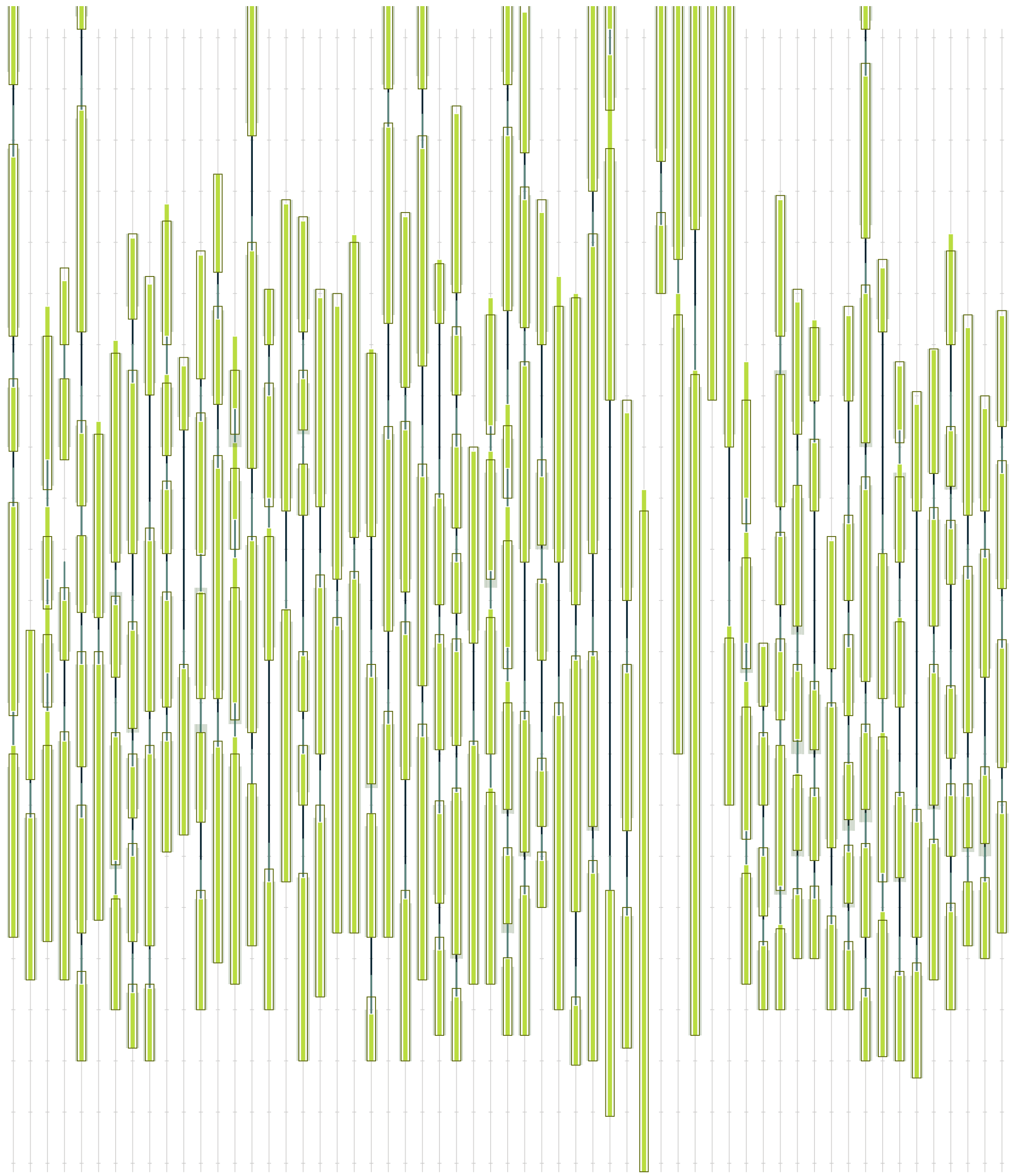


Figure 4-21: The BA_minTAD solution for March 1, 2008

arrival delays, relatively to the BA_minTAD solution? To begin with, recall that a total arrival delay is a function of the propagated delay from a preceding flight and the independent arrival delay of that flight (see Section 2.2.2). Because independent arrival delays are fixed in the AR and FR models, the total arrival delay minimization can only be achieved by minimizing the total propagated delay. On the other hand, the block time adjustment model allows aircraft connection slack to be converted into block time slack, and thus can potentially reduce independent arrival delays in addition to propagated delays.

If the contribution of independent arrival delays to the total arrival delay were to be small for this particular airline, the additional reduction of total arrival delay in the BA_minTAD solution would not be this significant. According to the performance of the Original schedule, propagated delays contribute only about one third of total arrival delay, and another two thirds is due to independent arrival delays of each flight. Moreover, the historical data indicate that planned block times are underestimated, on average, by almost 10 minutes. Over 50% of the flights flew longer than their planned block times more than 70% of the time. Figure 4-22 depicts the actual block time distributions of some problematic flights. A vertical bar in each plot denotes the planned block time for that flight. We can see that some flights always flew longer than their planned block times. Consequently, even though these flights experience no propagated delays, arrival delays are inevitable.

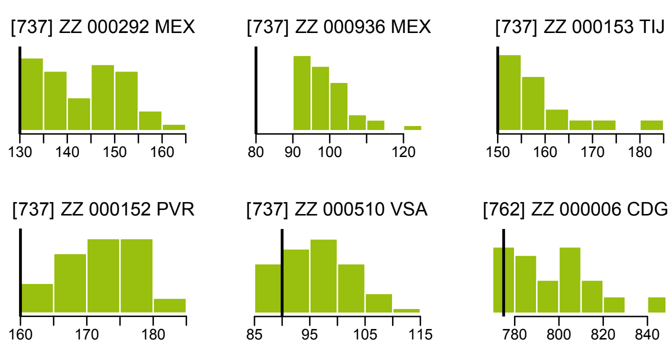


Figure 4-22: Actual block time distributions

To summarize, because the major contribution to the total arrival delay, for this particular airline, is due to independent arrival delays, the BA_minTAD solution can effectively use additional block time slack, converted from original ground time slack, to absorb independent arrival delays and achieve the very small total arrival delay.

Passenger Delay Analysis

As a result of the significant total arrival delay reduction, the total passenger delay is considerably decreased in the `BA_minTAD` solution. The number of disrupted passengers, however, increases by about 30%. Intuitively, because block times for each flight in the `BA_minTAD` solution are increased by 10 minutes on average, most of the passenger connection times become shorter. Consequently, connecting passengers are more likely to miss their connections.

In the `BA_maxEffPaxSlack15` solution, the number of disrupted passengers is reduced by 40%. Note that this reduction is larger than that achieved by the `FR_maxEffPaxSlack15_perfectInfo` solution. Similarly to the `FR_maxEffPaxSlack15` solution, the disrupted passenger delay reduction in `BA_maxEffPaxSlack15` still cannot make up for the increase in flight delays, and the total passenger delay is much larger than in the `Original` schedule.

4.4.2 Discussion on Models

Alternative Objectives

As discussed in Section 3.3.3, minimizing the total expected propagated delay or maximizing the total expected effective aircraft connection slack in the `BA` formulation is, in fact, not a good proxy for minimizing the total expected arrival delay. To illustrate the issue, Table 4.18 summarizes the performance of the `BA_minPD` and `BA_maxEffACSlack15` solutions.

The results show that the total propagated delay is significantly reduced in both solutions, compared to the corresponding `AR` and `FR` solutions. Again, the larger improvement in the `BA_maxEffACSlack15` solution confirms that maximizing the total expected effective aircraft slack is a better proxy for minimizing total propagated delay. Both solutions, however, result in considerably higher total arrival delay, which

	BA_minPD	BA_maxEffACSlack15
Schedule Statistics		
Total A/C Connection Slack (mins)	8303.40	7025.64
Total Absolute Block Time Change (mins)	2774.88	1876.00
Average Block Time Change (mins)	-6.72	-1.44
Flight Delay Statistics		
Total Propagated Delay (mins)	606.40	588.80
% of Flights with PD > 0	8.89%	8.17%
Total Arrival Delay (mins)	4705.08	3530.04
15-min On-Time Performance	63.31%	74.68%
60-min On-Time Performance	96.31%	96.73%
Passenger Delay Statistics		
Total Pax Delay (mins)	365306	294151
Total Disrupted Pax (pax)	71.20	73.96

Table 4.18: Average performance evaluation statistics over 25 days (March 1-25, 2008) for the BA models with alternative objectives

consequently leads to much larger total passenger delay. Additionally, having large total arrival delays while most flights can depart on time also causes more disrupted passengers. These results are consistent with our discussion in Section 3.3.3.

The schedule statistics indicate that block time of each flight is decreased, on average, by 6.7 and 1.4 minutes in the BA_minPD and BA_maxEffACSlack15 solutions, respectively, and the total aircraft connection slack increases in both solutions. These results are opposite to the BA_minTAD solution for which most of aircraft connection slack is transformed into block time slack. This difference reflects two common approaches to building robustness into airline schedules: 1) schedule padding (i.e., increasing block time slack); and 2) having large turn around time (i.e., increasing aircraft connection slack).

As discussed in Section 2.4, block time slack provides greater flexibility than aircraft connection slack because it can absorb not only propagated delay from prior flights, but also independent departure and arrival delay (such as delays due to Ground Delay Programs, taxi delays, and airborne delays); while aircraft connection slack can absorb only propagated delay from the preceding flight. However, it is generally more costly to add slack into block times (in particular because crew productivity is reduced and hence, crew costs are increased), and schedule padding may not be an appropriate approach for every airline. Interested readers are referred to Zhu (2009) [34]. In her work, she provides a comprehensive comparison of the

	BA_minTAD	BA_minTAD _expected	BA_minTAD _perfectInfo
Schedule Statistics			
Total A/C Connection Slack (mins)	4122.60	4756.52	5175.28
Total Absolute Block Time Change (mins)	2627.12	2039.52	1753.72
Average Block Time Change (mins)	10.55	7.93	6.20
Flight Delay Statistics			
Total Propagated Delay (mins)	827.76	911.84	629.00
% of Flights with PD > 0	14.18%	17.90%	9.14%
Total Arrival Delay (mins)	1873.48	2112.56	1631.12
15-min On-Time Performance	87.49%	85.12%	89.83%
60-min On-Time Performance	97.69%	97.42%	97.97%
Passenger Delay Statistics			
Total Pax Delay (mins)	178004	196815	159013
Total Disrupted Pax (pax)	62.12	70.12	52.64

Table 4.19: Average performance evaluation statistics over 25 days (March 1-25, 2008) of `BA_minTAD` solutions with different approaches of using historical data

performances of two airlines adopting these two different slack allocation approaches.

Solution Quality

As before, we exhibit the quality of the `BA_minTAD` solution by comparing with its corresponding `_expected` and `_perfectInfo` solutions. The performance of these solutions is summarized in Table 4.19.

According to Table 4.19, the `BA_minTAD` solution performs reasonably better than the `BA_minTAD_expected` solution with respect to every performance evaluation metric. This again demonstrates the benefits of using many different delay scenarios from historical data to capture the stochasticity of delays. Nevertheless, the `BA_minTAD_expected` solution, obtained by simply using *average* independent arrival delays of each flight, still results in a 32% reduction in the total arrival delay and a 24% reduction in the total passenger delay.

The performance of the `BA_minTAD_perfectInfo` solution indicates that, given perfect information, total arrival delay can be reduced by almost 50%, and total passenger delay can be reduced by almost 40%. The performance gap between the `BA_minTAD` and `BA_minTAD_perfectInfo` solutions with respect to total arrival delay and total passenger delay is about 10%.

Multiple Optimal Solutions

Typically, there are multiple optimal BA_minTAD solutions, as discussed in Section 3.3.4. In this section, we solve the robust block time adjustment problem again to select, among the optimal BA_minTAD solutions, the solution that

- minimizes the total expected propagated delay (BA_minTAD_minPD);
- maximizes the total expected effective aircraft connection slack with caps set equal to 15 minutes for every aircraft connection (BA_minTAD_maxEffACSlack15);
- maximizes the total expected effective passenger connection slack with caps set equal to 15 minutes for every passenger connection (BA_minTAD_maxEff-PaxSlack15); and
- minimizes the difference between the original and the optimal flight schedule (BA_minTAD_minBTChange).

The performances of these solutions are summarized in Table 4.20. Although all solutions minimize the expected total arrival delay, their evaluated performances are slightly different, depending on the secondary objectives. In particular, the total arrival delay modestly increases in every solution. The BA_minTAD_minPD and BA_minTAD_maxEffACSlack15 solutions have slightly less total propagated delay. The BA_minTAD_maxEffPaxSlack15 solution results in fewer disrupted passengers. Lastly, the BA_minTAD_minBTChange solution requires 1,000 minutes of total block time change less than the BA_minTAD solution. Additionally, as a result of the increases in total arrival delay, the total passenger delay is larger in every solution.

Given the limited improvements in the alternative optimal solutions, we conclude that, for this particular airline, solving a second-stage problem to obtain a "better" optimal solution with respect to another performance evaluation metric might not be worthwhile.

	BA_minTAD _minPD	BA_minTAD _maxEffACSlack15	BA_minTAD _maxEffPaxSlack	BA_minTAD _MinBTChange
Schedule Statistics				
Total A/C Connection Slack (mins)	4203.64	4188.64	4152.28	4221.00
Total Absolute Block Time Change (mins)	2553.12	2567.96	2597.04	2526.32
Average Block Time Change (mins)	10.22	10.28	10.43	10.15
Flight Delay Statistics				
Total Propagated Delay (mins)	817.40	815.40	828.60	820.08
% of Flights with PD > 0	14.15%	14.10%	14.30%	14.20%
Total Arrival Delay (mins)	1902.84	1894.12	1884.76	1905.44
15-min On-Time Performance	87.19%	87.26%	87.44%	87.14%
60-min On-Time Performance	97.67%	97.69%	97.67%	97.67%
Passenger Delay Statistics				
Total Pax Delay (mins)	179667	179198	178425	179249
Total Disrupted Pax (pax)	62.44	62.60	60.88	61.28

Table 4.20: Average performance evaluation statistics over 25 days (March 1-25, 2008) of the alternative optimal solutions to the BA models

Allowable changes in flight schedules

In the BA formulation presented in Section 3.3.2, there are six parameters– l_{x_i}, u_{x_i}, l_{y_i}, u_{y_i}, l_i, and u_i that limit the allowable changes in the departure time, arrival time, and block time of a given flight i. So far we assume a time window width of 30 minutes (± 15 minutes). In particular, for every flight, we set $l_{x_i} = l_{y_i} = l_i = -15$ and $u_{x_i} = u_{y_i} = u_i = 15$, except for the first and last flights of each string where we set $l_{x_i} = 0$ and $u_{y_i} = 0$, respectively.

To demonstrate the effect of a time window width on the performance of solutions, we solve for the BA_minTAD solutions again with time window widths of 20 and 10 minutes. Table 4.21 summarizes the performance of the solutions.

As the time window decreases, the flexibility of the model is limited, and thus the total arrival delay reduction decreases from 40% in the 30-minute time window case to 33% and 22% in the 20- and 10-minute cases, respectively. Despite the smaller reduction in the total arrival delay, the BA_minTAD solution for the 10-minute time window case still performs better, with respect to total arrival delay and passenger delay, than any AR and FR solutions presented in the previous sections. In summary, these remarkable improvements suggests that by increasing the block times less than 5 minutes on average, an airline can significantly improve its schedule performance.

	Time Window = ± 15	Time Window = ± 10	Time Window = ± 5
Schedule Statistics			
Total A/C Connection Slack (mins)	4122.60	4747.64	5616.48
Total Absolute Block Time Change (mins)	2627.12	1972.08	1079.72
Average Block Time Change (mins)	10.55	7.97	4.38
Flight Delay Statistics			
Total Propagated Delay (mins)	827.76	874.48	939.24
% of Flights with PD > 0	14.18%	14.78%	16.07%
Total Arrival Delay (mins)	1873.48	2100.88	2513.08
15-min On-Time Performance	87.49%	85.69%	82.51%
60-min On-Time Performance	97.69%	97.52%	97.27%
Passenger Delay Statistics			
Total Pax Delay (mins)	178004	192788	219148
Total Disrupted Pax (pax)	62.12	55.24	49.88

Table 4.21: Average performance evaluation statistics over 25 days (March 1-25, 2008) for the `BA` models with different time window widths

Chapter 5

Summary and Future Work

5.1 Summary

A myriad of uncontrollable factors in airline operations make delays and disruptions unavoidable. The impact of delays is exacerbated when they propagate to subsequent flights through an airline's interconnected network. In the past years, airlines have spent billions of dollars of operating costs incurred due to delays and disruptions. Most conventional scheduling models ignore the presence of uncertainties in actual operations in order to limit the complexity of the problem. This results in schedules that are vulnerable to disruptions. To overcome this shortcoming, there has been wide interest recently in building robustness into airline schedules, i.e., proactively making them more resilient to delays and disruptions.

The key challenge of robust schedule planning is to define *robustness* of a schedule such that it well reflects desired characteristics and can be captured in a tractable mathematical model. In Chapter 2, we review robust airline schedule planning approaches proposed in the literature. Apparently, there is no single consensus definition of *robustness* in the context of airline schedule planning. Additionally, one definition of *robustness* may lead to various mathematical *models*, often using different *proxies* to capture the robustness objectives and ensure model tractability. One of the most critical shortcomings in many works in the literature is that the performance of the resulting schedule is evaluated based primarily on the objective function values, rather than various performance evaluation metrics. Ignoring the trade-offs among different evaluation metrics makes the benefits of these *robust* schedule unclear.

In this thesis, we investigate slack allocation approaches for robust airline scheduling. Airlines have made numerous efforts to increase the utilization of all resources in their operations, often resulting in the minimization of schedule slack. Slack, however, is desirable in robust schedules as it can potentially absorb delays in an airline network. Therefore, we seek to re-allocate, rather than simply increase, the existing slack in the schedules such that the resulting distribution of slack is more effective in absorbing delays and minimizing disruptions. An example illustrating how we can strategically re-allocate slack in a schedule is provided at the end of Chapter 2.

In Chapter 3, we present a modeling framework for robust slack allocation in airline schedule planning. In particular, we propose three *models*: the robust aircraft re-routing model (`AR`), the robust fight schedule re-timing model (`FR`), and the robust block time adjustment model (`BA`), together with their variants. Different *proxies* are used as objective functions for each model. Importantly, we introduce a novel notion of *effective slack*, which is proved to serve as a good robustness proxy in many cases.

Using the data from an international carrier, we present proof-of-concept results in Chapter 4. We evaluate the impacts of the resulting schedules on various performance metrics, including passenger delays and delay propagation. The results show that minor modifications to an original schedule can significantly improve the overall performance of the schedule. Through empirical results, we demonstrate trade-offs between different performance metrics and provide a comprehensive discussion of model behaviors and how different characteristics of an airline can affect the strategy for robust scheduling. Our results are summarized in the next section.

5.1.1 Airline Strategy for Robust Schedule Planning

In the robust slack allocation framework presented in this work, there are different models and objective functions that an airline can mix and match to construct a robust schedule. As we discussed throughout this work, an airline's goals and characteristics are critical in determining which model and objective function are more appropriate and beneficial to the airline.

Models

In this framework, we propose three models, namely, the robust aircraft re-routing model (`AR`), the robust flight schedule re-timing model (`FR`), and the robust block time adjustment model (`BA`). The following are some criteria that may affect an airline's decision regarding which model is most appropriate.

Flexibility

The flexibility of the `AR` model is limited in the sense that it can affect only aircraft connection slack. Additionally, the extent of improvements in the resulting schedule hinges on the available re-routing opportunities, which depend largely on the airline's network structure. For example, an airline with a strong hub-and-spoke structure might find no re-routing opportunities for flights departing from spokes. Nevertheless, given the aircraft re-routing opportunities, an `AR` model is capable of removing large delays, which cannot be done in `FR` or `BA` models.

An `FR` model is more flexible than an `AR` model because it allows finer slack re-allocations. In particular, slack can be increased or decreased by any amount within the allowable time window. Moreover, it can also affect passenger connection slack. Therefore, it can be used together with a passenger-centric objective function to reduce passenger delays and misconnections.

Lastly, a `BA` model provides the greatest flexibility. Like an `FR` model, it allows finer re-allocation of slack, compared to an `AR` model. More importantly, ground time slack can be converted into block time slack in this model. Block time slack can absorb not only propagated delay from prior flights, but also independent departure and arrival delays (such as delays due to Ground Delay Programs, taxi delays, and airborne delays); while aircraft connection slack can absorb only propagated delays from the preceding flights. Therefore, this model is very useful for airlines that are facing large independent delays in their systems. One indicator of a need for this model is the contribution of independent arrival delays to total arrival delay. If it is large, then a BA model is the only model in this framework that is capable of reducing these independent delays.

Cost

Robustness in an airline schedule typically is achieved at a cost. Even in an AR model for which the flight schedule is fixed, changing aircraft routing can possibly affect crew duties. For instance, more crews might need to transfer between aircraft and require larger connection times. Consequently, crew costs might be slightly increased.

With the multi-faceted work rules such as minimum rest time, maximum flying time, and maximum duty period, crew duties are more affected when a flight schedule is changed in the FR or BA solutions. As a result, crew costs in FR and BA solutions may increase more than in an AR solution. Additionally, the schedule changes might also result in uneven utilization of airline resources and personnel such as gate agents, and ground and maintenance crews. These increasing planned costs, however, are limited by the restrictions imposed in our model allowing flight schedule changes within only a small time window.

Because crews are paid at least for the scheduled block time, block time slack is generally considered more costly than ground time slack. Additionally, longer block times can result in reductions in the number of flights that one crew can operate, or reductions in the number of possible crew connections, all resulting in reduced crew productivity.

Implementation

The AR formulation presented in this work focuses mainly on *re-routing* aircraft on a given day of operation, assuming that the resulting aircraft routes do not affect maintenance feasibility. Therefore, it is more applicable to use for short-term planning– before a day of operation, or use as part of a recovery tool where historical data are replaced by actual delays or expected delays that reflect the current delay situation. Nevertheless, the insights gained from this work can be extended to a general robust aircraft maintenance routing problem.

Unlike an `AR` model, the `FR` and `BA` models result in flight schedule changes. An airline's schedule is very critical to its competitive position and profitability. A schedule development process requires collaboration between many business units to resolve all tactical and operational issues that may arise from the resulting schedule. Therefore, the solutions to the `FR` and `BA` models can serve as guidelines for schedule changes, and an airline can use its decision support tools to analyze the impact of the changes and fine tune accordingly. Additionally, `FR` and `BA` models can be used as a part of recovery tool to provide airline operations controllers a "good" option to adjust scheduled operations. Also, an airline can specify allowable block time changes in a `BA` model according to possible aircraft speeds to identify some potential savings from reduced fuel burn.

Objective Functions

In Chapter 3 and 4, we introduce many objective functions that can be used with the slack re-allocation models provided in this framework. Ultimately, the decision of which objective function to use depends on the airline's goal. Some performance metrics that an airline might want to improve are total arrival delay minutes; 15-minute on-time performance rate; total propagated delay minutes; total passenger delay minutes; and total number of disrupted passengers. The following are some criteria that might affect an airline's objective function selection.

Passenger Delay

Total passenger delay is comprised of delays associated with non-disrupted and disrupted passengers. Although non-disrupted passengers contribute only their last flight's delays to the total passenger delay, almost all passengers are typically non-disrupted. On the other hand, a small number of disrupted passengers can possibly account for a large proportion of total passenger delay, because each of disrupted passenger likely has to wait for several hours, if not a full 24 hours, for the next available flights.

Because disrupted passenger delay calculation is complicated, we cannot directly minimize total passenger delay in our objective functions. Two possible proxies for minimizing total passenger delay are 1) to minimize some flight delay metric; or 2) to minimize passenger disruptions. However, the proof-of-concept results for the `FR` and `BA` models show that the solution that minimizes flight delay tends to result in more disrupted passengers, and vice versa. Therefore, in order to minimize total passenger delay, an airline needs to understand its characteristics and select an appropriate objective function. The percentage of connecting passengers and the percentage of total passenger delay due to disrupted passengers are two key statistics that are significant in this regard. For the airline we consider in this work, its total passenger delay is mainly driven by flight delays, rather than passenger misconnections, and thus minimizing flight delay is a more effective strategy in minimizing total passenger delay.

In the discussion of the `FR` models, we also propose the hybrid objective functions balancing total expected effective slack in aircraft connections and passenger connections. This objective can be use to balance the focus on minimizing flight delays and passenger disruptions by adjusting a weight to reflect an airline's characteristics.

Propagated Delays versus Arrival Delays

Recently, much work on robust schedule planning have focused primarily on reducing propagated delays, rather than arrival delays, and the total propagated delay metric has been used as a measure of the degree of "disruptions" in airline operations. Propagated delays and arrival delays are closely related. Small propagated delays generally lead to small arrival delays, and vice versa. However, there are subtle differences between propagated delays and arrival delays, as we discussed throughout this thesis.

In the `AR` and `FR` models where independent arrival delays are fixed, minimizing propagated delay is typically a good proxy for minimizing total arrival delay. In particular, decreasing a propagated delay from flight i to flight j, pd_{ij}, will result in less total arrival delay for flight j, *as long as the term* $pd_{ij} + IAD_j$ *remains positive.* The term $pd_{ij} + IAD_j$ becomes negative only when IAD_j takes a negative value, that is, when an airline pads its schedule to account for potential delays or can potentially fly the flight faster. In this case, further decreasing a propagated delay only makes the flight arrive before the scheduled arrival time and does not affect the total arrival delay. In fact, it might be desirable to *propagate* delays to those flights with negative IADs and let their block time slack help absorb the delays.

On the contrary, in a `BA` model where independent arrival delays can be altered, minimizing total arrival delay and minimizing total propagated delay lead to two different approaches to building robustness into airline schedules: 1) schedule padding (i.e., increasing block time slack); and 2) having large turn around time (i.e., increasing aircraft connection slack). In the first approach, block time slack is used extensively to absorb all kinds of delay, including propagated delay; while in the second approach, ground time slack is used extensively to absorb arrival delays to ensure that follow-on flights can depart on time.

In general, the objective of minimizing total arrival delay is more preferable because passengers are not bothered by late departures if arrivals at their destinations are on time. Additionally, on-time arrivals are beneficial to passengers and crews that need to make connections.

Effective Slack

In Chapter 3, we introduce a notion of *effective slack* and purpose an objective of maximizing the total expected effective slack to overcome the shortcomings of the objective of minimizing the total expected propagated delay. For each aircraft connection, we can specify a cap, for which effective slack in that connection can contribute to the objective value, to ensure that the model has no incentive to add more slack to connections that already have a reasonable amount of slack. This provides airlines flexibility to set cap values to reflect how much they are willing to protect against unexpected delay.

For example, when caps are set equal to zero for all aircraft connections, the model will focus mainly on the connections that have negative effective slack (or, equivalently, positive propagated delay), as indicated by the historical data, and ignore other connections for which effective slack is close to zero, that is, the delays from preceding flights are close to propagating.

The proof-of-concept results suggest that the objective of maximizing the total expected effective slack with appropriate caps is, regardless of the model used, more robust and yields smaller total propagated delay than minimizing the total expected propagated delay directly.

Moreover, we extend the notion of *effective slack* to passenger connections, as we did in the `FR` and `BA` models. The proof-of-concept results indicate that maximizing the total expected effective passenger connection slack is an effective proxy for minimizing the number of passenger misconnections.

Multiple Optimal Solutions

Typically, there are multiple optimal solutions to our robust slack re-allocation models. Although these optimal solutions give the same objection function value, they might not be equally effective with respect to other performance metrics. One possible approach to obtain an alternative optimal solution is to solve the model sequentially using different objective functions in order to select, among the optimal solutions to the initial objective, a solution that is optimal with respect to other objectives.

However, for a given airline, the number of optimal solutions increases with

- the extent of delays in the given set of historical data, which is partly impacted by the set of airports serviced by the airline;
- re-routing opportunities in the airline schedule, which are impacted by fleet homogeneity, network structure, connecting banks at hubs, etc.; and
- re-timing or block time adjustment opportunities in the airline schedule, which are impacted by allowable time windows for re-scheduling, the number of passenger connections, planned aircraft connection times in the given aircraft routes,

etc.

Therefore, given the same model and objective function, different airlines may obtain different numbers of optimal solutions. A decision to solve for an alternative optimal solution depends largely on the potential improvement, and hence the airline's underlying characteristics.

5.2 Future Work

The following are some possible avenues for future research extending the work presented in this thesis:

1. *Enforcing maintenance feasibility in the robust aircraft re-routing model*

In the `AR` model, we focus on re-routing aircraft on a given day of operation, assuming that the resulting aircraft routes do not violate maintenance feasibility. Fixing the number of aircraft departing from each airport at the beginning of the day and arriving at each airport at the end of the day can be quite restrictive, and the potential delay reduction is limited. Additionally, the assumption on maintenance feasibility might not be valid if we solve the `AR` model over a longer horizon. Therefore, it is of interest to extend the insights gained from this work to a general aircraft maintenance routing problem that explicitly considers aircraft maintenance requirements.

For instance, for a cyclic schedule, one can use the classical flight string model for aircraft maintenance routing proposed by Barnhart et al. (1998) [5] in conjunction with the objective function of maximizing total expected effective aircraft connection slack. For a dated schedule, Gronkvist (2005) [19] provides a formulation for a dated version of the aircraft maintenance routing problem, called the *tail assignment problem*, for which we can incorporate the robust objective functions presented in this work as well.

Alternatively, one can develop a simple heuristic based on the `AR` model that avoids violation of maintenance requirements. For example, we could just fix those aircraft that must be maintained that day to arrive where planned but allow the others to deviate. This will likely work well in a hub and spoke network.

2. Incorporating additional passenger information

We can further improve the `FR` and `BA` models that maximize the total expected effective passenger connection slack by incorporating additional passenger information. In our models, every minute of effective slack is equally weighted in every passenger connection, but differential weightings could be applied easily.

We could put different weights in the objective function according to the number of connecting passengers on each connection. Because different types of passengers (e.g., business and economy passengers) might cost an airline differently when they miss their connections, we can also put different weights based on the mix of connecting passengers on each connection. Moreover, some passenger connections connect to flights that operate only daily (e.g., most of the international flights), and these disrupted passengers might have to wait for a full twenty-four hours before they are re-accommodated. Others might connect to flights that are operated hourly, and the disrupted passengers are then re-accommodated within an hour or so. That being said, it might be preferable to explicitly incorporate estimated passenger re-accommodation times into the coefficients in the objective function.

3. Comparing different data utilization approaches

Through the empirical results, we have shown that by simply using average independent arrival delays as input to the model, an airline can still significantly improve its schedule performance. Further improvements can be obtained by using many different delay scenarios from historical data to capture the stochasticity of delays. These further improvements, however, come with a larger computational cost. Nonetheless, there are several other aspects of data utilization approaches that might be of interest.

In this work, we focus primarily on the sample average approximation approach, i.e., each day of operation in the historical data represents an instance of delay scenario, and we assume that each delay scenario is equally likely. One drawback of this approach is that it might be computationally expensive to include a large number of data points. In contrast, many works proposed in the literature model delays by fitting historical delay data to some standard distributions. For example, Lan (2003) [23] models total arrival delays using a lognormal distribution; Schaefer et al. (2005) [28] model flight delays and ground delays using gamma, Erlang, and beta distributions, depending on flight durations. This approach requires additional data pre-processing to identify and fit historical data to appropriate distributions. In addition, this approach typically ignores the correlation of flight delays on the same day of operation. Some interesting questions regarding these two approaches include, but are not limited to:

- How many data points are required to obtain a reasonably good solution? Apparently, in order to capture rare delay scenarios, a longer period of historical data is needed.
- Should all data points be weighted equally? The rare events might already be captured in the data set, but they are given equal probability to those events that are more common.
- Can standard distributions well capture the actual delay distributions? For example, what if a delay distribution is bimodal?
- Because eventually only expected values or sample averages are used in the objective function, do these two approaches result in significantly different schedules?

The answers to the questions above will partly depend on the sensitivity of the model to the input data. Additionally, they might be different for different airlines, depending on the variability of delays in their networks.

4. Investigating behaviors of the robust schedules when recovery is considered

In the evaluation process that we used to obtain the proof-of-concept results, interventions from airline operations controllers are not allowed. In particular, we assume no flight cancellations, aircraft swaps, early departures, or delaying flights for connecting passengers. To get a more realistic picture of our schedules' performances, it is necessary to simulate the schedules with a recovery module in place because it might be possible that the resulting schedules somehow limit the flexibility to recover in actual operations. This can be done using an advanced simulation tool for airline operations such as MEANS [14] or SimAir [27].

5. Quantifying the cost of schedule changes

Modifications of flight schedules might result in reduced crew productivity or uneven utilization of airline resources and personnel such as gate agents, and ground and maintenance crews. Although crew cost is one of the largest cost components in airline operations, we ignore the potential effects of our models on crew schedules because of lack of information. To make the robust schedules more appealing to airlines, it is crucial to quantify the cost of schedule changes and compare to the savings gained from delay reduction.

As discussed earlier, different robust slack re-allocation models provide different levels of flexibility at different costs. Thus, an ability to quantify the cost of schedule changes is also of great help in justifying which slack re-allocation scheme is the most cost-effective for a particular airline.

6. Repeating a similar analysis to the airlines with different characteristics

Throughout this work, we point out how an airline's characteristics can affect the strategy for robust schedule planning. It would be interesting to see how the proof-of-concept results change when we apply a similar analysis to airlines with different characteristics such as network structure, proportion of connecting passenger, percentage of total passenger delay due to disrupted passengers, and contribution of independent arrival delays to total arrival delay.

7. Multi-criteria optimization

Because of the multi-faceted nature of robustness, there are many trade-offs to be balanced, such as costs of schedule changes versus savings from delay reduction; and passenger delays versus flight delays. It would be very interesting to consider using multi-criteria optimization together with the robust objectives proposed in this work to obtain a schedule that performs relatively well with respect to different metrics.

Bibliography

[1] Ageeva, Y. 2000. Approaches to incorporating robustness into airline scheduling. Master's thesis, Massachusetts Institute of Technology. URL http://hdl.handle.net/1721.1/8930.

[2] AhmadBeygi, S., A. Cohn, M. Lapp. 2008. Decreasing airline delay propagation by Re-Allocating scheduled slack. Submission for Anna Valicek Medal.

[3] Ball, M., C. Barnhart, G. Nemhauser, A. Odoni. 2007. Air transportation: Irregular operations and control. C. Barnhart, G. Laporte, eds., *Handbooks in Operations Research & Management Science: Transportation*, vol. 14. North Holland.

[4] Barnhart, C., P. Belobaba, A. Odoni. 2003. Applications of operations research in the air transport industry. *Transportation Science* **37**(4) 368–391. URL http://transci.journal.informs.org/cgi/content/abstract/37/4/368.

[5] Barnhart, C., N. Boland, L. Clarke, E. Johnson, G. Nemhauser, R. Shenoi. 1998. Flight string models for aircraft fleeting and routing. *Transportation Science* **32**(3) 208–220. URL http://transci.journal.informs.org/cgi/content/abstract/32/3/208.

[6] Belobaba, P., A. Odoni, C. Barnhart. 2009. *The Global Airline Industry*. Wiley.

[7] Bertsimas, D., M. Sim. 2003. Robust discrete optimization and network flows. *Mathematical Programming* **98**(1) 49–71. doi:10.1007/s10107-003-0396-4. URL http://dx.doi.org/10.1007/s10107-003-0396-4.

[8] Bertsimas, D., M. Sim. 2004. The price of robustness. *Operations Research* **52**(1) 35–53. doi:10.1287/opre.1030.0065. URL http://or.journal.informs.org/cgi/content/abstract/52/1/35.

[9] Bratu, S. 2003. Airline passenger on-time schedule reliability: Analysis, algorithms and optimization decision models. Ph.D. thesis, Massachusetts Institute of Technology.

[10] Bratu, S., C. Barnhart. 2005. An analysis of passenger delays using flight operations and passenger booking data. *Air Traffic Control Quarterly* **13**(1) 128.

[11] Burke, E., et al. 2009. A multi-objective approach for robust airline scheduling. *Computers & Operations Research* doi:10.1016/j.cor.2009.03.026. URL http://www.sciencedirect.com/science/article/B6VC5-4W15KX6-1/2/2cdebbe7ef7d245c5fd615e310fbc1b3.

[12] Charnes, A., W. W. Cooper. 1959. Chance-Constrained programming. *Management Science* **6**(1) 73–79. doi:10.1287/mnsc.6.1.73. URL http://mansci.journal.informs.org/cgi/content/abstract/6/1/73.

[13] Charnes, A., W. W. Cooper. 1963. Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Research* **11**(1) 18–39. doi:10.1287/opre.11.1.18. URL http://or.journal.informs.org/cgi/content/abstract/11/1/18.

[14] Clarke, J.P., T. Melconian, E. Bly, F. Rabbani. 2007. MEANS: MIT extensible air network simulation. *Simulation* **83**(5) 385–399. doi:10.1177/0037549707063766. URL http://sim.sagepub.com/cgi/content/abstract/83/5/385.

[15] Eggenberg, N. 2009. Combining robustness and recovery for airline schedules. Ph.D. thesis, Ecole Polytechnique Federale de Lausanne.

[16] Ehrgott, M., D. Ryan. 2000. Bicriteria robustness versus cost optimisation in tour of duty planning at air new zealand. *Proceedings of the 35th Annual Conference of the Operational Research Society of New Zealand*. 31–40.

[17] Fearing, D., V. Vaze, C. Barnhart. 2010. Modeling passenger travel and delays in the national air transportation system. Working Paper.

[18] Gao, C., E. Johnson, B. Smith. 2009. Integrated airline fleet and crew robust planning. *Transportation Science* **43**(1) 2–16. URL http://transci.journal.informs.org/cgi/content/abstract/43/1/2.

[19] Gronkvist, M. 2005. The tail assignment problem. Ph.D. thesis, Goteborg University.

[20] Jacobs, P., A. Verbraeck, J. Mulder. 2005. Flight scheduling at KLM. *WSC '05: Proceedings of the 37th conference on Winter simulation*. Winter Simulation Conference, 299–306.

[21] Kang, L. 2004. Degradable airline scheduling: an approach to improve operational robustness and differentiate service quality. Ph.D. thesis, Massachusetts Institute of Technology. URL http://hdl.handle.net/1721.1/17659.

[22] Kohl, N., A. Larsen, J. Larsen, A. Ross, S. Tiourine. 2007. Airline disruption management–Perspectives, experiences and outlook. *Journal of Air Transport Management* **13**(3) 149–162. doi:10.1016/j.jairtraman.2007.01.001. URL http://www.sciencedirect.com/science/article/B6VGP-4NB38B5-1/2/7c7da3b05d93f8518d76445a2baace1f.

[23] Lan, S. 2003. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to achieve minimum passenger disruptions. Ph.D. thesis, Massachusetts Institute of Technology.

[24] Lan, S., J.P. Clarke, C. Barnhart. 2006. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. *Transportation Science* **40**(1) 15–28. URL `http://transci.journal.informs.org/cgi/content/abstract/40/1/15`.

[25] Marla, L., C. Barnhart. 2010. Robust optimization: Lessons learned from aircraft routing. Working Paper.

[26] Rosenberger, J., E. Johnson, G. Nemhauser. 2004. A robust Fleet-Assignment model with hub isolation and short cycles. *Transportation Science* **38**(3) 357–368. URL `http://transci.journal.informs.org/cgi/content/abstract/38/3/357`.

[27] Rosenberger, J., A. Schaefer, D. Goldsman, E. Johnson, A. Kleywegt, G. Nemhauser. 2002. A stochastic model of airline operations. *Transportation Science* **36**(4) 357–377. doi:http://dx.doi.org/10.1287/trsc.36.4.357.551.

[28] Schaefer, A., E. Johnson, A. Kleywegt, G. Nemhauser. 2005. Airline crew scheduling under uncertainty. *Transportation Science* **39**(3) 340–348. URL `http://transci.journal.informs.org/cgi/content/abstract/39/3/340`.

[29] Shebalov, S., D. Klabjan. 2006. Robust airline crew pairing: Move-up crews. *Transportation Science* **40**(3) 300–312. URL `http://transci.journal.informs.org/cgi/content/abstract/40/3/300`.

[30] Sherry, L., D. George. 2009. U.S. airline passenger trip delay report 2008. Tech. rep., Center for Air Transportation Systems Research, George Mason University.

[31] Smith, B., E. Johnson. 2006. Robust airline fleet assignment: Imposing station purity using station decomposition. *Transportation Science* **40**(4) 497–516. URL `http://transci.journal.informs.org/cgi/content/abstract/40/4/497`.

[32] The Joint Economic Comittee. 2008. Your flight has been delayed again. Tech. rep. URL `http://www.jec.senate.gov`.

[33] Yen, J., J. Birge. 2006. A stochastic programming approach to the airline crew scheduling problem. *Transportation Science* **40**(1) 3–14. URL `http://transci.journal.informs.org/cgi/content/abstract/40/1/3`.

[34] Zhu, Y. 2009. Evaluating airline delays: the role of airline networks, schedules, and passenger demands. Master's thesis, Massachusetts Institute of Technology. URL `http://hdl.handle.net/1721.1/47775`.