

Metrics and Methods for Improving Airline Schedule Reliability

By

Massimo Morin

M.S. Software Engineering
University "Ca' Foscari" of Venice, 1996

Submitted to the Department of Aeronautics and Astronautics in Partial Fulfillment of
the Requirements for the Degree of Master of Science in Transportation

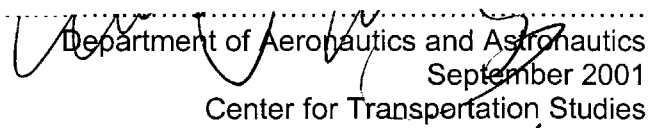
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

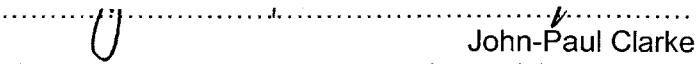
September 2001

© 2001 Massachusetts Institute of Technology
All rights reserved

Signature of the author.....

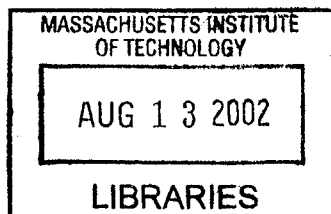

Department of Aeronautics and Astronautics
September 2001
Center for Transportation Studies

Certified by.....


John-Paul Clarke
Assistant Professor of Aeronautics and Astronautics
Thesis supervisor

Accepted by.....

Wallace E. Vander Velde
Professor of Aeronautics and Astronautics
Chair, Committee on Graduate Students



AERO

Metrics and Methods for Improving Airline Schedule Reliability

By
Massimo Morin

Submitted to the Department of Aeronautics and Astronautics in Partial Fulfillment of the Requirements for the Degree of Master of Science in Transportation

Abstract

Airline scheduling is a daunting task. Much time and resources are spent by airlines developing a schedule that meets expectations of profitability and competitiveness. Most of the time, however, the reliability aspect has a minor, if any, role in such a process. In reality disruption of the schedule occurs due to unforeseen events such as weather conditions, traffic congestion, and mechanical problems. The outcomes of these events are cancellations and delays. The impact that these disruptions have on airline operations is not only the increased cost for system maintenance and recovery, but also the loss of profitability and the perception of poor and unreliable service for the flying customer.

In this thesis we present an analysis of the schedule design process, highlight the drawbacks of the current proceedings and outline of new and more flexible framework for schedule design. We define a reliability measure, the Option Value, and a way of comparing flights based on the reliability they are providing, via the Option Disruption Value. The idea of reliability is based on the concept of flight performance: a flight is more reliable if it is able to match or outperform the on-time performance of the flights that leaves its origin station and arrives at its final destination at or near its arrival and departure times.

Based on these two measurements, we quantify the robustness and coverage of a sample schedule. Alternative passenger ratings are defined based on the concept of alternative itineraries (Coverage) and alternative independent itineraries (Robustness) that connect two locations. These are the Flight Options and the Flight Protection Options, respectively.

Fifteen methods to modify flight schedule are proposed. One method, Reduce/Increase Flight Slack Time (R/IPTS) was evaluated. Results indicate that R/IPTS was effective in increasing reliability in 70% of the flight considered, but that other methods need to be employed if reliability is to be increased further.

Thesis Supervisor: John-Paul Clarke

Title: Assistant Professor of Aeronautics and Astronautics

To My Parents and Friends, Here and Abroad.

Acknowledgements

My sincere gratitude goes to Professor **John-Paul Clarke** for his constant understanding and support. Without him this document could have taken much longer. Thank you, JP.

A big thanks goes also to my lab mates, especially the one in the Garden Corner, who so patiently and kindly help me out in the difficult and crazy times.

Above all, I owe a big thank you to:

1. **Francis Carr**. Without his help with Matlab and the patience to listen and explain statistical analysis to me, I would have been still trying to figure out what the heck is a convolution of functions.
2. **Terence Fan**. His deep understanding of airline operations and the transportation business has been a big boost to the development of the economic analysis about the protection options. Thank you Terrence, I know it takes a while but after all I think I got what you were saying ;-)
3. **Terran Melconian**. His support in configuring my computer and keeping my workstation and myself up to date has been a daunting, but I hope rewarding, task. Thank you Terran, also for reading this long thesis: your comments have been greatly appreciated.
4. **Jay Falker**. The very first person to give credit to my research, maybe because he only had a couple of hours to spare. I went a long way since our first talk!
5. Last but not least **Manuel "Chico" Roth**. Partner in crime during hard times with MS Windows™, MS Word™¹, deadlines, and critical issues that amazingly arise only when you are writing the final thesis draft. Writing a thesis is always a daunting task but it can be much easier if we are two struggling to meet the closing dates. Good luck, Chico, with your dissertation!

I need also to thank everybody who supported me, from my previous company boss, **John Pararas**, and my senior project manager **Albert "AI" Asber**, to the professors here at MIT who help me out understanding the deep implications of transportation issues, in particular in the Air Transportation Field.

I thank also the *Scopone Scientifico Knights at honorem*, **Chico**, my partner **Francis**, **Jonathan "The Canadian"** and **Nhut**. Guys finally you learned how to play a serious game. Thank for letting me winning all the time, though.

Finally big thanks go also to all the friends and people who directly or indirectly stayed near me during happy and less happy times. A very special thank you goes to **Irma** for being so patience and understanding with me. Thanks, dear, you made the difference!

In the end, I still own a big thank you, even if most of the time it is unspoken, to my parents, for the great opportunity and freedom they gave me, and the patience with which they are still waiting me back in Italy. **Mom, Dad**, do not worry, I will finish studying and I'll come back home... sooner or later ;o}

Massimo Morin

¹ MS Windows™, MS Word™ and all the other buggy software™ are a trademarks or registered trademarks of Microsoft™ Corporation.

Biography

Massimo Morin was born in Italy March 8 1971. He lived most of his time in the large farmhouse of his dad Antonio and mom Marta Morin in Codevigo, Padova, Italy. Since he was very young he has been interested in computer science: he obtained an associate degree in computer science at the "Zuccante" Technical State Institute in July 1990 at the age of 19. After working for one year as a software developer and customer support person at his computer science teacher company, IESSE s.r.l., he joined the Italian army. He completed his military service as lieutenant in July 1992.

In September 1992 he started the Master of Science in Software Engineering at the university "Ca' Foscari" in Venice. He completed all the required courses in May 1996. In November of the same year he graduated as one of the top three students of his and previous class. His final grade was of 110/110 cum laude and the thesis was about the quality of service offered by the ground side of airports, with analysis of Linate and Malpensa airport in Milan, Italy. The completion of the thesis did not stop him from continuing to do consulting and software development. In particular he devoted most of his time to analyzing and proposing software solutions the Crew Scheduling Problem for regional bus companies. He was continuously supporting software tools for the internationalization of civil engineering applications for Concrete s.r.l. in Padova, IT. Furthermore, he was involved with his former advisor in the development of SLAM (Simple Land-side Aggregate Model): a prototype of his master thesis ideas for the Milan Airport Authority.

In August 1997, he was hired by Scheduling Systems Inc. located in Waltham MA. After spending three years, all his energy and time in developing decision support systems for maintenance planning for various major US carriers, he decided to improve his transportation skills by starting a Master in Science of Transportation at the Massachusetts Institute of Technology, Cambridge MA.

He took the first courses in Air Transportation Economy, Airline Management and Design and Operate an Airspace Systems part-time, while still working at Scheduling Systems. In September 2000, he started full-time at MIT for the completion of his Master of Science of Transportation, with airline business and operations specialization.

Currently he is living in Belmont, MA, and he can be reached via email at mmorin@mit.edu

Contents

Acknowledgments	7
Biography	9
1. Introduction	17
1.1. The Airline Industry	18
1.2. Disruptions.....	21
1.3. The Disruption Effects.....	22
1.4. Dissertation Objectives and Outline.....	24
2. Irregular Operations: the New View	27
2.1. Approaches.....	27
2.2. The Novel Idea.....	29
2.3. Schedule Generation Strategy.....	31
2.4. Coverage and Robustness.....	32
2.5. Process Revised.....	35
2.6. Building Blocks Dependencies and Interactions.....	38
2.7. Scheduling Robustness Problem.....	40
2.8. New Process.....	41
2.9. Conclusions.....	44
3. Profile Analysis	45
3.1. Evaluation Process.....	46
3.2. Delay and Cancellation Considerations.....	48
3.3. Station Independence.....	49
3.4. Passenger Choice and Profile Hypothesis.....	52
3.5. The RSM Approach and Theoretical Justifications.....	54
3.6. The Profiles: Method and Analysis.....	56
3.7. Profiles: Example.....	62
3.8. Conclusions.....	66
4. Scheduling Problem: Basic Definition and Ideas	67
4.1. Basic Definitions.....	67
4.2. Robust Schedule.....	70
4.3. Coverage Example.....	76
4.4. Market Protection and Traffic Flow.....	78
4.5. Decision Making Process.....	79
4.6. Conclusions.....	81
5. Metrics and Reliability Measures	83
5.1. Reliability Measures.....	83
5.2. Option Value for a Leg	85
5.3. Slack Time and Behavioral Response.....	87
5.4. Option Value for a Connection.....	91
5.5. Flight Coordination and Rescheduling Costs.....	93
5.6. Option Value for a Flight.....	95
5.7. The Cost Term of a Trip.....	96

5.8.	Option Disruption Value.....	99
5.9.	Analysis of OV and ODV Characteristics.....	100
5.10.	A Posteriori Schedule Effectiveness Measures.....	101
5.11.	Conclusions.....	103
6.	Schedule Improvements	105
6.1.	Operation on the Network.....	105
6.2.	Flight Leg Changes.....	107
6.3.	Flight Changes.....	109
6.4.	Rotation Changes.....	111
6.5.	Action Location.....	112
6.6.	Stations Analysis.....	113
6.6.1.	Stations: Cancellations.....	114
6.6.2.	Stations: Delays.....	116
6.7.	Flight Leg Analysis.....	120
6.7.1.	Flight Legs: Cancellations.....	120
6.7.2.	Flight Legs: Delays.....	122
6.8.	Flight Analysis.....	128
6.9.	Areas of Concentration.....	129
6.9.1.	Areas of Concentration: Cancellations.....	129
6.9.2.	Areas of Concentration: Delays.....	130
6.10.	Rotation Analysis.....	131
6.10.1.	Rotation: Delays.....	133
6.11.	Leg Load Analysis.....	135
6.12.	Conclusions.....	136
7.	Model Interactions	137
7.1.	Robust Schedule Model Input, Output, Interactions.....	137
7.2.	Model Validation.....	142
7.3.	Validation Models: Expectations.....	143
7.4.	Increase/Decrease Flight Slack Time Problem.....	146
7.4.1.	High Load Legs.....	146
7.4.2.	Profile Analysis.....	148
7.4.3.	Slack Time Calculation.....	148
7.4.4.	Option Value.....	150
7.5.	Conclusions.....	150
8.	Conclusions	153
8.1.	Summary.....	153
8.2.	Model Results.....	155
8.3.	Additional Research.....	157
8.4.	Conclusions.....	159
	Bibliography	161

List of Figures

1.1	Schedule Planning Process. Canonical Form.....	19
1.2	Flight Disruption on Departure. Possible Cases.....	23
2.1	Isolated Station Example.....	34
2.2	Detail Model. Dependencies and Combination of the Building Blocks.....	38
2.3	Parts and Interaction of the Robust Scheduling Problem.....	41
2.4	Incorporation of the Robust Scheduling Problem in the New Framework.....	43
3.1	Network as a Source of Disruptions.....	49
3.2	Lognormal Probability Distribution Function Chart.....	55
3.3	Visualization Tool for Taxi-out time in EWR July 2000.....	58
3.4	Removal of Last Observation.....	59
3.5	Trip Profile Visualization July 2000 Newark (EWR) to All Stations.....	60
3.6	Lognormal Stations Fit on Station With Less Than 70% Profiles Fit.....	61
3.7	Normal Distribution Fit on Station With Less Than 80% Profiles Fit.....	62
3.8	Gate Profile EWR 8:10 on Mondays July 2000.....	63
3.9	Taxi out Profile EWR 8:10 on Mondays July 2000.....	64
3.10	Trip Time Profile Newark Denver (EWR-DEN) on July 2000.....	64
3.11	Taxi-in Profile DEN 10:30 on Mondays July 2000.....	65
4.1	Network Topology.....	70
4.2	Example of Coverage and Robustness.....	72
4.3	Flow Controller Decision Framework.....	78
5.1	Arrival Time Profile. Lateness Analysis.....	88
5.2	Profile Layout EWR-DEN Flight Leg Departing at 8:15.....	90
5.3	Comparison Convolution and Normality Assumption for Arrival Time Profile.....	90
5.4	Measurement of Schedule Delay: Passenger's Decision-Making Process.....	97
6.1	Total Cancellation Distribution July 2000.....	114
6.2	Distribution First 80% of the Cancellation in July 2000.....	115
6.3	Cancellation Compared to Station Traffic July 2000.....	116
6.4	Departure Schedule Station Variability.....	117
6.5	Departure Schedule Station Average Variability.....	118
6.6	Arrival Schedule Station Variability.....	119
6.7	Arrival Schedule Station Average Variability.....	119
6.8	Leg Cancelled.....	121
6.9	Proportion of Leg Cancelled.....	121
6.10	Leg Departure Variability. All Stations.....	123
6.11	Leg Departure Variability. Most Delayed Legs.....	123

6.12	Leg Arrival Variability. All Stations.....	124
6.13	Leg Arrival Variability. Most Delayed Legs.....	124
6.14	Most Unreliable Legs.....	126
6.15	Most Delayed Legs.....	127
6.16	Leg Airborne Variability.....	128
6.17	Areas of Concentration: Cancellations.....	130
6.18	Areas of Concentration: Delays.....	131
6.19	Rotation Visualization Tool	132
6.20	Rotation Delay Variability.....	133
6.21	Rotation Variability by Carrier.....	134
6.22	Rotation Variability by Equipment Type.....	134
7.1	The Schedule Robustness Problem	138
7.2	Interactions of the Core Parts.....	139
7.3	RSM with Highlighted Conceptual Division and Sub-Problems.....	141
7.4	Legs Load on Continental Airline Network July 2000.....	147
7.5	Load Distributions on the 1000 Flight Legs with Higher Load.....	147
7.6	Slack Time Distributions Across the 1000 Most Loaded Legs.....	149
8.1	Dispatcher Interactions and Rerouting Activity.....	158

List of Tables

2.1	Input and Output of the Building Blocks	39
4.1	Coverage Example.....	76
4.2	Coverage Reference.....	76
4.3	Reference Flight.....	77
4.4	The $C_{120}(\text{EWR-LAS}, f)$ Set.....	77
6.1	Flight Legs Actions: Impact on Schedule Design Process	108
6.2	Flight Leg Actions: Schedule Improvements.....	109
6.3	Flight Leg Actions: Airline Costs and Restrictions Implications.....	109
6.4	Flight Actions: Impact on Schedule Design Process.....	110
6.5	Flight Actions: Schedule Improvements.....	110
6.6	Flight Actions: Airline Cost and Restriction Implications.....	111
6.7	Rotation Actions: Schedule Improvements.....	112
6.8	Rotation Actions: Airline Cost and Restriction Implications.....	112
6.9	Departure Variability.....	125
6.10	Arrival Variability.....	125

Chapter 1

Introduction

Airlines obtain their revenue by providing a “connectivity service” for people and goods. The service follows a published schedule: that is determined via a complex decision-making process that attempts to generate the optimal schedule for the service. Unfortunately, the condition in which the service is planned is very different from the circumstances in which it operates. This uncertainty related to the environment in which the airline has to function results in a fragile schedule that cannot be recovered when, as is often the case, the schedule is disrupted, by delays and cancellations.

One of the most important characteristics of the airline schedule is its **reliability**. The level of reliability or flexibility and adaptability to the changing environment drives the airline performance, and at the same time dictates the passengers' satisfaction.

In this chapter we describe the process that every airline goes through while designing the service schedule. This is the most important process for any airline, but its complexity precludes full consideration of all possible conditions that may be present during operation.

Following this description, there is a short characterization of the disruptions: what their characteristics are and how they impact the airline business. This is followed by a description of the effect on the airline service, and how the airline operates in these situations. Finally, there is an outline of the thesis and a description of the contents of the chapters to come.

1.1 The Airline Industry

The airline industry provides connectivity between people and places. Every airline offers a service: moving passengers from location a to location b following a publicly available *schedule*. Given a schedule it is possible to identify cycles of service offered that are usually developed for a day or a week. These periods, or cycles, define the airline operations for the market $a b$.

The collection of plane movements on the connected markets is called the **airline operations**. Historically the planning of these operations, or **schedule planning process**, is achieved throughout a sequence of steps. Different approaches are taken by different airlines; thus, the steps depicted here should be interpreted as a representative approach. Nevertheless, all the different variation of the schedule planning process can be obtained from the process described here. This is, in fact, called the *schedule planning process in the canonical form* and it is composed of market identification, schedule design, fleet assignment, aircraft routing and crew scheduling.

The first step is **market identification**: based on historic data, analysis of the attractiveness of the market and the competitor's situation, the airline has to determine if it is advantageous to serve a given market. These considerations are based on expected demand and consequent expected profit. If during a complete period of operations (*cycle*) the service is not profitable, then it is not worthwhile to introduce or maintain such a service.

Once profitability is determined, the next step is **schedule design**: when to fly aircraft from a to b . This is critical because the demand is not evenly distributed along the cycle. There are peaks when most people want to fly. These are the most attractive times to offer flights, but, of course, all flights cannot be offered at the same time. Other factor must be also taken into consideration, such as service frequency and aircraft utilization. The presence of competitors may increase the complexity of this step: their service and their competitive response must be accurately predicted. The result of this phase is a set of legs to fly.

The third step consists of the **fleet assignment**: determination of the aircraft type to fly. A fleet is a group of aircrafts that have common characteristics like special capability (like over water), range and capacity. The whole network, therefore, is divided into sub-networks that a single fleet can handle. The network subdivision and association of the fleet to the network depends of the characteristics of the fleet and the aircraft availability.

The problem is to assign an aircraft type to every leg that has to be flown, considering capacity needed and aircraft available.

The next step is called **aircraft routing**. A rotation is a series of flights that are flown by a single aircraft. The problem is to concatenate flights into rotations and associate every rotation to a single aircraft. This has to be done while maintaining the planned service and allowing the aircraft to undergo maintenance. The maintenance planning is required because an aircraft cannot fly for more than a fixed number of hours between maintenance layovers. The maintenance is performed at checkup locations. The locations vary by geographical areas and aircraft types. These locations may not be in either *a* or *b*; therefore a proper *visit* to a third location has to be planned, without disrupting the service.

The last step is the **crew scheduling**: assign a crew to every flight that is going to fly the aircraft determined in the previous phase. The crews are subjected to various restrictions and costs; therefore it may happen that the same crew does not always fly the same flight or the same aircraft. The problem gets more complicated when considering aircraft certification, rests, connections, contracts and seniority. The objective is to build up rotations for the crews. In this case, every leg has to have a crew (pilot, copilot and flight attendants) without exceeding the number of crew available, allowing for their connections and maintaining policies and contractual agreements.

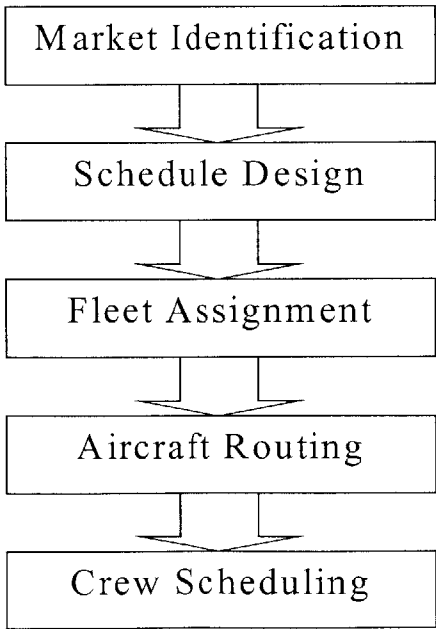


Figure 1.1 Schedule Planning Process. Canonical Form

The frequency that the airline goes through these steps varies: market identification usually happens once a year and rarely changes, while crew scheduling occurs every couple of months or every change of the season. The season, the optimization of the service, and competitive response are other factors that contribute to schedule adjustment.

The schedule is the final product of an airline. This is used as an input to the solution of the aircraft routing and crew scheduling problems. The airline is very sensitive to the solution of these two problems because equipment and crew costs are the highest expense after fuel costs.

The aircraft routing problem has two aspects: an aircraft costs the airline in insurance, maintenance and ownership (commonly defined as *fixed costs*) even if it does not fly. It is clear that the more a plane flies, the less these costs impact on the total aircraft operating costs on a per hour basis. What it is advisable, therefore, is having a high utilization period. It is typical to find, in fact, utilization of 12 to 14 hours a day on a domestic service (international service can rise up to 18 hours a day). The second issue is related to the maintenance: if the aircraft does not go into maintenance every 60 flying hours, it has to be grounded. In such a case, the mechanics have to be taken with their equipment to the location of the plane. The grounding of the plane causes disruption of the service, reduction of revenue due to equipment unavailability and added cost due to labor and crew relocation. In conclusion these two problems address the need that the plane must fly and must be maintained.

The crew scheduling problem is important because crews are very expensive (second largest expense after fuel cost). Their compensation is regulated by complex contracts: flying time, minimum benefits due, out of *residence* fees and minimum rest, are only some of the criteria used for determining such amounts. As in the case of the aircraft, a crew is paid the minimum wage independently of their fly time. There is also a contractual fly time that determines the maximum number of flying hours: if the effective flying time exceeds such limit then the crew gets paid overtime.

In normal operations both these problems get solved in a very efficient and effective way. However, the solutions work only in the ideal condition. When disruptions occur these solutions are highly inefficient and a recovery plan has to be put in place.

1.2 Disruptions

Service disruption is more the norm than the exception in the airline industry. It is so frequent that the planned/published schedule is only indicative: the real schedule, called *operating schedule*, is adjusted constantly, and it is the one that is usually followed by the airline. The departure and arrival time presented on the schedule in use are not even unique: for every flight there are *scheduled*, *expected* and *actual* departure and arrival times.

Disruptions, or service irregularities, occur for three main reasons: **weather**, **air traffic control (ATC)** and **equipment malfunction**. The weather is the predominant factor and most of the time causes or aggravates the others.

All of these events have specific characteristics:

1. **Predictability**: when and how often it occurs.
2. **Duration**: how long its action persists.
3. **Intensity**: how severe the event is.
4. **Impact**: what area it is affecting.
5. **Duration of recovery**: how long the recovery from the disruption will take.

Predictability is the ability to forecast the events. Weather usually follows patterns: it is easy to identify the overall period in which they may occur (e.g., seasons). The weather is continuously monitored. There is, therefore, a continuous and steady flow of information about the weather situation going to the *Airline Operation Control Center* (AOCC). The degree of accuracy of the prediction of the evolution of the situation depends on the time horizon. Air traffic control is less predictable and is based on congestion information. As with the weather condition, the ATC situation is continuously updated and therefore reasonable predictions can be obtained (on the order of hours). Equipment failures are less predictable, hence less tractable.

The duration of the event indicates how long such an event affects the system. The extent of the event is something that can be forecast and can give an indication of how long an irregularity will persist.

The intensity specifies the area covered, while impact tells how much it is affecting the airline operations. The impact is related to the type of operations the airline is managing, the sensitivity of the area, the type of event and its duration. Weather is the most problematic because it moves and can change intensity easily. The airline hubs and their surroundings are the most important stations because they can affect the whole network. An equipment problem is more limited because it involves a single

aircraft; nevertheless, more than one flight can be involved depending on the length of the problem, the sequence of flights the aircraft has to cover, and the recovery procedure.

The duration of the recovery varies in relation to all the types of disruption. The recovery procedure that the airline uses is focused on restoring the regularity of the operation. This involves modification of the rotation, the crew shifts, relocation of the equipment (ferrying), and adjustment of the maintenance schedule. The procedure can take from hours to days; the longer the recovery time, the greater the disruption that the airline sustains.

1.3 The Disruptions Effects

When one or more such events occur, there are two negative effects on the airline operation that are produced: **delays** and **cancellations**.

A flight is considered delayed when it does not reach its destination within 15 minutes of the *scheduled arrival time (SAT)*. The time loss can happen before departure, during the trip or just before landing. There is no penalty for late departure which does not affect the landing time. When the problem is localized to a single aircraft (equipment malfunction) then the flight can be delayed or cancelled. In the latter case, passengers are moved to other flights or to other airlines. In the case of weather or ATC problems, an extensive area and multiple stations can be blocked, and cut out of the network.

Every possible situation can be summarized by considering one of more airports closed for one or more flights and for a variable period of time. The station closure can be for inbound flight, outbound flight or both. In all cases it is a decision of *the scheduler*, an operator located at the AOCC, whether to delay the flights or cancel them.

As reported in Mathaisel (1996), if the station is blocked for inbound flights, various options are available to the scheduler:

1. **Delay of departure:** a flight that has the destination station closed is delayed before departing for a variable amount of time.
2. **Cancellation of the flight:** a flight is cancelled and passengers are diverted to other flights (not necessarily on the same airline).
3. **Preemptive take off:** the flight may be delayed but takes off anyway, in the hope the airport will be available at the *actual arrival time (AAT)*.

4. **Overfly:** this is typical for a flight going from *a* to *b* and then to *c*. The flight takes off but it does not stop at the interrupted station (*b*): it continues to the next destination (*c*).
5. **Diversion:** the flight takes off but the final destination is changed to a different one during the trip.
6. **Return to station:** the flight takes off but during the trip the arrival destination is not available and therefore it returns to the original station.

In case of station closure during departure, the situation is simpler: the flight is delayed or cancelled depending on the duration of the disruption.

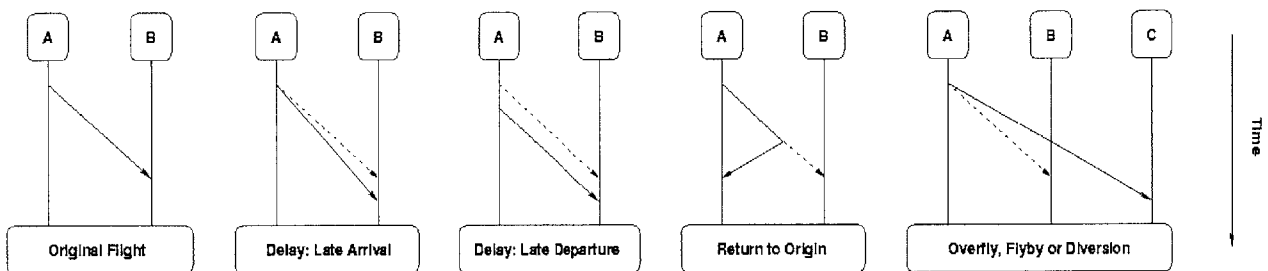


Figure 1.2. Flight Disruption on Departure. Possible Cases.

All the possible scenarios are depicted in Figure 1.2. The situation can be relieved by proper readjustment of the flight at the departing station: it may be worth it to modify the aircraft rotation by allocating the available aircraft to departing flights; another option is to move the delayed aircraft to a later flight and use an available aircraft for the continuation of the service; yet another option is to start delaying the connecting flights for a limited amount of time to allow the connecting passengers to transfer to the second flight.

All this juggling of equipment, crew and passengers leaves behind spare aircraft that need to be ferried to some other station, cancellation of flights, rotation readjustment, and angry passengers who are forced to go home or to stay overnight in an undesirable location. This last point is particularly important because airlines have notoriously poor policy for dealing with passengers. When passengers miss connections, experience long delays, and/or arrive at a different location, airline official usually compensates them for it in a variety of ways. The compensation is defined to be *hard* when a sum of cash is handed out to the passenger, based on the kind and the severity of the delay. Most of the time, instead, there is a so-called *soft compensation* that includes expense coverage, a free ticket, free relocation on a different flight, upgrade to a higher class of

service, extra frequent flier miles, and bonuses. The habit of using both soft and hard compensation heavily affects the revenue stream of the airline. What is worse is that passengers become dissatisfied with the service and then tend to **spill** to the competing airlines. In the long run this greatly affects airline profitability.

1.4 Dissertation Objectives and Outline

As seen, service disruption is to be avoided as much as possible. Ranges of procedures have been studied, some more successful than others.

The problem is indeed quite complex: there are solutions for recovering when a disruption occurs. However, none of these solutions take advantage of schedule structure, or impose a special arrangement of the schedule. What is missing is a clear redefinition of the schedule planning approach with the purpose of minimizing the passengers' disruption impacts.

In this thesis this aspect is analyzed in more detail. A new approach for coping with the airline schedule design process will emerge. A consequent new way of measuring the disruptions will be outlined as well. This will help to formulate some solutions to the problem and will indicate a better way for passengers to select their trips.

Chapter 2 is a review of the recovery process, and the introduction of a new idea for modeling the schedule planning process. This is the first key part of the new scheduling process.

Chapter 3 presents a performance-based characterization of the network and the concept of the time performance profile for a flight. This is the second cornerstone of this document.

Chapter 4 illustrates a more formal way to describe the elements that we are dealing with. This will give us a powerful instrument for writing down the ideas expressed in Chapter 2.

Chapter 5 pushes forward the idea of performance-based characterization of the network and describes a way of measuring the reliability of the service offered by an airline in a more formal way.

Chapter 6 shows how to take advantage of the reliability measures via *grading* the flights, and then proceeds with a variety of ideas on how to modify the schedule for increasing reliability.

Chapter 7 focuses on one of the detailed models outlined in Chapter 6 for improving the schedule. The concept of schedule slack time is studied. A way of validating the new schedule and the consequent expectations coming from it is presented.

Chapter 8 concludes this document with an analysis of the results obtained by the model presented in Chapter 7. Suggestion for future directions and research are derived from the results of this initial solution.

Chapter 2

Irregular Operations: the New View

In this chapter there is a presentation of the actions that the airline starts when a severe schedule/network disruption occurs. There is an introduction to the airline irregular operation problem and a description of the importance and implication of such a problem for the airline operations.

A presentation of the steps followed during the process of schedule planning and a more formal characterization of them follows. The definition of coverage and robustness of the schedule with a description of the implications on the schedule planning process is introduced.

In the end, after a presentation of a revised model and the notation for representing the schedule design problem, there is a definition of the Schedule Robustness Problem (SRP) and how its solution will impact the solution of the Airline Irregular Operations Problem (AIOP).

2.1 Approaches

Much of an airline resources, expertise and computational time are invested in development of an operable schedule that maximizes airline profitability. When an unforeseen event, such as severe weather patterns, unexpected aircraft failure or a more generic, irregular event occurs, the airline operations are greatly impacted.

In the past years, airlines have become more concerned about developing optimal flight schedules, with very little opportunities to accommodate system failures. The

resulting schedule is therefore not flexible enough to adapt to these unforeseen events (*system shocks*) in the system in which the airline operates.

Substantial research has been done in the area of schedule design i.e. solving to optimality based on airline characteristics and strategies, but very little has been done to address the impact of irregular operations.

The approaches used to address the issue of irregular operations tend to reschedule the flights after the disruption occurs. The rescheduling is done in a robust, efficient and real-time mode to reassign operating aircrafts to flights and concurrently reconstructs the residual airline network. As reported by Clarke (1997) some research has been done in this sector by various institutes and airlines.

The first to propose an approach to this problem were Teodorovic and Gubernich (1984), while working at the Research and Development Department of United Airlines. The solution they propose tends to minimize the overall passenger delay in the aftermath of a scheduled perturbation. They attempt to determine the least expensive set of aircraft routings and schedule plan using a branch and bound procedure.

Teodorovic (1985) proposes another algorithm for measuring and improving airlines schedule reliability related to meteorological conditions. This is done via identifying the adaptability of the airline schedule to weather conditions; after that, he tries to minimize the number of aircraft required to accommodate the passengers' traffic resulting during such a weather event.

Teodorovic and Stojkovich (1990) and (1995) follow with a discussion of an algorithm for rescheduling aircraft via minimizing the number of cancellations. They attempt also to incorporate dispatcher experience in the decision process regarding traffic management.

Decision support systems have also been introduced. They do not contain any smart algorithm but only some aid for the dispatcher for rescheduling the various operations. Jarrah (1993), Yan and Yang (1996), Cao and Kanafani (1997a and 1997b), and Mathaisel (1996) all present such an idea basing all the considerations on software programs designed around network flow theory.

Other approaches range from Arguello (1997) that presents real-time optimization models for reconstructing the aircraft schedule based on forecast of delays, to Clarke (1997) that incorporates aircraft routing and spillage considerations to limit the "*lost passengers*" in the recovering process.

In recent years, a hybrid trend has emerged in which combinations of aircraft assignment and routing problems, or fleet assignment and crew scheduling have been considered (see Barnhard (1997), Talluri (1996), Soumis (1980) and Berge and Hoperstand (1993)).

The common idea is that all of these methods try to recover the system to its initial configuration in the fastest way possible while reducing the cost of such actions. Airlines are very cost conscious so they realize that if the initial configuration of the system is *optimal* and a perturbation occurs, then the best way to keep the costs under control is to return to the configuration in use before the inconvenience. The cost considered varies: it may be crew cost, which focuses on the recovery by allowing crews to continue their service; another is cancellation: reduce the number of flights cancelled, allowing the minimum relocation of aircrafts; another is spillage: avoid having passengers to reschedule their trip with another airline.

The recovery operation can be more or less efficient but there are two concerns that arise:

1. Passengers are taken into account only after they have initiated their trip and they are already affected by the disruption. Hence they already experience some delay or cancellation.
2. The recovery is independent from other disruptions. This means that when the recovery plan is initiated there is an assumption that no other disruptions are occurring in the system.

These hypotheses limit the effectiveness of the approaches used for recovery. All the models try to integrate various variables, like passenger spillage, airline image, crew, and fleet costs, but the deep correlation is not easily controllable. The models available all try to deal with the aftermath of the disruptions. The major inconvenience is, therefore, that all the approaches try to recover the system but not to reduce the impact of the disturbance before they occur. In practice all the procedures for recovery are not taking full advantage of the information that is available to the AOCC before the disruption takes place.

2.2 The Novel Idea

The novel idea is to consider a schedule not as fixed but as a result of a series of steps, some of which have the purpose to add flexibility and robustness to it. The schedule

itself should have some reliability and give options for cancellation and for delays and recovery without incurring all the drawbacks mentioned previously. The new schedule is going to be different from the one coming out of the normal scheduling design process. Part of the optimality in terms of profitability, aircraft utilization, strategic position and dominance is going to be exchanged for a faster and safer recovery opportunity. Notice, however, that the faster and safer recovery translates into better passenger service, reduction of recovery costs and an improvement to the airlines image. There is, therefore, a trade-off to consider between the optimality of the solution and its adaptability.

This is achievable considering that during usual operations there is a flow of information that is collected in the Airline Operation Control Center, and for example this information may be used to determine if a forecasted weather event can cause problems for the network. It is clear, therefore, that if the schedule can adapt to that, via reducing the flow to the concerned area, then when and if a disruptions occurs, its impact will be minimal.

The information coming in will give some indication of the situation. From such indication proper action can be taken. There are two ways to deal with the schedule:

1. **Passive Way.** The information related to an area indicates that some flights are likely to be affected by a disruption. The schedule is not changed but passengers are diverted to other flights that have higher odds of reaching destinations.
2. **Active Way.** The schedule is modified. Traffic to the sensitive area is reduced gradually; aircraft are rerouted or cancelled and passengers are diverted as in the case above.

In both cases if a disruption occurs, the passengers affected by the disruption are reduced. In the first case, passengers are moved and no action is taken for the crew of aircraft rotation. In the second case, the operations involved in the concerned area are adapted.

The flow of information arriving at the AOCC gives only a forecast of a situation that is more or less probable. Based on that, it is possible to gradually reduce the percentage of the flow in the area of interest. When and if the condition improves, then the flow is gradually increased to its original and normal value. In the passive case, the flow considered is passenger flow: the passengers are moved to other flights. In the active case, instead, the same consideration involves fights, and therefore preemptive cancellation may occur.

In all the cases, whenever a disruption occurs, all the approaches and methods used for recovering are still valid: there is no need to develop new models. The important difference is that the system has already reduced the number of aircraft or passengers affected, and therefore the total “cost” of the recovery is going to be reduced.

2.3 Schedule Generation Strategy

There has been some work done to automate the generation of the schedule and, although some success has been accomplished, integration of all the steps has not been achieved yet. In Antes (1996) there are some examples of such approaches.

Etschmaier and Mathaisel (1985) are among the first to approach such a problem. They explain a stepwise method that is called *frequency planning* or *frequency optimization*. After this initial step they proceed with the definition of the departure and arrival schedule; they use the solution of the crew scheduling problem only to evaluate the generated schedule.

Ghobrial et al. (1992) explains another approach used on small airlines with a single aircraft type. Given the market to serve and the set of candidate routes, they determine the flight frequency for every leg. Due to the fact that the frequency can go to 0, they are able to identify which legs to fly and which ones to remove: this process identifies the market that can be served. Thereafter, they establish the departure time of every flight based on a *time-of-day demand model*. In the end, they solve the aircraft routing problem neglecting the maintenance requirements.

Additional refinements are provided by *Suhl (1995)*. As a first step, he estimates the capacity of the legs that are going to serve some predefined markets. Afterwards he gives two options: the first strategy is to solve the flight scheduling problem, followed by a combination of the fleet assignment and aircraft routing. The second strategy instead solves first the fleet assignment problem and then follows with a combination of the flight scheduling and aircraft routing. The final step for both strategies is to solve the crew scheduling problem.

USAir sponsored the research of *Rushmeiner and Kontogiorgis (1997)*. They came out with a seven step process: the first one is called *market planning*; the second is the departure schedule design; the determination of the arrival schedule directly follows;

afterwards the fleet assignment problem and the aircraft routing problem. In the end, there is the crew scheduling problem and the scheduling of the ground staff.

Another approach is proposed by Zils (1996) via the integration of the strategies proposed by Suhl; the innovation is the utilization of genetic algorithms for the combination of arrival and departure scheduling problem, aircraft routing and fleet assignment problem.

All the approaches above try to automate in multiple steps the schedule planning process but none of them add flexibility to the schedule once a market selection has been determined. The complexity of the problem allows only minor adjustment and, therefore the introduction of new characteristics like reliability, robustness, and flexibility needs to be done in a more formal way. The approach for doing this, needs to address:

1. **Characterization.** A formal characterization of the reliability and flexibility of the schedule to incorporate in the canonical process.
2. **Formalism.** A new way of representing the schedule design process. This should identify the key parts of the various stages and the relationship between them.
3. **Interaction.** The interaction between the parts composing the schedule design process (SDP) and the contribution that each of them gives to the final design process.

2.4 Coverage and Robustness

The key part of the schedule that needs to be exploited is its flexibility and reliability. The schedule flexibility can be achieved via a proper characterization of the schedule itself.

The fundamental properties that a schedule has to have for incorporating flexibility are:

1. **Multiple Flight Redundancy.** There has to be multiple flights that connect market a b . In this case if something happened to any of such flight then there is always others available to divert the passengers.
2. **Multiple Independent Flight Redundancy.** The flights should go from station a to b via different routes: if anything happens to any station involved in a route, then the other flights that follow different routes (through other intermediate stations) can still reach the same destination.
3. **Station Isolation.** Isolate the flight going to a sensitive station; hence rotations disrupted on that station would not impact the whole network.

The path redundancy is very important because modern schedules rely on a hub and spoke topology. This topology implies that most of the airline traffic is concentrated on multi-stop flights. The intermediate stops are done in special stations called **hubs**. If a disruption affects a hub, then the whole network is impacted. If it is possible to equilibrate the flow across different hubs or **large gateways** (backup hubs), then the propagation of problems to other hubs, gateways and networks is reduced.

The application of the multiple flight redundancy to a flight or a market can be commonly referred as having the market, or flight, covered; hence this characteristic will be referred as **market coverage**, or simply **coverage**.

When there is multiple independent flight redundancy for a flight or a market, the market or flight will be referred as robust; hence the characteristic will be indicated as **market robustness**, or simply **robustness**.

The multiple flight redundancy and the multiple independent flight redundancy are focused on specific flights that are considered critical. The definition of *critical flights* depends on the airline objective: it may be profitability, connectivity, load or prestige. The fundamental tenet is that there is only a subset of all the scheduled flights that are worth protecting. The protection of these flights is achieved through the first two options. Given a flight f with origin a and destination b , the flight redundancy characteristics impose a scheduling of different multiple flights departing from a and arriving to b close to the original schedule arrival time of f . If any of them is affected by delays then there will be another flight shortly after.

The differences between the multiple flight redundancy and the multiple independent flight redundancy consist of having different paths (routes) that connect a to b . This is achieved when a flight departing from a directed to b has to stop at an intermediate third station c . If the problem is at station c , the availability of multiple flights after the one in consideration will not help, because they still have to pass through the disrupted station c . If, instead, at least one of the next flights is going through a different station (c') then the flight in question is more protected.

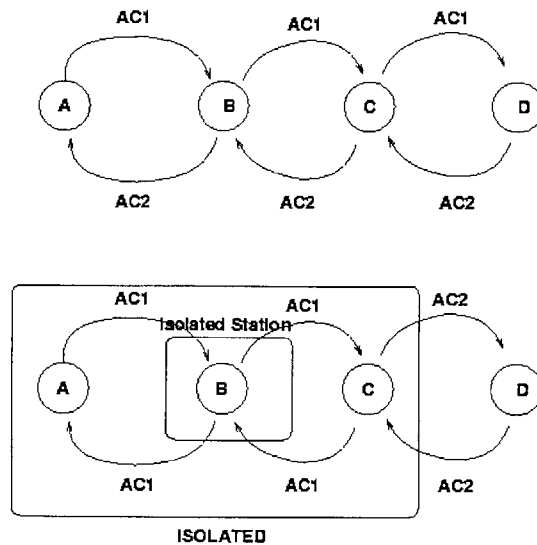


Figure 2.1. Isolated Station Example

The station isolation is related to station problems; as before, it is appropriate to identify some *critical station* (e.g., the hubs and gateways). A station becomes critical if delays experienced in that station are affecting flights that have no apparent relationship between them (no connecting passengers). For example, consider a flight going from *a* to *b* and then to *c* and a second flight going from *c* to *d*. If the aircraft used is the same, a problem involving station *b* that grounds the plane will affect the second flight. The direct connection between the flights is, therefore, only on the aircraft rotation. If instead the station *b* was isolated, e.g. the aircraft of the first flight was looping *a b c b a b c...* then the second flight would have been unaffected. It is clear, therefore, if it is possible to isolate the flights going through a station, then the network impact of such delays is reduced.

Notice that the introduction of a new flight option for protecting a flight or a market will give the characteristic of coverage, robustness or isolation to the market depending on the station where the aircraft will make the intermediate stop.

The set of critical flights (stations) is going to be a subset of the total flights (stations). It is unrealistic to try to protect all the flights because it would mean saturating the system with redundant flights: this is not only unprofitable, but also unfeasible due to limited number of aircraft available. However, due to cycles during the year of weather and air traffic control congestion, it is possible to define by *season* which flights and stations are critical. In such a way, it is possible to put in place a set of schedules that have the coverage and robustness of the market for the desired areas. For example it

could be more worthwhile to protect the flight to the southeast US during summertime than wintertime due to the seasonal prevalence of thunderstorms and hurricanes. On the other hand it is worth protecting the northeast during wintertime due to snowstorms. All this is achievable via different schedules for different seasons according to the seasonal weather patterns.

2.5 Process Revised

The introduction of robustness and/or coverage steps in the schedule has not been considered so far. The problem, as outlined before, is related to:

1. Lack of flexibility in the canonical schedule design process. It is difficult to extend the SDP due to its monolithic form of “only” five major steps.
2. Lack of identification of the core component of the SDP. The schedule design problem incorporates all the problems that a modern airline faces during the design of the service to be put in place; however, most of them interact with each other and usually they are treated together even if it could be more advantageous to separate them into simpler and more tractable problems.
3. Complexity of the problem. The problem itself is complex and extending it further can drive the current algorithm to the limit of its solvability.
4. Black box decision. The solution is usually handled by different sectors of the airline organization (marketing, strategy, operations) that have different priorities on their agenda. Hence every one of them is introducing only its specific optimization leaving the solution of the “more complicated problem” to the remaining departments.
5. Separation of the solution. Usually a solution is taken and worked out without generating a *what-if* scenario due to the unclear result of every sub-problem. In practice, every step receives a schedule in input and generates another schedule in output, without outlining the core modification and changes.

Recently, there has been a trend toward integrating the various parts of the scheduling planning process and then generating a mega-problem, the **scheduling planning problem** (SPP): solving the SPP could definitely return the optimal and ideal result of the schedule planning process because there are more *dimensions* in which the analyst/optimizer can work. There are, however, two basic difficulties: the first one is that the problem becomes intractable. Its complexity and the sheer number of variables that

must be considered literally explode. The second, and subtler, is that it leaves little space for extension and the introduction of special cases.

What is needed is instead a more flexible framework that allows “plug-in” problems: every component will be a “*problem*”. The modules will have a clear interface of data input, a solid, selected and narrow defined challenge to solve and a clear limited output. The unit is easier to handle and to integrate. This is beneficial for a more generic and effective solution of schedule design process.

Hence, a redefinition of the process with which the airline puts the service in place can help to extend the model and allow the insertion of extra steps, such as the schedule robustness one, into it. Antes (1996) underlines this fact and puts emphasis on the schedule planning process as a structured system of smaller sub-parts. Redefinition and isolation of the core parts of the various steps allows the definition of so-called *building blocks* that can be switched and recombined to generate different approaches to solve portions of the master problem.

The building blocks are the following:

1. **Market Evaluation Problem (MEP)**: identify which market it is worthwhile or desirable to enter;
2. **Market Selection Problem (MSeP)**: determine a pool of market that can be served;
3. **Airport Selection Problem (ASeP)**: select the airports that are serving the market desired to enter;
4. **Leg Selection Problem (LSeP)**: determine how to connect the airports selected;
5. **Leg Frequency Assignment Problem (LFAP)**: determine how many times to fly the leg in one day;
6. **Leg Capacity Assignment Problem (LCAP)**: identify the optimal capacity for the aircraft to fly each leg;
7. **Crew Pairing Problem (CPP)**: determine the series of flights that a crew has to fly;
8. **Crew Assignment Problem (CAP)**: choose the pairings for each individual crew such that every crew gets a plan of action for a month and each leg has been covered;
9. **Departure Scheduling Problem (DSP)**: determine the optimal departure time of each leg;
10. **Arrival Scheduling Problem (ASP)**: identify the optimal arrival time of each leg;

11. **Aircraft Rotation Problem (ARP)**: assign a tail number to each flight such that the aircraft can go to maintenance;
12. **Fleet Assignment Problem (FAP)**: determine the fleet to which each leg has to be assigned based on the capacity required.

These steps are not necessarily followed in this order; for example the departure scheduling problem (DSP) and arrival scheduling problem (ASP) are influenced by airport and crew restrictions and this can influence both the crew assignment problem (CPP) and the leg selection problem (LSeP).

The combination of these small blocks can form more complex problems. The traditional method combines them in this way: market evaluation (MEP), market selection (MSP), and airport selection (ASP) generate the market identification problem. The schedule design problem is a combination of the leg frequency (LFAP), arrival scheduling (ASP) and departure scheduling (DSP) problems. The aircraft routing is a combination of the leg frequency (LFP), leg capacity (LCAP) and aircraft rotation (ARP) problems. The crew scheduling is a combination of the crew pairing (CPP) and crew assignment (CAP) problems.

It is possible to define new problems based on new combinations of the building blocks. The relation of the various building blocks and the initial problem are depicted in Figure 2.2.

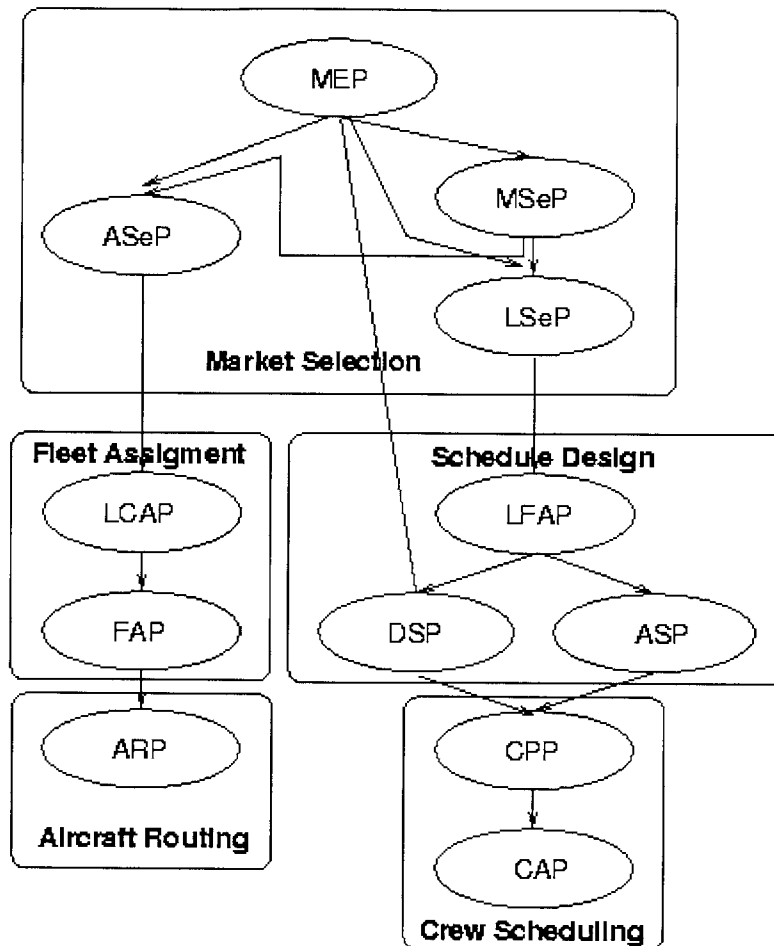


Figure 2.2 Detail Model. Dependencies and Combination of the Building Blocks.

2.6 Building Block Dependencies and Interactions

The building blocks have some dependencies based on the input they are receiving and the output they are generating; some other dependencies, instead, are dictated by pure airline operations logic.

There is input/output dependency because the output of one block is the input of another. If that is the case, the “*input generating*” problem needs to be solved before the “*output consumer*” problem. For example the Crew Pairing problem generates the pairing that defines the service that the pilot has to fly. The Crew Assignment Problem is assigning the pairing to a crewmember and creates the shift that the crew has to fly. Obviously CAP cannot generate any solution if CPP does not generate the pairings.

On the other hand, the Fleet Assignment Problem is independent of the Crew Assignment Problem; however, it is logical to assign an aircraft to a route before assigning a pilot to avoid restriction conflicts when generating the services.

Each block is characterized by its input and output. Some inputs are mandatory and others are optional. The mandatory input is the fundamental data needed to solve the building block sub-problem. The optional input is not necessary to the solution of the problem, but it can give insight to the solver for generating more appropriate solutions.

One of the most important inputs, and the starting point of most of the building blocks, is the O&D Matrix: this is literally a matrix that represents the unconstrained origin-destination demand in a market.

The complete list of the input and output is reported in Table 2.1.

Table 2.1 Input and Output of the Building Blocks

Problem	Mandatory Input	Optional Input	Output
MEP	Historic Data Scenario	Customer Preference	O&D Matrix
MSeP	O&D Matrix		Markets
ASeP	O&D Matrix	Markets	Airports
LSeP	O&D Matrix	Airports Markets	Legs
LCAP	O&D Matrix Legs	Frequencies Markets	Capacities
LFAP	O&D Matrix Legs	Capacities Pairings Markets	Frequencies
FAP	Legs Capacities	Pairing Frequencies D/A Timetable	Fleeting
CPP	Legs	Fleeting Frequencies D/A Timetable	Pairings
DSP	O&D Matrix Legs Frequencies	Fleeting/Rotations Pairings Markets Arrival Timetable	Departure Timetable
ASP	O&D Matrix Legs Frequencies	Fleeting/Rotations Pairings Markets Departure Timetable	Arrival Timetable
ARP	Fleeting Legs	Frequencies D/A Timetable	A/C Routing
CAP	Pairings Legs	Frequencies D/A Timetable	Crew Assignment

2.7 Scheduling Robustness Problem

The introduction of the **Schedule Robustness Problem** (SRP) can be considered an extra problem to insert in the pool of building blocks. The objective of this problem is to increase the reliability of the schedule. The flexibility of the schedule is obtained via the implementing of the three criteria introduced before: multiple flight availability, multiple station redundancy and station isolation. All the criteria can act on different stages of the schedule planning process.

The input of this problem is a set of information including:

1. The original schedule.
2. The critical flight and station to protect.
3. The characterization of the disruption to handle.
4. The leg available.
5. The aircraft rotation.

The output of it is:

1. A new schedule that includes coverage and robustness properties.
2. The new rotations that include the isolation of the station.

Due to its complexity and the various aspects that it covers, it is more convenient to separate such a problem into subparts: **critical input identification** and the **schedule modification**. The input manipulation process is a preprocessing of the schedule that is targeted to identify the resources to be optimized. The critical input identification can be isolated into:

1. **Critical Flight Selection Problem** (CFSP): identifies the flights to be protected during the irregular operation situation or to avoid using as protecting options.
2. **Critical Station Selection Problem** (CSSP): analogous to the critical flight selection problem, it identifies the stations to be isolated during the irregular operation situation or to avoid using them for flight protecting options.

The two problem solutions provide the input to:

1. **The Flight Leg Adjustment Problem** (FLAP): given the critical flight legs to protect, it generates adjustments that improve their reliability.
2. **The Multiple Flight Allocation Problem** (MFAP): given the critical flight to protect, this step determines the allocation of extra flights to cover the market in consideration;

3. **The Multiple Station Redundancy Problem (MSRP):** given the critical station to protect, it generates the set of alternative stations and flights that contribute to the robustness of the schedule;
4. **The Station Isolation Problem (SIP):** given the schedule and the aircraft rotations, it readjusts them to isolate the critical station.

All of the problems presented are going to affect the schedule planning process in various ways, via increasing its scope. Some of the problems above are related, like the critical station selection problem (CSSP) and the multiple station redundancy problem (MSRP); however, the introduction of all of them is not mandatory.

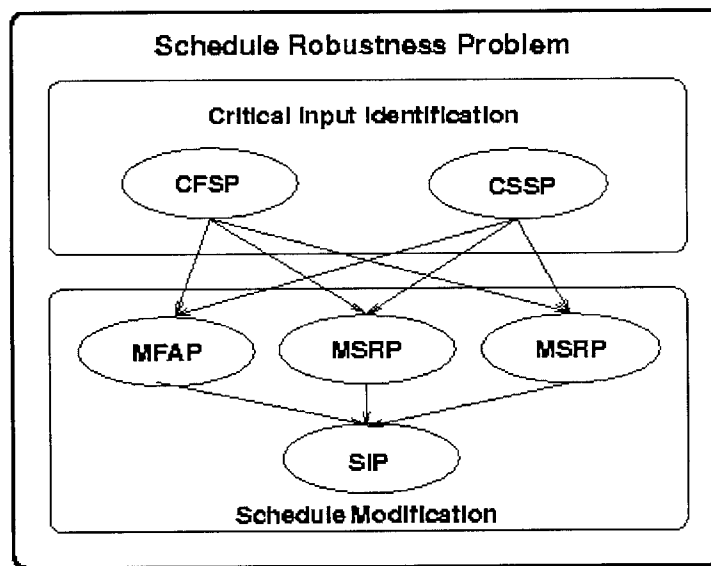


Figure 2.3 Parts and Interaction of the Robust Scheduling Problem

2.8 New Process

Considering the various blocks and their combination, it is convenient to use the notation defined by Antes (1996): this notation is more compact and it describes the problem interactions. The notation is composed of only 2 operators: a combination of block A and B is denoted by $A \oplus B$. Sequential treatment of block A followed by block B is denoted by $A \rightarrow B$. The operator \oplus has higher precedence over the operator \rightarrow .

Using such notation, all the algorithms and models for treating the airline schedule planning process become easy to describe. The process of schedule planning seen in the literature can be expressed in the following way:

1. **Stepwise Approach** by Etschmaier & Mathaisel (1995):

$$\text{LSeP} \oplus \text{LFAP} \rightarrow \text{DSP} \rightarrow \text{ARP} \rightarrow \text{CPP} \oplus \text{CAP}$$

2. **Small Airlines** by Ghobrial et al. (1992):

$$\text{LSeP} \oplus \text{LFAP} \rightarrow \text{DSP} \rightarrow \text{ARP}$$

3. **Strategy 1** by Suhl (1995):

$$\text{MEP} \oplus \text{SeP} \oplus \text{LSeP} \oplus \text{LFAP} \oplus \text{LCAP} \rightarrow \text{FAP} \rightarrow \text{DSP} \oplus \text{ASP} \oplus \text{ARP} \rightarrow \text{CPP} \oplus \text{CAP}$$

4. **Strategy 2** by Suhl (1995):

$$\text{MEP} \oplus \text{MseP} \oplus \text{LSeP} \oplus \text{LFAP} \oplus \text{LCAP} \rightarrow \text{DSP} \oplus \text{ASP} \oplus \text{FAP} \oplus \text{ARP} \rightarrow \text{CPP} \oplus \text{CAP}$$

5. **Integrated Approach** by Zils (1996):

$$\text{MEP} \oplus \text{MSeP} \oplus \text{LSeP} \oplus \text{LFAP} \oplus \text{LCAP} \rightarrow \text{DSP} \oplus \text{ASP} \oplus \text{FAP} \oplus \text{ARP} \rightarrow \text{CPP} \oplus \text{CAP}$$

6. **USAir Approach** by Rushmeier & Kontogiorgis (1996):

$$\text{MSeP} \oplus \text{LSeP} \oplus \text{LFAP} \rightarrow \text{DSP} \rightarrow \text{FAP} \oplus \text{ARP} (\oplus \text{ASP}) \rightarrow \text{CPP} \rightarrow \text{CAP}$$

Going back to the canonical form of the Schedule Planning Process, the schedule design problem can be expressed as

$$\text{LSeP} \rightarrow \text{LFAP} \rightarrow \text{ASP} \oplus \text{DSP}$$

The introduction of the general SRP on the schedule design problem can be denoted as follow:

$$\text{LSeP} \oplus \text{SRP} \rightarrow \text{LFAP} \rightarrow \text{ASP} \oplus \text{DSP}$$

The SRP itself can be modeled in the following way:

$$\text{CSSP} \oplus \text{CFSP} \rightarrow \text{FLAP} \oplus \text{FAP} \oplus \text{DSP} \oplus \text{ASP} \rightarrow \text{MSRP} \oplus \text{LSeP} \rightarrow \text{MFAP} \oplus \text{DSP} \rightarrow \text{SIP} \oplus \text{ARP}$$

Due to the intrinsic complexity of the schedule robustness problem, and the fact that some parts are optional, it is better to introduce the building block of the SRP in the different stages of the canonical problem. In detail:

Introduction of the MFAP:

$$LSeP \oplus CFSP \rightarrow LFAP \oplus MFAP \rightarrow ASP \oplus DSP$$

Introduction of the MSRP:

$$CSSP \rightarrow MSRP \oplus LSeP \oplus ARP$$

Introduction of the FLAP:

$$LSeP \oplus CFSP \rightarrow FLAP \oplus FAP \rightarrow ASP \oplus DSP$$

Introduction of the SIP:

$$CSSP \oplus CFSP \rightarrow MSRP \oplus LSeP \rightarrow MFAP \oplus ASP \rightarrow SIP \oplus ARP$$

Figure 2.4 describes how these parts interact.

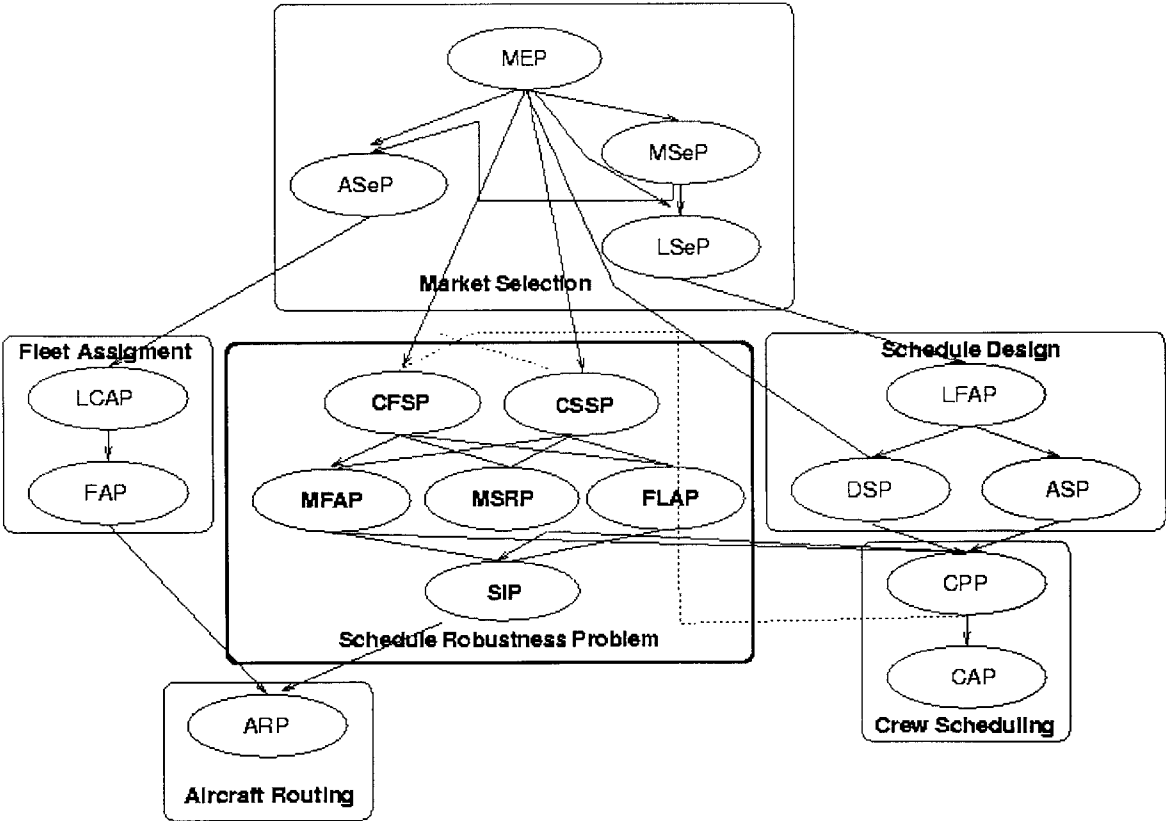


Figure 2.4 Incorporation of the Robust Scheduling Problem in the New Framework

2.9 Conclusions

The outcome of the current, complex and daunting Schedule Design Process is a schedule that is optimal for the airline objectives, but it is poor from the reliability perspective. When a disruption occurs, a posteriori recovery plans are executed to limit the impact of delays and cancellations. This, however, is done only after the facts. During disruptions, the airline incurs in extra costs that are limited only by the adoption of effective and multiple recovery plans. Thus, it turns out that most of the effort and energy spent in producing the “optimal” schedule is lost during such irregular operation.

Various solutions for recovering from a disruption have been presented, but none so far have tried to better characterize the schedule and redefine the schedule planning approach to minimize this impact of disruptions. The idea of robustness and the definition of the SRP can help to relieve such a problem.

The introduction of the small building blocks in a plug-in framework allows the narrow definition of small parts of the schedule design process. The notation introduced by Antes (1996) is very useful for defining new problems and the interaction between them. The rearrangement of such problems can generate new problems to solve with different objectives in mind. In the end, the use and definition of building blocks for the schedule design process allows the classification and comparisons of different approaches. It facilitates the interpretation and discussion of strategy, problem description, models, and algorithms.

Chapter 3

Profile Analysis

An analysis of an airline's performance is the initial step in understanding how to improve its schedule. Evaluating the schedule and considering its base components, the flights and flight legs, is a starting point for introducing robustness in to the system.

After an evaluation of a generic schedule, we proceed with a discussion about the schedule reliability characteristics. These characteristics will be summarized into numerical value representing airline performances. These values are the flight **Option Value (OV)** and the **Option Disruption Value (ODV)**. The Option Disruption Value, in particular, will have a major role in the flight reliability measurements, as will be outlined in Chapter 7.

In this chapter, the network independence concept and the time profile behavior as a measure of flight leg reliability are introduced. The subdivision of a flight leg trip in its base components, gate time, taxi-out, trip time, and taxi-in time, allows a full characterization of the flight leg performance. For each of these key components the time profile has been studied, implemented and analyzed under various hypotheses. The final time profile coming out from the performance of each station has been generated from the scheduled service flown in the domestic US in July 2000 (the process can easily be extended for any month).

A detailed description of the process used to identify and determine the profiles will be presented. Furthermore, an analysis of the algorithm used is offered. In the end of the chapter there is an example profile calculation for a flight leg (EWR-DEN). At the concluding part there is the introduction, via simple statistical considerations, of the concept of **Flight Leg Slack Time**: this concept will have a central role in the calculation of the OV and will be discussed in detail in the next chapter.

3.1 Evaluation Process

The evaluation of a schedule is the first step in solving the Flight Leg Adjustment Problem (FLAP), the Multiple Flight Allocation Problem (MFAP), the Multiple Station Redundancy Problem (MSRP) and the Station Isolation Problem (SIP). Given the output of the Critical Flight Selection Problem (CFSP), the Critical Station Selection Problem (CSSP) and an existing schedule, the analysis of its coverage and robustness is the initial aspect to consider for improving the original schedule and then to provide effective solutions to the Schedule Robustness Problem (SRP).

The first observation is that, given a critical flight or station, if the disruption is occurring at the final destination, there is very little that can be done. The same can be said if it is the origin station that is disrupted. No matter where the passengers are coming from, they will not be able to reach their destination easily. If the flight in consideration is composed of multiple legs, it will be considered that the disruption occurs on the first leg; when it occurs in an intermediate leg, then the sub-trip starting from the first disrupted leg can be analyzed.

Non-stop flights are excluded from this dissertation for 3 reasons:

1. The protection of a direct flight can be effective only through another direct flight. The disutility imposed on passengers to accept a protection flight option having any stop is greatly reduced, compared to the disutility of another direct flight departing even hours after the one delays/cancelled. This will be highlighted when comparing the OV of a flight and a flight leg in the next chapter.
2. Delay and cancellation affect more passengers flying on connecting flights than passengers traveling on direct flights, because a cancellation or a delay of a leg on a connecting flight can cause a connection to be missed, and hence it can generate a so-called **knock-on delays** (cascade of delays) situation.
3. Non-stop flights usually serve very competitive and highly valuable markets, and hence passengers are more prone to choose them, even if provided by a competing airline, than a connecting one.

Given any flight f , its **flight protection options**, or FPO set, is a set of flights that are leaving the origin station near the departure time of flight f . The number of options available to f as a protection can run up to hundreds. They can, however, be narrowed down to very few by considering the arrival time of flight f . Indeed, passengers are very

interested in arriving at a destination as close as possible to the original arrival time independently of the route followed and the departure time. In such a case, if the arrival time window considered is of four hours centered on the original arriving time ($t \pm 2$ hour), then the number of options is reduced to around 4 to 6 flights or options depending on the market considered. These are the key protecting options of flight f : the higher this number is then higher the degree of protection. Furthermore, if a narrow time window is considered and the number of option does not decreases then a higher degree of protection is offered as well.

Notice that the time window is not necessarily symmetric on the original arrival time: it may be interesting, for example, to consider a window of two hours before and one hour after the arrival time, depending on flight characteristics and passenger preferences.

The options obtained have to be evaluated based on:

1. Their effectiveness.
2. Their availability (e.g., enough seats).
3. The protection level reached by the flight to safeguard.

The goodness of an option is determined by analyzing its characteristics. The important aspects are:

1. Physical characteristic of the flight such as number of stops, ground time and flight time.
2. On-time performance, connectivity probability, and cancellation rate.
3. Seat availability.

These data are combined to give an **Option Value (OV)**: the higher this value, the less appealing is the option. If this is compared with the option value associated with the flight to protect f , by including also the difference between their respective arrival times, the result is the **Option Disruption Value (ODV)** of the option in consideration.

The availability is translated into **Option Seats (OS)**: the number of opportunities is related to the available seats for such an option. The Option Disruption Value associated to every Option Seat is identical to the one associated with the flight protection option considered.

The **Flight Protection Level (FPL)** is fixed by the airline and it is based on the value of the flight to protect. The FPL is quantified in a combination of FPO/OS and ODV. A proper protection level is reached if FPO/OS are adequate and if OPV matches

a given value or distribution. An example of FPL is 5 FPO in 2 hours or 120 OS in 1 hour. An example of ODV is no more than 10% for all the options available.

When the proper FPL requested is not reached, then a model for increasing FPO/OS and reducing ODV has to be developed. This can be achieved:

1. If FPO/SO matches, but ODV level is too high then it can be lowered by shifting the in time the existing FPOs.
2. If FPO matches but OS does not, then, for example, the introduction of larger/smaller aircraft can increase/reduce the number of seats available.
3. If FPO does not match, then it is proper to introduce new flight options.

These considerations can be obtained by evaluating the characteristics of the various options; however, while doing this the delay profile and the cancellation rate presents some problems.

3.2 Delay and Cancellation Considerations

The option value is based on the physical characteristics of the flight, the on-time performance and the seat availability. The physical characteristics such as departure time, flight time, connecting time, arrival time and stops are all determined when the option is identified. The seat availability is obtained from historic data or via a demand modeling approach. The on-time performance is determined via the profile analysis.

The characteristics of every option in terms of delay distribution and cancellation rate can be derived from historic data analysis. It is, however, difficult to analyze a single flight because data for the flight may not be available over a long enough time window to develop consistent distributions.

In the extreme case, if a new option has to be introduced, there is no distribution for its delay performance and cancellation rate: this implies that distributions and rates developed for existing options cannot be used for *new* options. Having the distributions for existing options, and, instead, estimating distributions for *new* options, cause diverse level of analysis accuracy between flights. In particular it will not be possible to meaningfully compare new and old options at the same time. It is proper, therefore, to use aggregate quantification of delay performance and cancellation rate. This is obtained by the **profile analysis**. This means that only flights departing from the origin station and flights arriving to the destination station of flight in consideration are examined. In this way the origin and destination airport are considered to be gates in the

network. The network, therefore, will shape the performance of the option in consideration.

Taking into account every leg in a fixed time window and analyzing, for each of them, the departure and arrival performances of the origin and the destination station, determines the departing and arriving performance profile, respectively. The time window is centered (symmetrically or asymmetrically) on the departure and arrival time of the leg f being considered: this defines a set of flights referred to as the **flight neighborhood** of f .

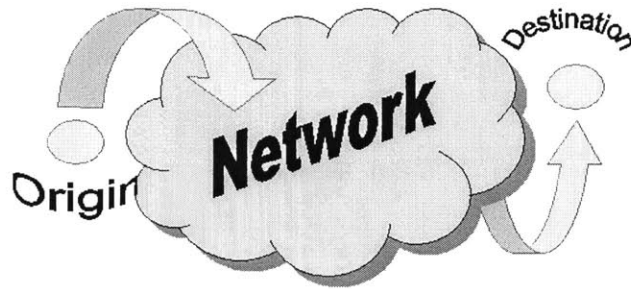


Figure 3.1 Network as a Source of Disruptions

The profile obtained considering all the flights leaving the origin station and the profile of all the flights arriving at the destination station, will be used to characterize the arrival and departure profile of the option being considered. These values are to characterize all the options being considered, thus putting the existing options and the new options (**shadow options**) at the same level.

3.3 Station Independence

The key assumption described in the previous section is that the behavior of a flight/flight leg between two stations will follow the same behavior as all the flights leaving the origin station and arriving at their destination in their respective departing and arrival time.

For example if there is a need to put a flight f between two stations, say Boston (BOS) to Chicago O'Hare (ORD), it is assumed that the departure behavior of this flight will follow the same behavior of the flights departing from BOS, in, for example, the hour before and after the scheduled departure time of f . The same reasoning can be applied to the destination station: the flight in consideration will arrive, and therefore will have the

same performance and delay, as the various flights arriving in the two hour window centered on the scheduled arrival time of the flight f .

This reasoning leads to a **leg independence and a station dependent analysis**: i.e. knowing the behavior of the flights at the origin station and the flights at the destination station is enough to predict the on-time performance of the flight being considered. An analysis of this type has several advantages. The first and most important of them is that the flight being considered does not need to exist in the current schedule: in fact, once the departure time and the arrival time are fixed, the profile of the flight's departure and the flight's arrival are determined. A second benefit is that it is easier to add a flight in any desired time without considering any major restriction: the profiles are sufficient to determine the behavior of the new flight. This last point allows us to predict the performance of any flight and to *successively adjust* the departure and arrival time until its behavior matches a desired one.

Because the flight profile calculation considers the origin and destination separately, then the calculation of any conditional probability of cancellation or delay between multiple legs, becomes easier: the profiles can be considered independently when evaluating a single flight.

The component that connects the departure and arrival profile is the profile of the flight time. From data in the ASQP database, it is possible to estimate a distribution of flight times.

A single flight leg, therefore is divided into 3 time profile parts:

1. The departure profile.
2. The coupling profile.
3. The arrival profile.

The departure profile is further divided into the **Gate Profile** and the **Taxi-Out Profile**. The first shows the behavior of a set of flights that want to leave the gate in a specified time window. This model gives the delay in departure, and also can be used to quantify the probability of connecting passengers *making their connection*. The second is the profile that the flight legs will have after leaving the gate until they take off (wheels off time). This will include *Ground Delay Programs* (GDPs), queuing and delays due to departing airport conditions.

The coupling profile is the one that allows a full characterization of what is happening on route. This is generated considering the flight time of all the flights connecting the two station of the flight in consideration, in a fixed period of time (e.g., one month). This

coupling profile, also called **In-Air Profile**, is, therefore, the time profile of a generic flight leaving the origin and reaching the destination during that month.

The arrival profile models the behavior of a flight between touch down and reaching the gate (de-boarding operation are not considered). For this reason this profile is called the **Taxi-In Profile**.

Another useful property of station independence is the cancellation rate. Given the origin station and a fixed time window, there are two cancellation rates: the **Departure Cancellation Rate** and the **Arrival Cancellation Rate**. The first one is calculated using the departing flight in the specific time window; the second is obtained using the arriving flights in the considered time window.

These four profiles and the cancellation rates can fully describe any flight from origin a to destination b independently of its existence and its scheduling time. Recombining the various profiles and cancellation rates gives the behavior of not only a flight leg but also a flight, and then of an itinerary, as well. As a result it is possible to obtain the final profile of the duration of the flight or flight leg and, expressed in probability, if, how and when it will arrive late.

It is possible to consider options that are not actually present on the current schedule. These options, called **shadow options**, can be introduced into the system when and if they improve the coverage and/or the robustness of certain markets. This can be integrated into the model; the basic characteristics of the new flight are determined a priori: departure time, the number of stops, fly time, and ground time can be empirically calculated. The departing time and arrival time should also consider limitations like slot and gate availability. The arrival time is simply obtained considering departure time plus the fly time. The new options provide all their capacity to the system: these new flights are, therefore, included into the schedule, and hence will be used by the flying public. After considerations of load balancing between old and new options the final number of seat options introduced by the new flight is lower than their total capacity; nevertheless, more seat will be available for passenger relocation on the whole network.

While the basic information is empirically determined, a more detailed analysis of delays and cancellations is needed. The on-time performance is obtained by the reliability performance of the stations touched by the shadow option. Every station, in fact, given any time window determines the departure, coupling and arrival profiles. These options, therefore, can be handled exactly like any other existing flights.

3.4 Passenger Choice and Profile Hypothesis

The idea of using the network as a source of disruption is not new. A similar approach to the one presented here has already been applied to the public transportation network in the Netherlands as described in Bruinsma (1999) and Rietveld (2001).

In both papers there is an outline of how the delay impacts a passenger's trip. Results from these studies indicate that passengers are very risk averse and that passengers take into consideration the odds of delays while planning their trips. The passenger's decision is highly impacted by the odds of so-called non-recurrent and unforeseeable delays. The presence of recurrent delays, instead, has a marginal effect in their choice.

The passenger adversity to non-recurrent delays can be described by the fact that unreliability, as related to their trip time, increases the chances of them missing their connections. In fact, potential travelers make their choices so as to reduce the chance of an unreliable movement. Information about reliability is an important factor in the passenger decision process. However, most people make transport decisions with incomplete information. One might argue that the regular traveller is well aware of unreliability. However, he or she usually has little information about alternative routes or trips using other carriers or transport modes.

This is particularly important in itinerary-based decision-making when multiple carriers, multiple modes or multiple vehicles need to be considered. In these papers there is specific reference to transportation chains that can be translated to the airline business as multi leg flights or itineraries.

High levels of un-reliability can greatly affect the operation of hub-and-spoke systems. A trip is considered unreliable in cases of positive delays (late arrival) as much as negative delays (early arrival). In various studies there is a strong indication that early departure is an inconvenience (even if less than late arrival) and not a benefit.

Travel time itself can be divided into 3 categories

1. Official travel time.
2. Actual travel time distribution.
3. Perceived travel time distribution.

The *official travel time* is the travel time according to the published timetables. The official travel time may be defined as the travel time needed given an ideal condition in which there is no congestion, or no major disruption.

The *actual travel time distribution* represents the travel time outcomes of actual trips. Delays may occur due to a large number of factors. In addition, there may also be travel times that are (slightly) shorter than the official travel times.

The *perceived travel time distribution* refers to travel times as perceived by the traveller. For experienced travellers, the distributions of actual and perceived travel times may be about equal. However, when travellers do not have experience or lack information about a particular trip, the two distributions may be quite different.

Departure times are usually strongly asymmetric. There may be some probability of early departures, but the probability of late departures is usually much higher. Therefore, natural candidates for departure time distributions are the Gamma, Lognormal and Weibull distributions. These are all characterized by a curve where the mode is clearly smaller than the median and the mean travel time. These distributions are, at the left hand part of their range, characterised by an increasing probability of departure per time unit (given that they have not yet departed), followed by a decreasing probability of departure at the right hand part of their range.

The distributions above-mentioned have a common feature in that they have a unique mode and finite expectation and variance. Whereas the standard gamma, lognormal and weibull distributions are all defined on the positive half axis, there is the introduction of a third parameter to all distributions; the location parameter θ . This parameter is chosen such that the difference between the data and the distribution is minimal. If the shape parameters are correctly chosen (i.e., smaller than 1), the modes of the gamma and the Weibull distribution are not equal to the minimum value. For the lognormal distribution this is always the case. In practical situations this is often a desirable property both for departure and arrival distributions, whereas it may allow early departures and arrivals.

The formal expressions for the distributions is:

$$\text{Lognormal: } f(x) = \frac{e^{-\frac{1}{2} \left(\frac{\log(x-\theta)}{m\sigma} \right)^2}}{(x-\theta)\sigma\sqrt{2\pi}} \quad \text{for } x \geq \theta \quad (3.1)$$

$$\text{Gamma: } f(x) = \frac{\left(\frac{x-\theta}{\beta} \right)^{\gamma-1} e^{-\frac{x-\theta}{\beta}}}{\beta \Gamma(\gamma)} \quad \text{for } x \geq \theta \quad (3.2)$$

Weibull:
$$f(x) = \frac{\gamma}{\alpha} \left(\frac{x-\theta}{\alpha} \right)^{\gamma-1} e^{-\left(\frac{x-\theta}{\alpha}\right)^\gamma} \quad \text{for } x \geq \theta \quad (3.3)$$

Having a choice of distribution enables a near perfect description of the various times. In Carey (1994), the model for the arrival time is based on the simple formula of a linear combination of the departure time and of the trip time:

$$t_a = t_d + t_t \quad (3.4)$$

There is an implicit assumption that the covariance between the deviation from the scheduled departure time and the deviation from the scheduled travel time is equal to zero:

$$\text{Cov}(t_d, t_a) = 0 \quad (3.5)$$

This assumption is made because it might be the case that there are no paired data available on departure and arrival time deviations. At first sight the assumption may seem too restrictive, but it is important to realize that a zero covariance is not equivalent to independence of the stochastic variables. It is highly possible that opposite effects cancel each other out: on one hand, the pilot may be able to recover the departure delay via speeding up during the trip, on the other hand, it often happens that delays accumulate.

The conclusion of Bruinsma (1999) and of Rietveld (2001) shows that the Lognormal distribution is a good fit for such times. The description of the trip times is used for justifying changes in the schedule that increase reliability; It is clear that if the reliability is increased a better service is provided, hence the tariff can be increased as well.

3.5 The RSM Approach and Theoretical Justifications

The approach presented here follows the concept presented in the above papers but with some minor changes. Instead of considering the arrival time as a simple combination of trip time and departure time, the arrival time is a combination of gate time, taxi-out time trip time and taxi-in time:

$$t_a = t_{gt} + t_{to} + t_{tr} + t_{ti} \quad (3.6)$$

The other difference is, instead of outlining the profiles using the full set of trips serving the same market, we isolated the various stations and analyzed the performance

of the single location for all the flight legs departing or arriving in a fixed time window (**flights neighborhood**). This gives more strength to the zero covariance assumption.

Given such a situation, the same distributions have been analyzed and two methods of estimation have been used for their parameters: the first method is the *Best Fit Estimation*; the second is the *Maximum Likelihood Estimation*. Both methods gave good results, but the second is computationally less demanding.

The gamma and lognormal distribution have been used and both the Kolmogorov-Smirnoff and the X^2 goodness fit test has been applied.

The distribution of choice has been the lognormal for the gate, taxi-out, and taxi-in time. The normal distribution has been used for the trip time.

EXAMPLES OF LOGNORMAL FAILURE PDF'S

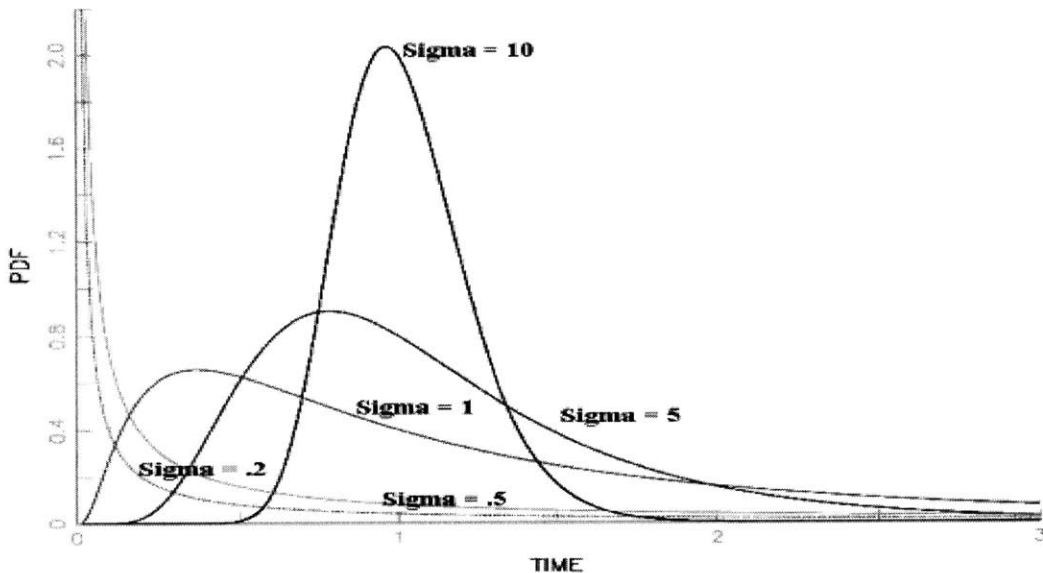


Figure 3.2 Lognormal Probability Distribution Function Chart

Both the gamma and lognormal distributions had a very good fitting test; however the lognormal is flexible enough to make it a very useful empirical model. In addition, the relationship to the normal (just take natural logarithms of all the data and time points and you have "normal" data) makes it easy to work with mathematically, with many good software analysis programs available to treat normal data. The lognormal model can be theoretically derived under assumptions matching many failure degradation processes common to various failure mechanisms.

A brief sketch of the theoretical arguments leading to a lognormal model follows. As Kolmogorov states "Applying the Central Limit Theorem to small additive errors in the log

domain and justifying a normal model is equivalent to justifying the lognormal model in real time when a process moves towards failure based on the cumulative effect of many small "multiplicative" shocks. More precisely, if at any instant in time a degradation process undergoes a small increase in the total amount of degradation that is proportional to the current total amount of degradation, then it is reasonable to expect the time to failure (i.e., reaching a critical amount of degradation) to follow a lognormal distribution" (Kolmogorov, 1941).

In a more formal way we can describe this distribution from the following situation: let y_1, y_2, \dots, y_i be measurements of the amount of degradation for a particular failure process taken at successive discrete instants of time as the process moves towards failure. Assume the following relationships exist between the y 's:

$$y_i = (1 + \xi_i)y_{i-1} \quad (3.7)$$

where the ξ_i are small independent random perturbations or **shocks** to the system that move the failure process along. In other words, the increase in the amount of degradation from one instant to the next is a small random multiple of the total amount of degradation already present (accumulation of delay, degradation of airport performance, weather perturbation on the trip, etc). This is what is meant by **multiplicative degradation**. The situation is analogous to a snowball rolling down a snow-covered hill; the larger it gets, the faster it grows because it is able to pick up even more snow.

It is possible to express the total amount of degradation at the n -th instant of time by a multiplication of all the previous perturbation, back to the first one. This first one is a constant perturbation that is continuously pushed by small random shocks. Since failure occurs when the amount of degradation reaches a critical point, the time of failure will be modeled successfully by a lognormal for this type of process.

3.6 The Profiles: Method and Analysis

A set of automatic tools have been designed and implemented to isolate the profiles. A thorough analysis of the schedule for the major US airlines in the US domestic market has been performed. The month analyzed was July 2000, and a fitting process for the various distributions was determined. A time window of 2 hours centered on the departure and arrival time was used. This time window defines the set of neighborhood flights. Considering the variability of the data, the following process was used:

1. The real histogram was generated a priori, with real value of mean and variance. The histogram represents how many flights spend n minutes in the relative portion of their trip (gate, taxi-out or taxi-in from their scheduled departure time).
2. The histogram was truncated at the 150th minute (everything that spends more than 2 hours and 30 minutes is considered too extreme to be evaluated).
3. Extreme values located at minute -30^{th} and 150^{th} were removed. In this way the data is not affected by enormously delayed flights.
4. If less than 4 observations were available, then the mean and variance recorded was the mean and variance of these observations (different from the whole sample, and of the limited histogram).
5. If there were more than 3 observations, and if the observations returned the same value (multiple repetition of the same result), then there was no distribution and the mean was the single sample value, variance is 0.
6. If in the sample, the distance between the last element and its predecessor was greater than the distance between $n-1$ -th and $n-2$ -th and the distance of $n-2$ -th and $n-3$ -th, then the last observation was removed as well (too far from the bulk of the sample).
7. The estimation of the distribution was done through MLE and the data-shift parameter θ was determined such that the distance between the data and the shifted distribution was minimal.

This process utilized the ASQP database for the month of July 2000, for the gate, taxi-in and taxi-out profile. The automatic tools put in place performed the above procedure for a time step of 10 minutes starting at midnight (00:00, 00:10, 00:20... 23:50). The histogram was generated for 2-minute interval time steps starting from -30 up to 150 minutes ($-30, -28, -26 \dots 148, 150$). The calculation of the flight neighborhood spanned across days as well. This process is automatic and can be applied to any ASQP file. The calculation of the specific time profile was done via linear interpolation of the two closest profiles (e.g., for the 8:15am profile it interpolates the 8:10am and 8:20am profile). A generic visualization tool was also developed, as reported in Figure 3.3.

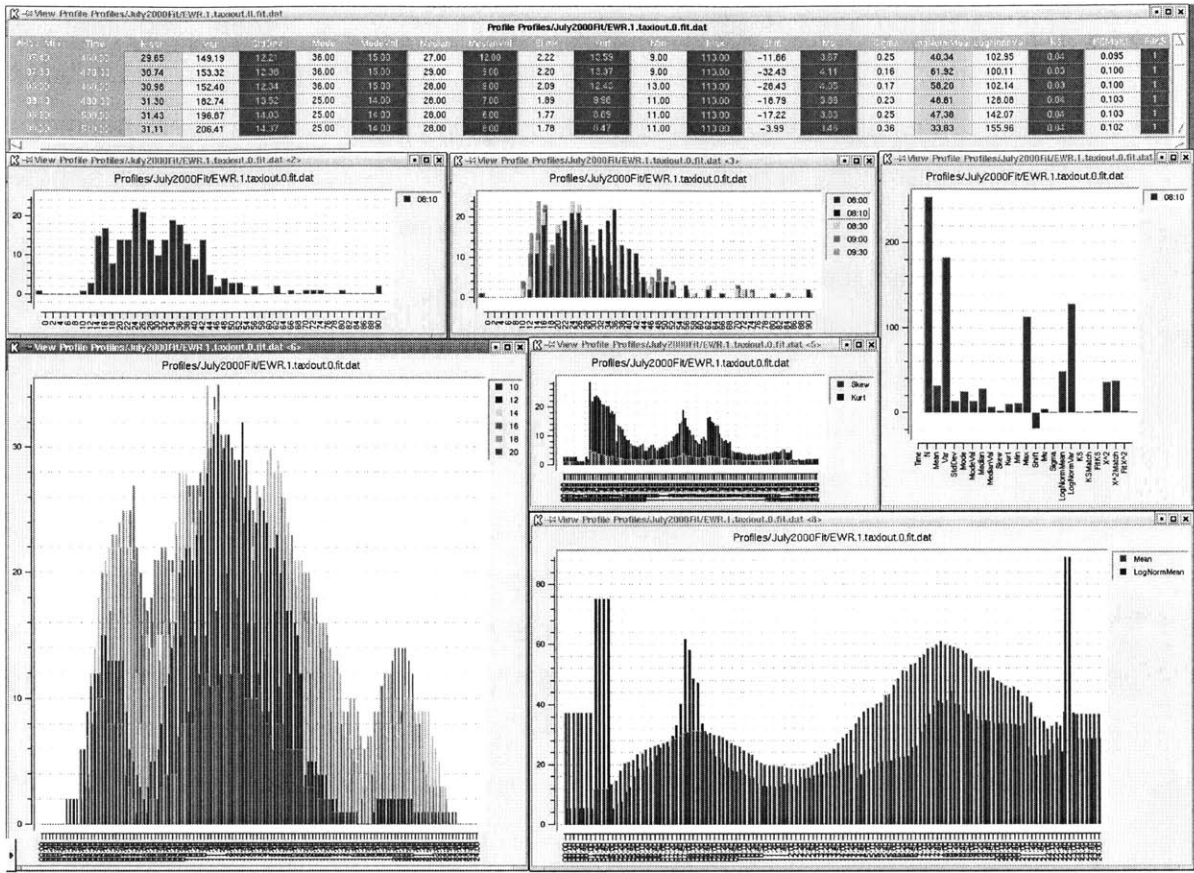


Figure 3.3. Visualization Tool for Taxi-out Time in EWR July 2000. From Top Clockwise Till the Center: Data Spreadsheet. Profile Descriptive Statistics at 8:10am. Standard Mean and Lognormal Mean of Data Across the Day. Number of Flights Spending 10, 12, 14, 16, 18, and 20 Minutes in Taxi-Out Time. Taxi-Out Histogram at 8:10am. Taxi-Out Time at 8:00am, 8:30am, 9:00am and 9:30am. Distributions of the Skewness and Kurtosis of the Profiles Along the Day.

In the process, the removal of the last element of the very heavy tail (point 7) on the gate profile was very effective. It created a general improvement of 0.77% points on the general fit test, but some stations were greatly improved. LGA, for example, improved from a 61% to a 65% profile fit. Figure 3.4 depicts the most affected stations.

The process described above has been applied on the schedule available and, even if in taxi-out and taxi-in time there is no noticeable difference, the lognormal distribution fits 4% better than the Gamma distribution on the gate profile. Across the board, the lognormal fits the gate profile 85%, the taxi-out profile fits 95.5% and the taxi-in profile fits 95.6% of the time.

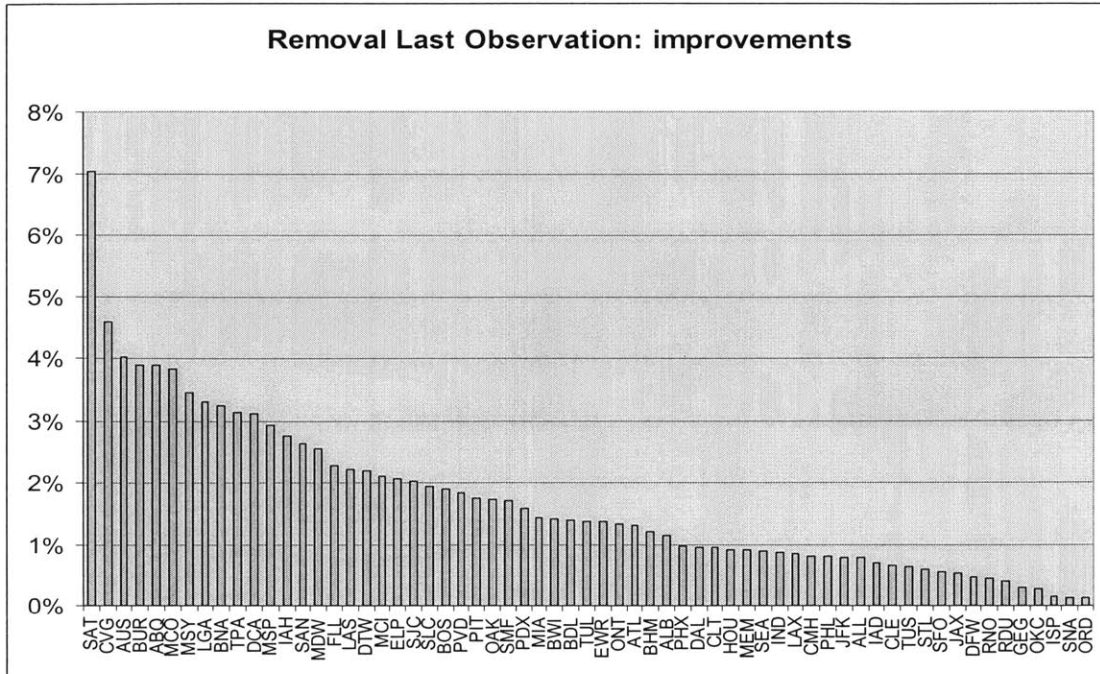


Figure 3.4 Removal of Last Observation

The process used for the trip time is analogous to the previous process for departure time, but no points were removed. The trip distributions have been recorded centered on their calculated mean. The visualization tool gives insight on the distributions as reported in Figure 3.5.

The normal distribution fits the trip time distribution on 80% of the samples.

The profile distributions have been tested against the Kolmogorov-Smirnoff and X^2 goodness fit test with a confidence level of 99%.

The X^2 test gave positive result where the Kolmogorov-Smirnoff test was positive; however, the sparsity of the data denied its applicability in certain situations.

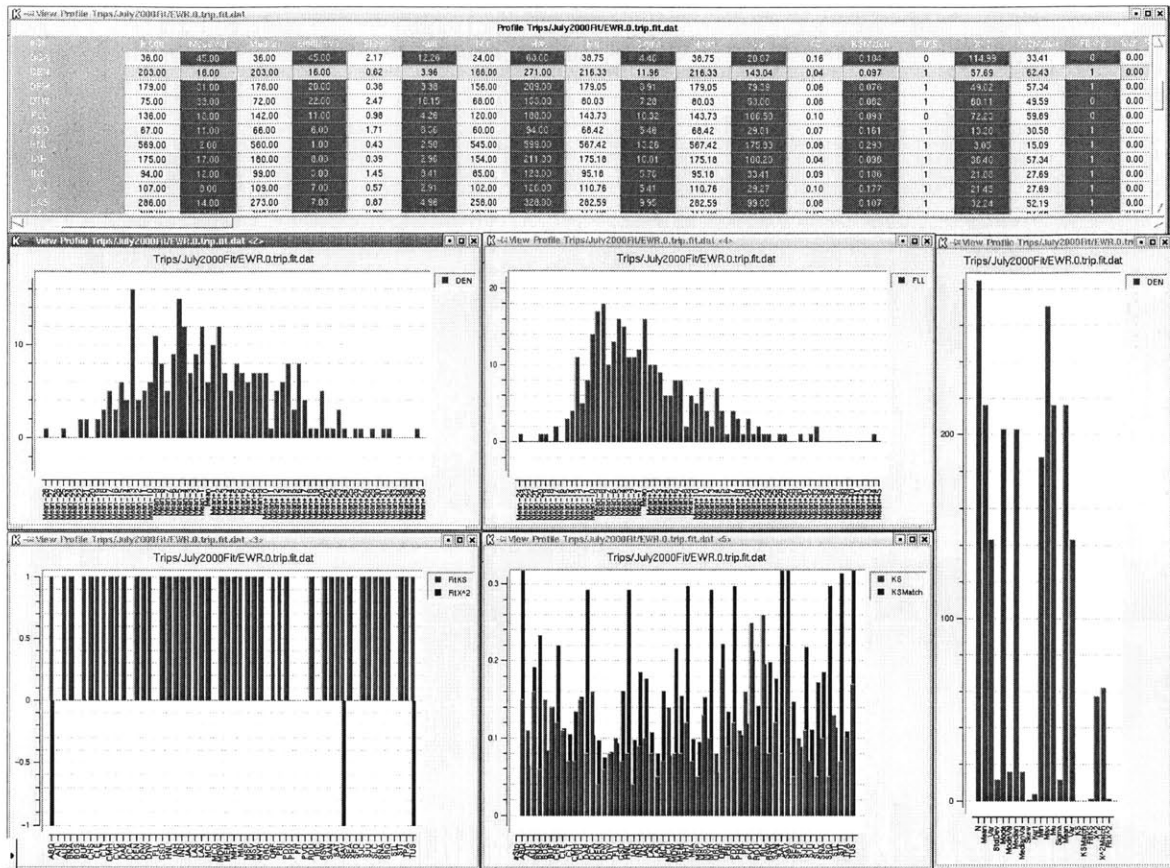


Figure 3.5. Trip Profile Visualization July 2000 Newark (EWR) to All Stations. From Top Clockwise to the Center: Data Spreadsheet. Descriptive Statistics of EWR-DEN Trips. Comparison of Kolmogorov-Smirnoff Values with the Critical Ones. Fit Test Results for the KS and χ^2 Test on All Trips (Fits=1, No Fit= 0, No Enough Data = -1). Trip Profile for EWR-DEN (Normality Fit Test is Acceptor). Trip Profile EWR-FLL (Normality Fit Test is Rejected).

The profiles with the lowest percentage of fits are the gate time (only 85% of the time) and the trip time (80% of the time). The situation in which there is no fit occurs because:

1. There are too few observations: only a few intervals had positive response.
2. The distribution is severely impacted by extreme values. In this case there may be a spike in the tail of the distribution (gate case) or the normal distribution is too flat (trip case).
3. There is multi-modality: the distribution of the events has two or more peaks. This goes against the uni-modality assumption of the Lognormal and Normal distribution.
4. There are too many observations around the mode of the distribution: this peak, even if positive, causes the KS test to fail.

The following charts show the stations with the lowest percentage of fits for the gate time (left) and trip time (right). The trip times highlighted are only for the stations that are connected with more than 10 other stations.

Notice that, based on the previous assumptions the above distributions are the best possible fit for the available data.

LogNormal Fit on Gate Profile: Highlight on <70% Fit

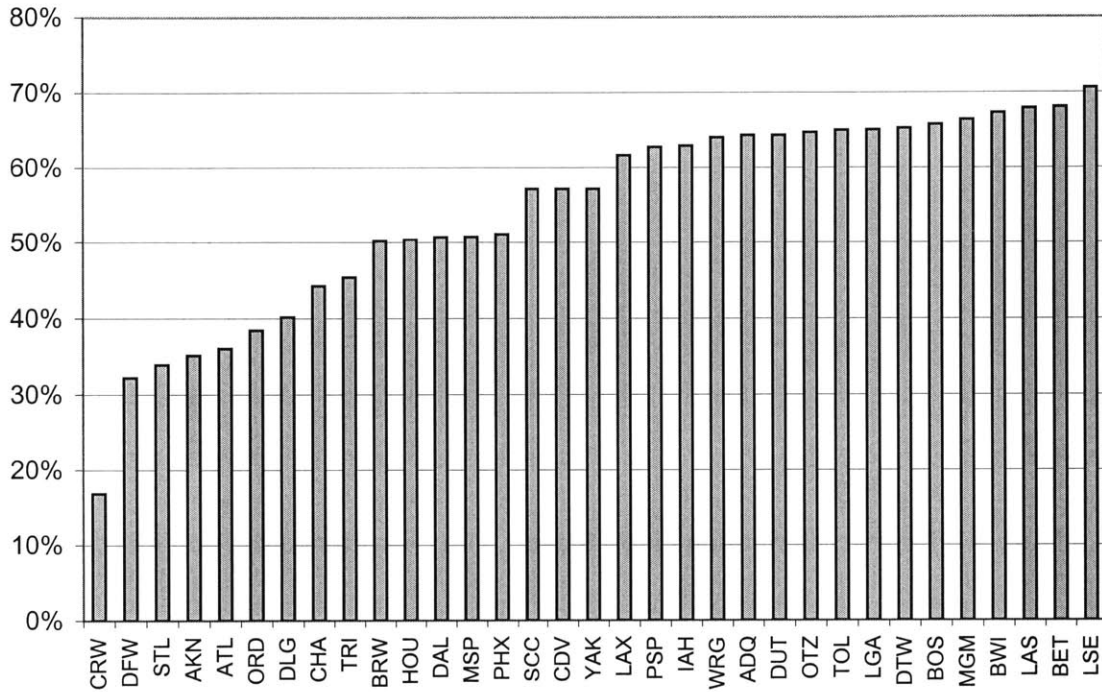


Figure 3.6 Lognormal Gate Fit on Station With Less Than 70% Profiles Fit

Normal Distribution Trip Time

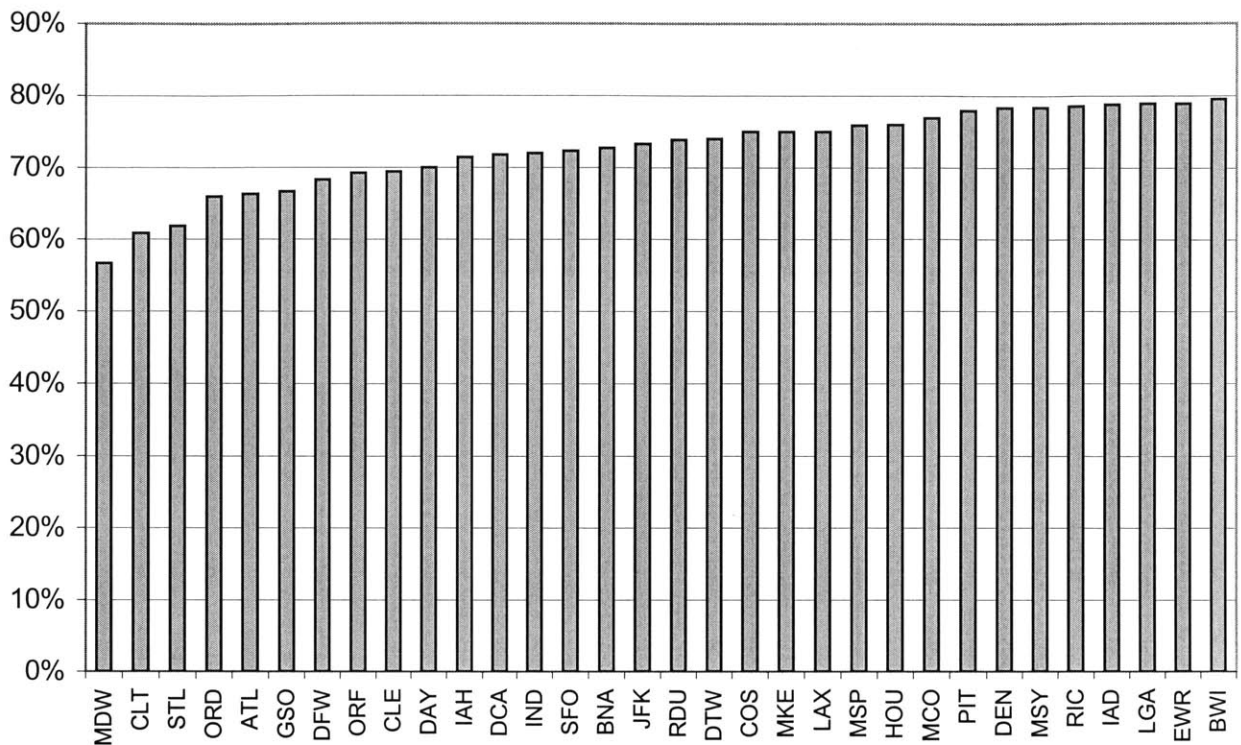


Figure 3.7. Normal Distribution Fit on Station With Less Than 80% Profiles Fit

3.7 Profiles: Example

Below, we present an example of a profile on the flight leg Newark (NY) to Denver scheduled in July 2001 departing on Mondays at 8:10 am and arriving at 10:30am.

Considering all the flights leaving Newark between 7:10am and 9:10am during the Mondays of July 2000 the departure profile is determined. In this 2-hour window 254 flights were flown (roughly 50 every Monday).

The connecting profile is obtained by analyzing all the flights connecting Newark and Denver in the month of July 2000. There are a total of 285 flights going from EWR to DEN.

The arriving profile is derived from the analysis of 184 flights arriving in Denver between 9:30am and 11:30am during the Mondays of July 2000.

The departure profile is divided into gate profile and taxi-out profile. The chart of the gate profile is in Figure 3.8. As depicted by the chart, most of the flights leave the gate

between -10 and +10 minutes of the schedule departure time. The distribution of the data fit the lognormal distribution very well, with parameters $\text{LogNorm}(2.30, 0.87)$. The shift parameter is -8.33 minutes. On average the flights leave 6 minutes later than the schedule departure time, with a variation of 15 minutes.

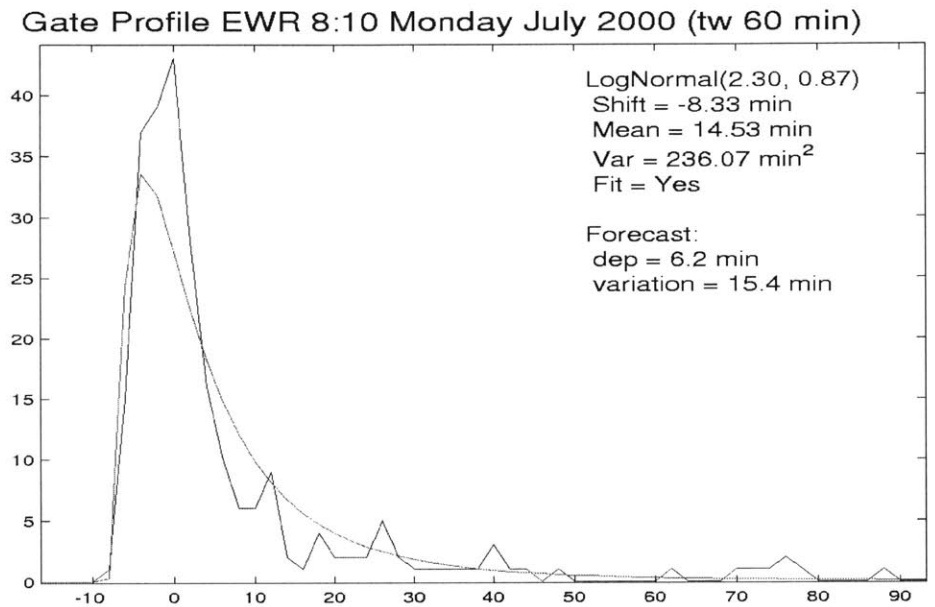


Figure 3.8 Gate Profile EWR 8:10 on Mondays July 2000

The taxi-out profile for the same set of flights is reported in Figure 3.9. Most of the flights spend between 10 to 50 minutes in taxi-out time. Even in this case, the lognormal distribution fits the data well. The parameters are $\text{LogNorm}(3.83, 0.25)$. The shift parameter is -17 minutes. This implies that the average taxi-out time is 30 minutes with a variation of 12 minutes.

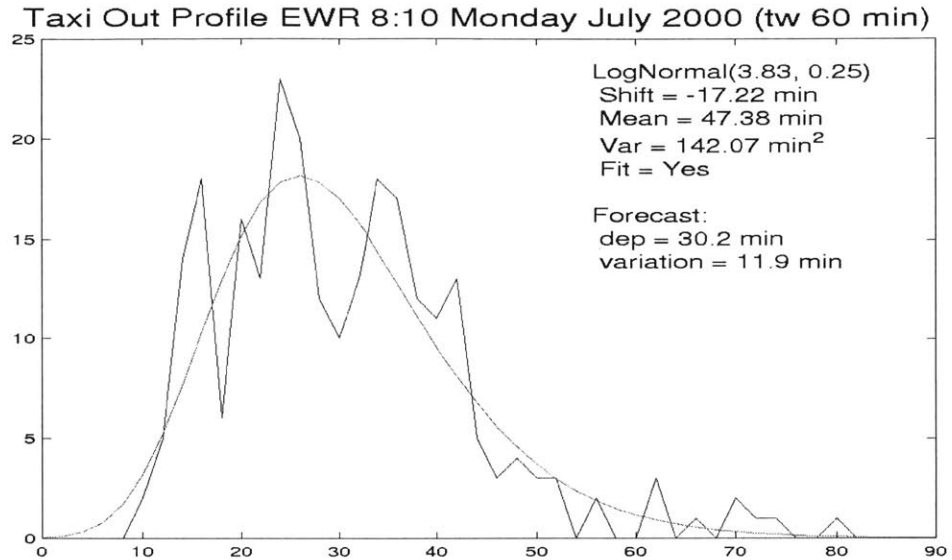


Figure 3.9 Taxi-out Profile EWR 8:10 on Mondays July 2000

The flight time profile appears to be normally distributed. The trip time profile chart is shown in Figure 3.10. The normal distribution fits both the Kolmogorov-Smirnoff, and the X^2 test. There is no distinction between equipment types when computing the profile behavior. This may be the reason for the multiple spikes in the chart. Nevertheless, the test indicates a good fit for a normal distribution of parameters Normal(216.33, 11.96). This implies an average trip of 216 minutes with 12 minutes of variation.

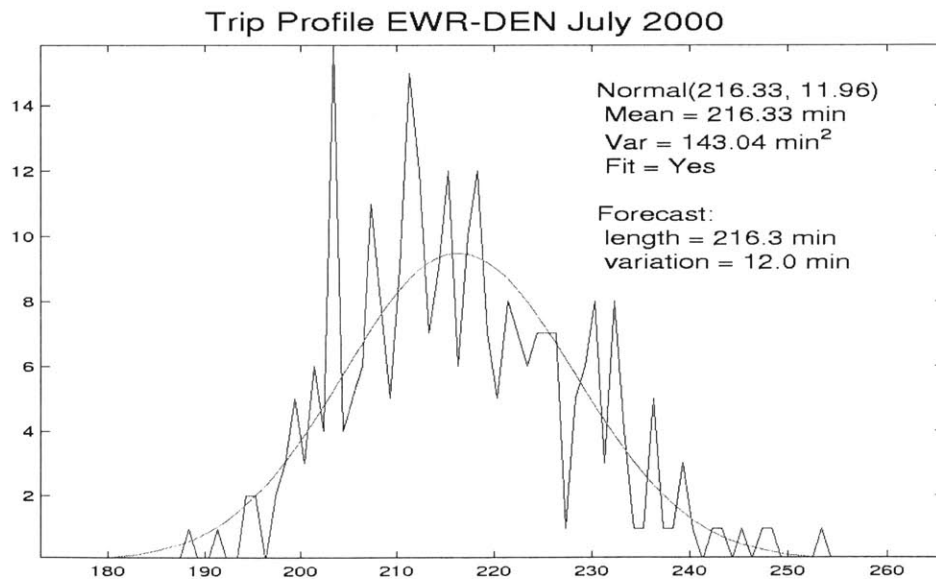


Figure 3.10 Trip Time Profile Newark Denver (EWR-DEN) on July 2000

The taxi-in in Denver around 10:30 am is depicted in Figure 3.11. The log normal distribution fits these data as well. The parameters are LogNorm(1.82,0.77), with a shift of 2 minutes. The average taxi-in time is of 10 minutes and the variation is 8 minutes.

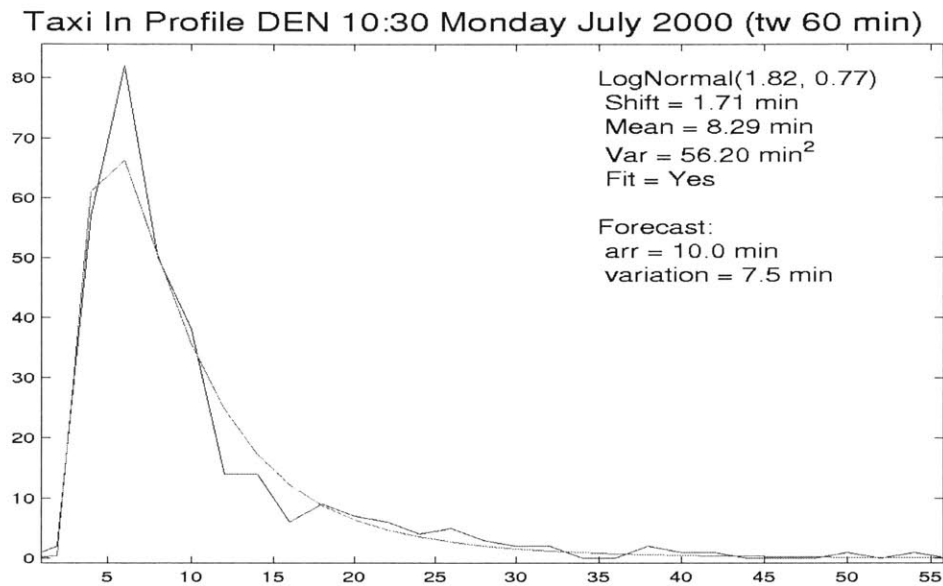


Figure 3.11 Taxi-in Profile DEN 10:30 on Mondays July 2000

From the above data, it is possible to summarize some observation related to this flight. The scheduled time of this flight is a 260 minutes (3:20 hours). If the flight follows the means of the profiles exactly, it should, more realistically, take

$$30 + 216 + 10 = 256 \text{ min} \quad (3.8)$$

of trip time and should depart on average 6 minutes later that its schedule departure time.

The time difference between the scheduled trip time and the effective trip time is positive 4 minutes. This time, called **Flight Leg Slack Time** is inserted a priori into the schedule for dealing with variability and unpredictability of events occurring in the network. From this data, it appears that the scheduler overestimated the performance of such a flight: on the average the flight will be always 4 minutes shorter in trip but will arrive

$$(256 - 260) + 6 = 2 \text{ min} \quad (3.9)$$

late at destination.

This flight slack time is obtained considering only the average trip time. In reality the variation coming out of the profile is

$$\sqrt{15^2 + 12^2 + 12^2 + 7^2} = 24 \text{ min} \quad (3.10)$$

If it is assumed that the arrival time will be normally distributed and it is required that 82% (the $\mu + \sigma$ percentile of any normal distribution) of the flights arrive on time, then the scheduled flight arrival time should be of

$$256 + 24 + 6 = 286 \text{ min} \quad (3.11)$$

after the scheduled departure time. This difference in slack time

$$286 - 260 = 26 \text{ min} \quad (3.12)$$

will assume a great importance as illustrated in Chapter 5 and 6.

3.8 Conclusions

In this chapter the concepts of station independence and of the network as source of disruption for airline operations were presented as a means of characterizing network performance.

A thorough analysis of the key components of a flight leg was also performed. The results confirm what is found in the literature: the time profile describing the flight leg performance fits a lognormal and a normal distribution most of the time.

These are the key components for defining a reliability measurement. Defining a proper way of grading the various options will be key for driving the adjustment of the schedule. It is clear, however, that this single process is not enough for understanding the whole benefit to the airline operations: integration with the crew, rotation and fleet assignment and an analysis of the impact on the other part of the airline system are needed.

In the next chapter, we introduce, the Option Value, as a flight reliability performance indicator, and a quantification of the Option Disruption Value as a way of selecting possible options for improving the schedule at hand.

Chapter 4

Scheduling Problem: Basic Definition and Ideas

In this chapter, we formally introduce the elements composing the schedule and the basic ideas behind reliability and how to put it in place.

The chapter is divided into two parts. The first presents the key element of a schedule: notions of market, legs and flights. A formal definition of the set of options and how to isolate them given a schedule and a generic flight is also presented.

The second part presents ideas on how to take advantage of the options available during the process of passenger relocation. This is a schematic view that will be fully described in the last chapter of this dissertation.

4.1 Basic Definitions

The first entities handled during the schedule planning process are the basic elements composing the airline network: stations, flights and flight legs. The next definitions give formal specification and explanation of the importance of such entities in the particulars of the problem at hand.

DEFINITION 4.1: STATION

A station *stn* is a location at which an aircraft *ac* can stop for any kind of operation such as maintenance, passenger loading, and passenger unloading.

It is clear that associated with every station there is a location, but for a location there may be zero or more stations associated with it. The relevant locations are the ones that have at least one station. For example Lowell (MA), Boston and New York are

location. Lowell does not have a station associated with it, Boston has one (Logan International Airport, BOS), New York has three (JFK, NWK, LGA). For simplicity the notion of location and station are interchangeable.

DEFINITION 4.2 - MARKET & LOCATION

A market m is characterized by a pair of locations a and b . This market will be identified as $m_{a,b}$. The opposite market for $m_{a,b}$ is the market $m_{b,a}$.

Let M be a set of markets. Inherently M defines also a set of locations $L(M)$ that is the union of all the pairs of locations of every market:

$$L(M) = \{ l \mid \exists m_{l,x} \vee m_{x,l} \in M \}$$

Generally given two locations a and b , there are both the direct and the opposite market ($m_{a,b}$ and $m_{b,a}$).

DEFINITION 4.3 - FLIGHT LEG

A flight leg fl is a non-stop flight departing at a certain time t_0 from a station a and arriving at station b at a time t_1 with $t_1 > t_0$. The departing station of fl is denoted as $\overline{fl} = a$. The arriving station is denoted as $\underline{fl} = b$.

Associated with the flight leg are the departure time and the arrival time.

Let $fl^- = t_0$ be the departure time for fl , and $fl^+ = t_1$ the arrival time for fl . Given two flight legs fl_1 and fl_2 there is an implicit order:

$$fl_1 \leq fl_2 \Leftrightarrow fl_1^- \leq fl_2^+$$

DEFINITION 4.4 - FLIGHT

A Flight, denoted as f , is an ordered set of consecutive flight legs

$$f = \{ fl_i \mid [fl_0] = a \wedge [fl_i] = [fl_{i+1}] \wedge [fl_{n-1}] = b \wedge fl_i < fl_{i+1} \quad \forall i = 1..n-2 \}$$

The $|f|$ is the number of flight legs inside f (in previous case $|f|=n$).

The redefinition of the extremes of a flight is $\overline{f} = fl_0$ and $\underline{f} = fl_{n-1}$.

$$S(f) = \{ stn \mid \exists [fl] = stn \vee \exists [fl] = stn \quad fl \in f \}$$

The station touched by that flight is defined as

The $|f|$ can be 1 or greater than 1. In the first case, the flight is the flight leg composing it. In the second case we have a sequence (or chain) of flight legs to follow.

DEFINITION 4.5 - FLIGHT AND MARKET

Given a market $m_{a,b}$ a flight serving it is a flight f which is $f \supseteq a$ and $f \supseteq b$.

Clearly, $|f|=n$ then there are $n(n+1)/2$ markets served (all the locations touched). For simplicity it is assumed there are no excluded markets.

DEFINITION 4.6 - SCHEDULE

A schedule s is a set of flights.

$M(s)$ is the set of markets served by schedule s .

$$M(s) = \bigcup_{f \in s} M(f)$$

$L(s)$ is a set of locations served by s

$$L(s) = \bigcup_{m \in M(s)} L(m)$$

Let S is a set of schedules, $M(S)$ is the set of all markets served by the schedule in S

$$L(S) = \bigcup_{s \in S} L(s)$$

$L(S)$ is the set of location covered by $M(S)$

$$L(S) = \bigcup_{m \in M(S)} L(m)$$

DEFINITION 4.7 – NETWORK TOPOLOGY: FULLY CONNECTED

A network composed of n markets is **fully connected** if given any market there is a direct flight from all other remaining markets.

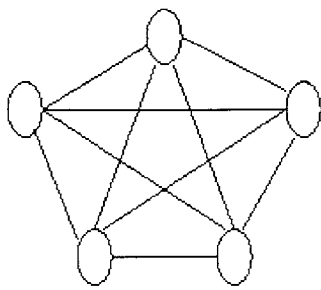
DEFINITION 4.8 – NETWORK TOPOLOGY: HUB AND SPOKE

A network composed of n markets is **hub and spoke** if one of the stations is used as intermediate stop for connecting every market.

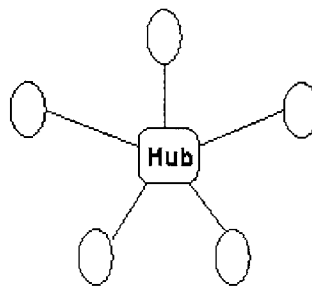
DEFINITION 4.9 – NETWORK TOPOLOGY: MIXED

A network composed of n markets is **mixed** if there are sub-network that are either hub and spoke or fully connected.

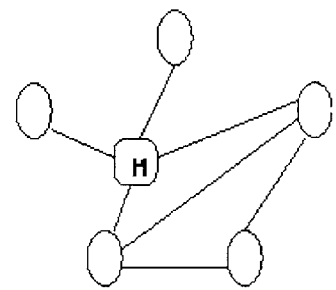
Notice that in the first case having n market, there are $n(n-1)/2$ flights. In the case of hub and spoke, there are $n-1$ flights only. Intrinsically the hub and spoke offers connectivity with less flights, however this efficiency has the disadvantage that every flight has an extra stop. Figure 4.1 gives a visualization of the topology.



Fully Connected Network



Hub and Spoke Network



Mixed Network

Figure 4.1 Network Topology

4.2 Robust Schedule

A schedule, or timetable, is a collection of information about times and places: the schedule reports the time at which a flight is going to depart from a location and arrive at a second location. Other information can be present on a schedule, like equipment type, elapsed time, flight number and so on, but the fundamental information is departure station, departure time, arrival station, arrival time.

There is an intimate relationship between the schedule and the airline network. The airline network can be represented as a graph in which the nodes are the stations and the arcs are the flight legs. In such a perspective, the schedule is a “picture” of the network at a given time. Extrapolating this concept, the combination of the static network and the schedule is representing the whole system as a network that spans in three dimensions, in which the third dimension is time. Notice that the schedule has a cycle embedded in it (repetition of service) and therefore the 3D network notation is a finite representation of the system if a single cycle is considered.

Robustness is a broad property of the schedule and it is defined along the time axis. In particular given an interconnected 3D network, an itinerary is a set of continuous arcs that connect an origin node to a destination node. The itinerary is equivalent to a flight (notice that it might be also a set of continuous flights). The schedule is said to be robust if cutting one arc of an itinerary there are other itineraries available that connect the two nodes. In common airline terminology, this means that there are other flights that connect the origin station to the destination station. The cut arc is the leg that is cancelled or delayed. The itinerary having the cut arc is the interrupted itinerary that is

translated into the disrupted flight. The second itinerary that allows the connectivity between origin and destination is called redundant itinerary. In this dissertation the redundant itinerary is called **flight alternative option**.

Robustness has two characteristics: Temporal and Spatial.

The **temporal robustness** characteristic refers to the distance along the time axis of the interrupted itinerary with the redundant one. The airlines refer to it as the time window in which the service can continue or be reestablished from the origin to the final destination. This is usually the elapsed time that a passenger has to wait to access another flight that will conduct him to the destination.

The **spatial robustness** refers to the redundant itineraries. There is spatial robustness if the original itinerary and the redundant itinerary do not share any arc. This indicates that an option flight has spatial robustness if it is going through different station than its alternative options.

Based on such distinctions, here are some handy definitions.

DEFINITION 4.10 - MARKET MULTI-COVER IN $[t_0, t_1]$

A market m is said to be multi-covered of degree k in $[t_0, t_1]=\Delta t$, if there is a set of at least k flights that serve this market in the period Δt . The market m is said to be k -covered in $[t_0, t_1]$.

DEFINITION 4.11 - MARKET MULTI-COVER IN t

A market m k -covered in $[t_0, t_1]=\Delta t$ is said to be k -robust in t , if it is k -covered independently from the origin of the interval t_0 .

DEFINITION 4.12 - COVERAGE FUNCTION FOR A MARKET

The coverage function (or coverage) for a market is represent by a chronological

$$C_t(m) = \{f \mid f_0^+ - f_{k-1}^+ < t\}$$

ordered set of flights serving such a market

The number of flights in such set defines the degree of coverage. If $|C_t(m)|=k$ then the market m is said to be k covered in t .

DEFINITION 4.13 - DISJOINT FLIGHTS

Given a set of flights F , the disjoint of this set is a second set containing flights that are not touching common stations.

$$D(F) = \bigcup_{f \in F} S(f) \setminus \bigcap_{f \in F} S(f)$$

DEFINITION 4.14 - ROBUSTNESS FUNCTION FOR MARKET

The robustness function (or *robustness*) for a market m is represented by a

$$R_t(m) = D(C_t(m))$$

chronological ordered set of flights serving such markets that are disjoint.

As in the previous case, the number of flights in the market robustness set defines the degree of robustness.

The multi-coverage property is not enough to characterize the robustness of a market. For example the following three flights serving market $m_{a,b}$, as reported in Figure 4.2

$$f_1 = \{fl_{11}, fl_{12} \mid \lceil fl_{11} \rceil = \lfloor fl_{12} \rfloor = c\}$$

$$f_2 = \{fl_{21}, fl_{22} \mid \lceil fl_{21} \rceil = \lfloor fl_{22} \rfloor = c\}$$

$$f_3 = \{fl_{31}, fl_{32} \mid \lceil fl_{31} \rceil = \lfloor fl_{32} \rfloor = d\}$$

$$f_1^-, f_2^-, f_3^- \in [t_0, t_1] = \Delta t = t$$

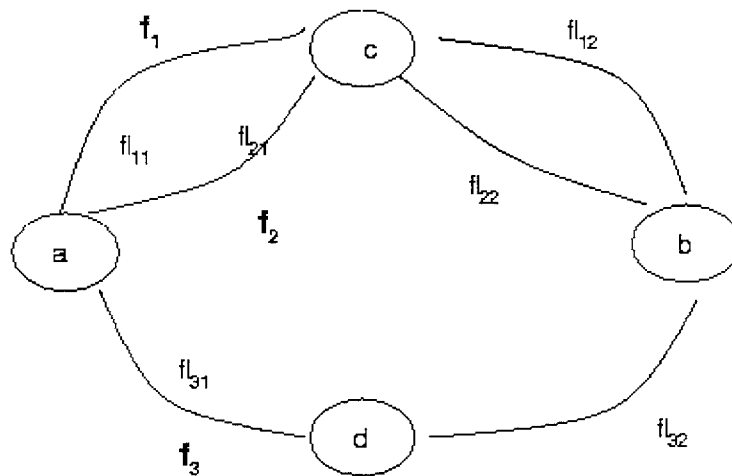


Figure 4.2 Example of Coverage and Robustness

Then market $m_{a,b}$ is 3-covered. In this case it is possible to have repetition of flights having the same flight legs but scheduled at different times

If flight f_1 , for example, is late or cancelled, then passengers can be relocated in f_2 or f_3 . If instead there is a problem at station C then not only is f_1 affected, but f_2 as well. In this case the only available option is f_3 . In such a situation it is said that the market $m_{a,b}$ is 3-covered, but 2-robust.

The characteristic of robustness is related to a market and it is bi-dimensional in the degree of coverage (k) and time of coverage (t). Given two markets m_1 and m_2 it is possible to decide which one is more robust in two possible ways:

1. Robustness in time t :

$$R_t(m_1) < R_u(m_2) \Leftrightarrow |R_t(m_1)| < |R_u(m_2)| \wedge t=u$$

This means that the degree of robustness of the first is less than that of the second when considering the same time interval. In this case, it is said that m_1 is less robust than m_2 in time $t=\Delta t$.

2. Robustness in coverage k :

$$R_t(m_1) < R_u(m_2) \Leftrightarrow |R_t(m_1)| = |R_u(m_2)| \wedge t < u$$

This means that the time interval of robustness of the first is less than that of the second one when considering the same coverage degree. In this case, it is said that m_1 is less robust than m_2 in degree k .

The characterization of robustness in this way can fulfill two purposes: optimize coverage for time, or optimize coverage for degree. The first case can be useful when degradation of operations happens fast in time (ATC problems). Optimizing imposing a small interval of coverage allows an airline to cut service or to absorb reduction of service very quickly.

The second case, instead, is important when a change of condition or reduction of service happen slowly (weather conditions). In this situation it is reasonable to have a high number of rearrangements available for serving the market in consideration.

The definitions above are very useful for characterizing a market on a typical service, but unfortunately there are 2 considerations that make this definition impractical:

1. The coverage set and the robustness set for a market are not unique. Depending on the shifting on the time window during the day, the flights in these sets are changing even if their total number (k) is not.
2. Flights are not uniformly distributed across the day. It is clear, then, that it can be more useful to define coverage of a market given a flight to recover.

This flight will be the one that is going to be examined during any disruptions and, therefore, the operator can use the robustness of the schedule itself.

Thus, the normal implementations of these definitions must be done by considering a reference flight.

DEFINITION 4.15 - COVERAGE FUNCTION FOR A MARKET GIVEN A FLIGHT

The **coverage function** (or **coverage**) for a market given a flight \hat{f} is represented by a chronologically ordered set of flights serving this market that arrive at the destination up to t minutes after the arrival time of \hat{f}

$$C_t(m, \hat{f}) = \{f \mid \hat{f}^+ \leq f_i^+ \leq \hat{f}^+ + t\}$$

DEFINITION 4.16 - INVERSE COVERAGE FUNCTION FOR A MARKET GIVEN A FLIGHT

The **inverse coverage function** (or **inverse coverage**) for a market $m_{a,b}$ given a flight \hat{f} , is represented by a chronologically ordered set of flights serving this market that depart from origin up to t minutes after the departure time of \hat{f}

$$C_t^{-1}(m, \hat{f}) = \{f \mid \hat{f}^- \leq f_i^- \leq \hat{f}^- + t\}$$

It is also possible to extend the definition of coverage and inverse coverage function, considering the flight in consideration as centered in the reference interval. In this case the flights considered will be the ones departing or arriving before and after the flight to protect.

DEFINITION 4.17 - EXTENDED COVERAGE FUNCTION FOR A MARKET GIVEN A FLIGHT

The **extended coverage function** for a market $m_{a,b}$ given a flight \hat{f} , is represented by a chronologically ordered set of flights serving this market that arrive t minutes before or after \hat{f}

$$CE_t(m, \hat{f}) = \{f \mid \hat{f}^+ - t \leq f_i^+ \leq \hat{f}^+ + t\}$$

DEFINITION 4.18 - EXTENDED INVERSE COVERAGE FUNCTION FOR A MARKET GIVEN A FLIGHT

The **extended inverse coverage function** for a market $m_{a,b}$ given a flight \hat{f} , is represented by a chronologically ordered set of flights serving this market that depart t minutes before or after \hat{f}

$$CE_t^{-1}(m, \hat{f}) = \{f \mid \hat{f}^- - t \leq f_i^- \leq \hat{f}^- + t\}$$

DEFINITION 4.19 - ROBUSTNESS FUNCTION FOR MARKET GIVEN A FLIGHT

The **robustness function** (or **robustness**) for a market m given a flight \hat{f} is represented by a chronologically ordered set of flights serving this market that they are disjoint.

$$R_t(m, \hat{f}) = D(C_t(m, \hat{f}))$$

Notice that:

1. The robustness degree for a flight is less than the robustness of the market

$$|R_t(m, \hat{f})| \leq |R_t(m)| - 1$$

2. The set of composing flights of the robustness set for a flight \hat{f} is equal to the robustness of a market minus the flight \hat{f} .

$$R_t(m, \hat{f}) = R_t(m) / \hat{f}$$

In this way there is a third dimension for coverage and robustness of a flight. The three dimensions are:

1. The time interval for which the coverage is requested: the shorter, the better.
2. The number of flights available in such coverage: the higher, the better.
3. The reference flight: this adds only some specification when we are deciding on the service to generate.

The flight reference adds a different cut on the quality of a schedule. Due to the fact that the flights are not equally distributed during the day, the specification of the flight reference helps to isolate the time window of interest.

For example, suppose there are the flights reported in Table 4.1.

Table 4.1 Coverage Example

Flight	Origin	Time	Destination	Time
F1	BOS	8:00 am	ORD	10:00 am
F2	BOS	9:00 am	ORD	11:00 am
F3	BOS	4:00 pm	ORD	6:00 pm
F4	BOS	5:00 pm	ORD	7:00 pm

The generic coverage on the day is 8 hours with 2 flights. A shorter interval is going to have no coverage: for example, considering the interval 9:30am to 3:30pm there is an interval of 6 hours but no flights. Considering one flight at a time, there is a different picture as shown in Table 4.2.

Table 4.2 Coverage Reference

F1	1 in 1 Hour
	2 in 8 Hours
	3 in 9 Hours
F2	1 in 7 Hours
	2 in 8 Hours
F3	1 in 1 Hour
F4	Uncovered

As seen, the introduction of the reference flight allows us to isolate the flights that have no cover (f_4) and also the minimum expected time for covering a flight (the first flights).

4.3 Coverage Example

Below there are examples of the sets outlined above. The market considered is the New York - Las Vegas market. The airports involved are Newark, NJ Newark International Airport (EWR) and Las Vegas, NV Mc Carran International Airport (LAS). The airline considered is United Airlines. Neither EWR nor LAS are hubs for United. During a typical day in July 2000 there are 299 connecting flights. The parameter used for determining the number of flights is:

1. Maximum number of stops: 2.
2. Minimum connecting time: 30 minutes (1/2 of a hour).

3. Maximum connecting time: 180 minutes (3 hours).

Consider the flight in Table 4.3 that will be the reference flight f .

Table 4.3 Reference Flight

EWR-LAS	08:15	12:15				
1 st flight	0347	N594UA	EWR	DEN	08:15	10:34
2 nd flight	2709	N371UA	DEN	LAS	11:25	12:15
Stops	2					
Trip Time	420 Minutes					
Fly Time	369 Minutes					
Rest Time	51 Minutes					

Considering also a time interval of 2 hours, the following sets of flights are outlined:

1. $C_t(m, f) = C_{120}(EWR-LAS, f)$

Considering the flight arriving at LAS in 2 hours from the scheduled arrival time of the original flight f , reduces the number of flights to 4. Excluding the one that has the same first leg reduces the set to 2 flights. They are reported in Table 4.4.

Table 4.4 The $C_{120}(EWR-LAS, f)$ Set

EWR.LAS	08:15	13:30				
1 st Flight	639	N7285U	EWR	ORD	08:15	09:41
2 nd Flight	721	N1835U	ORD	LAS	11:50	13:30
		Stops = 2	Trip Time = 495	Fly Time = 366	Rest Time = 129	

EWR.LAS	09:15	13:30				
1 st Flight	641	N7274U	EWR	ORD	09:15	10:31
2 nd Flight	721	N1835U	ORD	LAS	11:50	13:30
		Stops = 2	Trip Time = 435	Fly Time = 356	Rest Time = 79	

2. $CE_t(m, f) = CE_{120}(EWR-LAS, f)$

This set is identical to the previous because no flight arrives 2 hours before f .

3. $C^{-1}_t(m, f) = C^{-1}_{120}(EWR-LAS, f)$

There are 79 flights all departing between 8:15 and 10:15. Removing the one having the same first leg reduces the number to 72.

4. $CE^{-1}_t(m, f) = CE^{-1}_{120}(EWR-LAS, f)$

This set extends the previous set of 72 flights with another 59 flights. All these flights do not have the same first leg and depart between 6:15 to 10:15

The first set, $C_{120}(EWR-LAS, f)$, is the most important because passengers are focused on arriving at their destination as close as possible to the original arrival time.

4.4 Market Protection and Traffic Flow

Given the above sets, it is possible to determine if a flight is protected enough to allow perturbations of its service. The protection of a flight is related to the options, or alternatives, that are available in place of that flight.

From here on, the most interesting set is the coverage set: $C_i(m,f)$, the robust set can be derived from the coverage set via the application of the disjoint operator.

Every flight indicated in the above sets is an opportunity for a passenger to relocate if the flight is disrupted or an opportunity for the scheduler if a preemptive delay/cancellation has to be assessed.

Given a multi-covered market, then in any of the stations that the flight is touching there is a need to determine which path is the best to follow to reach the destination. The process of making an effective decision is the other side of the problem: when provided enough options or coverage, what is the best way of taking advantage of it?

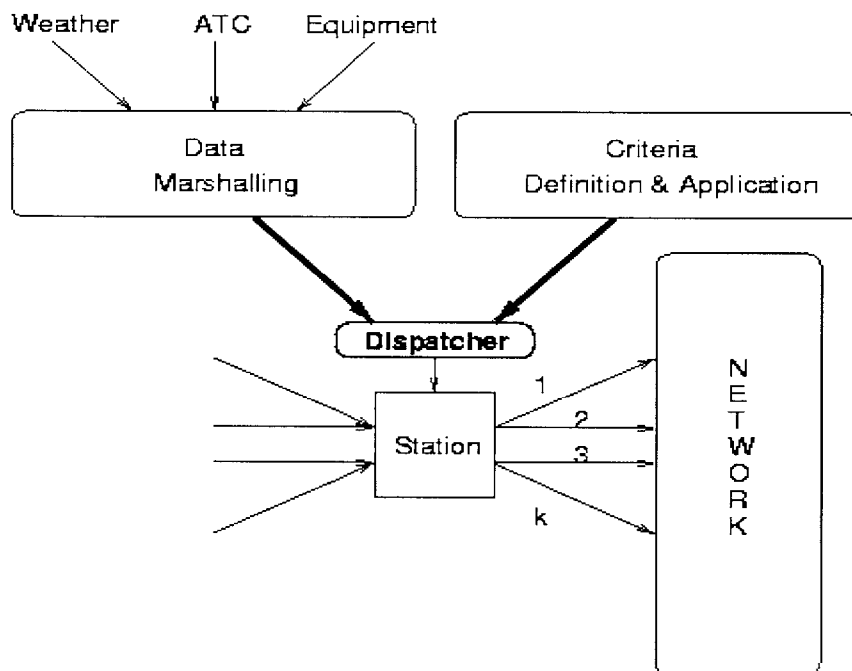


Figure 4.2 Flow Controller Decision Framework

Figure 4.1 shows a first approach. A *flow controller*, or more generally, the **dispatcher** makes the decision, based on information available: the situation at the station, throughout the network and at the final destination.

The information comes from air traffic control, weather forecast, and equipment maintenance and availability department. Based on these data and some logical criteria,

the controller determines in real time which path is appropriate for the passenger/aircraft. This path is the one that has been generated during the planning process.

There is, therefore, another side to the Robustness Scheduling problem that is the **Decision Making Problem**. In this case, information must be collected, weighted, marshalled, summarized and expressed in a clear way. How to formulate and quantify the criteria that allows the controller to find out the proper decision is another critical issue. The decision making process can be better defined, depending the level of details it operates, and the goals that the dispatcher is trying to achieve.

4.5 Decision Making Process

Suppose that there is one passenger p in $m_{a,b}$. He planned to take the flight f that originated from a and has destination b departing at a certain time t_0 . The options available to p for reaching the destination in the future time window t (t_0+t), are k , the $C_i(m,f)$.

The first flight f in the coverage set is the flight that p is supposed to take (f). Reducing the flow to b or to any of $S(f)$ is imposing different decision opportunities on p :

1. If f is not affected, p has 100% willingness to take such a flight.
2. If f is affected then there is a reduction of opportunity of x -percent for such flight. This implies that p has $100-x$ -percent willingness to take that flight. The remaining x -percent is allocated to the next flight in the coverage set if:
 - a. Capacity allows this allocation
 - b. Restriction allows it
3. If there are no other flights available on the coverage set, then p has x -percent probability of not arriving in b with a substitution in the time t . In this case variation of t can increase p 's chances.

The order in which the substitute flights are set is based on the policy for recovering the operation. Examples of that can be

1. lower number of connections
2. closest arrival time to the original one
3. first available
4. higher reliability
5. lower Disruption Value

The same reasoning can be applied when the reduction of flow is done to any flight going to any of the stations touched by the flight f . p 's willingness to take a flight that goes to an interrupted station is 0, and therefore all the flights going to such a station are removed from the coverage set. In this case the coverage for a market is not enough. In this case $m_{a,b}$ is robust only if there is a reallocation possibility.

The controller in this case is at passenger level. The considerations are based on:

1. capacity constraints.
2. cabin and class constraints.
3. willingness of the passenger to move.
4. level of service offered.
5. higher reliability.
6. lower passenger disutility.

The advantage of such a method is that there is a very fine control over the passengers' flow. It is possible to reroute every single passenger, and via proper decision on which passengers to move, give different levels of service.

The disadvantages of such a method are:

1. The equipment is not affected. If a flight is delayed, this approach does not reduce its delays. The number of passengers affected is smaller, but nevertheless the equipment, and so the service, is still disrupted.
2. If a station "goes down" then this application of robustness does not help to recover the operation. Nevertheless, the passenger can still use the alternative paths, but now the capacity constraints are more *stringent*.
3. The decisions are made *run time* on every station for every passenger, and that can be computationally intensive. It could be appropriate to have human intervention in the process.
4. The single decision of which passenger to move can be troublesome. The result of such a process can be that 5% of the passengers of such a flight have to be moved to a second flight. Well, which of the current passenger are going to be moved? Families, groups, class distinction, time opportunity, intrinsic cost ("*I will arrive late to my meeting!*") and value ("*I will not be stuck somewhere!*") of moving should be considered.

This approach can be useful in association with the level of service offered by the airline. It does not affect operation at all. The variation of the system condition is only passively

observed, and no effort to adapt the network to the disruption is assumed. This is very simple and may be put in place easily.

4.6 Conclusion

The basic terms and definitions and the introduction of the coverage and robust set given a market and a flight have been proposed. An example of identification and utilization of the coverage set has been presented.

After introducing the goodness of the coverage set, an indication of how to take advantage of such an opportunity has been highlighted. The process that the dispatcher should follow is outlined as well.

The next chapter will highlight how to increase the options in the coverage set, and how to grade them to aid the dispatcher in making the proper choice during the decision-making process.

Chapter 5

Metrics and Reliability Measures

Having identified the core part of the Schedule Design Model, presented the key element of the Robust Schedule Model, and characterized the element that needs to be manipulated (the flights), a way of grading and comparing various flight options is needed.

The introduction of the reliability measure is a way of doing this. The **Option Value (OV)** is the means of grading and selecting the various options. The **Option Disruption Value (ODV)** is a way of comparing the options. Both the option value and the option disruption value are formally introduced here. These metrics form a set of metrics that represent the reliability of a flight. They are compared with what has been studied and developed for other means of transportation (trains, and buses).

The Option Value is composed of two key elements: the **Option Value for a Flight** and the **Option Value for a Connection**. A formal investigation of the weighting values of the composing terms of the formula for the OVs is presented, with some insight on the meaning of each terms.

In the end of the chapter there is a summary of alternative and meaningful ways of measuring the reliability of the schedule. This last step is needed to understand the quality of any algorithm focused on improving the schedule reliability.

5.1 Reliability Measures

As described in Carey (1999), reliability is a measure of the quality of service provided by an operator, which needs to be constantly monitored.

In particular an accurate reliability measurement is valuable because:

1. **Operators** use it in planning, management, control, dispatching and marketing of the service.
2. **Users** use it for making travel choices.
3. **Regulators** need to check if operators are delivering the promised or contracted quality of service. Typically operators are required by law to publish their punctuality performance.

There are various methods for quantifying reliability like, for example:

1. Probability of early departure.
2. Mean difference between expected arrival time and scheduled arrival time.
3. Mean delay of an arrival given that one arrives late.
4. Mean delay on arrival given that one arrives more than x minutes late.
5. Standard Deviation of the arrival time.
6. Adjusted Standard deviation.
7. Square difference between scheduled and actual arrival time.
8. Percentage of "on time" arrivals;

More generally, there are analytic methods that are easy to implement and very good for simple systems; there are also simulation methods: they are more accurate but time-consuming. In recent years the utilization of *ad hoc measures* has emerged, which are a combination of the above plus a combination of so-called "*rules of thumb*".

Another categorization of these tools is based on the period in which the estimates are done; this can be *a posteriori estimates* or *a priori estimates*. In general, the measure of observed delay is widely used as a posteriori estimate of delay, but is of little help to forecasts. The a priori, instead, must be based on a probabilistic approach. This approach is used for forecasting and can be also used on the design stage.

Another important characteristic is that the reliability metric should be uniform across markets, trip time and location.

A good reliability indicator should have the following characteristic:

1. **Simplicity.** It should be easy to calculate with the data available. Only meaningful data that is easy to observe and collect should be used. The concept on which it is based should also be intuitive and easy to quantify.
2. **Statistical Aspects.** The metric should use some statistical consideration that should be formally tested. These postulations also need to be validated. At the

same time the assumptions should not be too generic or too specific. This is to avoid the possibility that every event fits, or very few events fit the hypothesis.

3. **Rules of Thumb.** There should be some logic between the terms involved in the reliability equation. On the other hand, there should be integration with terms so that even if they express different aspects of the event under examination, they can give at the same time a good understanding of the phenomenon in consideration.
4. **Market/trip Independence.** The reliability measure should be easy to implement independently of the data available. If the data refers to the same phenomenon, then very similar reliability indication should emerge.
5. **Uniformity.** The indicator should allow direct or indirect comparisons of events and schedules. It should be easy to determine if a situation is more reliable than another without any contradiction or uncertainty.
6. **Unit Independence.** The reliability should be expressed in terms that can be easily translated into monetary value, time indication, or any other meaningful unit needed for proper analysis.

Having listed the characteristics of any good reliability indicator, below is the formal definitions of the reliability indicator for flight legs, connections, flights and schedules.

5.2 Option Value for a Leg

As said, the reliability measure is the **Option Value (OV)**. A formal definition of it is:

$$OV_{leg}(fl) = (w_{gt}\bar{t}_{gt} + w_{to}\bar{t}_{to} + w_{tr}\bar{t}_{tr} + w_{ti}\bar{t}_{ti} + w_s |t_s - t'_s|) e^{\frac{p_c}{1-p_c}} e^{1-l^3} \quad (5.1)$$

where

- $w_{gt}\bar{t}_{gt}$ is the average gate time and the relative weight.
- $w_{to}\bar{t}_{to}$ is the average taxi out time and the relative weight.
- $w_{tr}\bar{t}_{tr}$ is the average flying time and the relative weight.
- $w_{ti}\bar{t}_{ti}$ is the average taxi in time and the relative weight.
- $w_s |t_s - t'_s|$ is the slack adjustment introduced in the schedule by the scheduler and its weight. This value should be as close as possible to the expected one. In detail

- t_s is the **Flight Leg Slack Time** and it comes as a difference between the scheduled flight time and the summation of the gate time, taxi out, fly, and taxi in time:

$$t_s = t_{gg} - (\bar{t}_{gt} + \bar{t}_{to} + \bar{t}_{tr} + \bar{t}_{ti}) \quad (5.2)$$

- t_s' is the **Expected Leg Slack Time** determined or imposed for the slack time.
- p_c is the probability of being cancelled. This probability is calculate as follows:

$$p_c = p_{cd} + p_{ca} (1 - p_{cd}) \quad (5.3)$$

where p_{cd} is the probability of being cancelled on departure, and p_{ca} is the probability of being cancelled on arrival. The above formula can be spelled out as: *the probability for a flight leg to be cancelled is equal to the probability of a flight leg to be cancelled in departure plus the probability of being cancelled in arrival given it departed.*

- l is the leg load factor.

The Option Value indicates the quality of an option. The lower the value better the option.

The slack times have a special role in the formula: the t_s inserted in the formula is used for compensating unreliability of the schedule; when t_s is greater than, or smaller than t_s' then the flight is under- or over-compensated for reliability.

There is minimum value for the OV that is called **Ideal Option Value (I-OV)**. The Ideal Option Value is obtained in the case of perfect compensation of the schedule ($t_s' = t_s$), lack of cancellations ($p_c = 0$) and availability of seats for passengers' relocation ($l = 0$). This is calculated as

$$OV_{leg}^{ideal}(fl) = w_{gt} \bar{t}_{gt} + w_{to} \bar{t}_{to} + w_{tr} \bar{t}_{tr} + w_{ti} \bar{t}_{ti} \quad (5.4)$$

Given a flight leg fl , the I-OV is the lowest and best value that can be achieved for fl . Using the Ideal Option Value as a base line, the **OV Ratio (OV-R)** may be determined as an indication of the quality of the flight:

$$R_{leg}(fl) = \frac{OV_{leg}(fl) - OV_{leg}^{ideal}(fl)}{OV_{leg}^{ideal}(fl)} = \frac{OV_{leg}(fl)}{OV_{leg}^{ideal}(fl)} - 1 \quad (5.5)$$

An OV Ratio that is a positive number can be seen as the percentage of increased burden of a flight in **utils (utility units)** on a generic passenger.

In certain situations, too many compromises may be required to achieve the I-OV; thus, it may be appropriate to interpret this value situation by situation (competitiveness reasons, etc) and act accordingly.

As Carey (1994) states, *a common goal in service planning is often “of satisfying” than “of optimizing”: that is, satisfy certain pre-specified measures of reliability or performance.*

5.3 Slack Time and Behavioral Response

When a schedule is implemented the time that actually is taken for any activity is subjected to variation; hence one strategy for improving reliability is to allow more time in the timetable, by adding some slack time.

As reported in Carey (1998) to reduce such over-runs or lateness, and improve reliability and costs, some extra time is usually allowed for some or all activities in the schedule. However, he points out that it is a known fact that if more time is allocated for an activity, then the activity often tends to take longer. Because of this behavioral response, some or all of the benefits (in reliability, costs, etc.) of the extra time allowance are lost. This is called **behavioral response** because it is intrinsic in operators, dispatcher and system having more time availability. Some managers, therefore, consider responding to this phenomenon with a “tight” schedule to avoid what is perceived as an unnecessary waste of time.

The cost of time plays a major role here. This cost is given as a summation of cost of the trip plus the cost of lateness (if any) minus the cost of earliness (if any). Even if the perception of these times is different, and even if sometimes earliness may be seen as a benefit rather than a penalty, studies on the British Rail (1980) show that early arrival is without doubt a disutility for the traveling passengers. In general, however, the cost of earliness is much less than the cost of lateness, because a late arrival can generate a cascade of delays (**knock-on delays**).

Given the probability density distribution of arrival time $f(t)$ (let $F(t)$ be the arrival time cumulative distribution), the cost of unit of trip time c_t , and the cost of unit of lateness c_l , the cost of the lateness in respect to the arrival time T is calculated as

$$C_l(T) = E[(t - T)c_l | t \geq T] = \int_T^{\infty} (t - T)c_l f(t) dt \quad (5.6)$$

This is a convex function; hence it has a global minimum. When $c_l \geq c_t$ (cost of lateness is greater than trip cost) then the optimal probability of lateness is obtained as

$$[1 - F(T^o)] = \frac{c_t}{c_l} \quad (5.7)$$

Hence the optimal probability of arriving on time or early is

$$F(T^o) = 1 - \frac{c_t}{c_l} \quad (5.8)$$

And the optimal schedule time is

$$T^o = F^{-1}(1 - c_t/c_l) \quad (5.9)$$

where T^o denotes the optimal value of T .

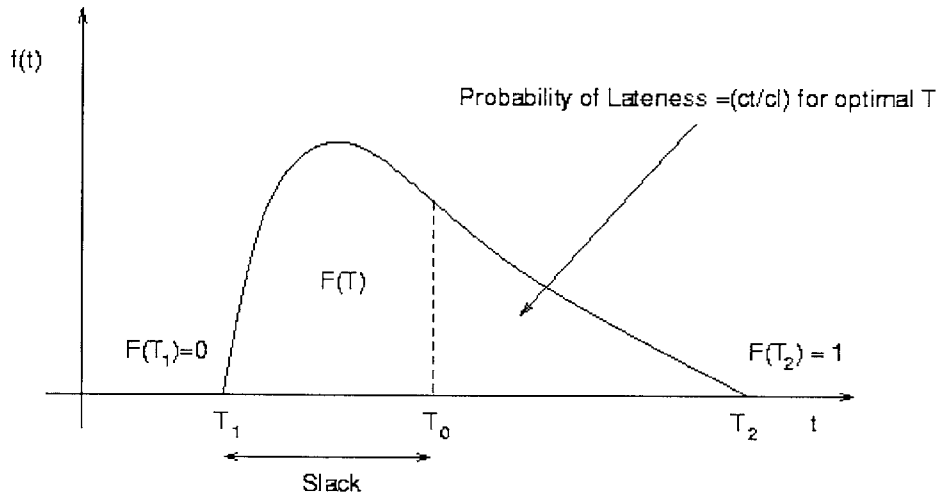


Figure 5.1 Arrival Time Profile. Lateness Analysis

Given the arrival distribution, the relationship between the value of c_l and c_t determines the proportion of acceptable late arrivals. For example:

1. If $c_l = 2 c_t$ then $F(T^o) = 0.5$. The optimal scheduled time T^o is the median of the arrival distribution, which means that 50% of arrivals reach destination later than scheduled.
2. If there is no cost associated with lateness, $c_l \rightarrow c_t$ then $F(T^o) = 0$ and $T^o = \min(t) = T_1$. Given that there is no extra cost when arriving late, the optimum is obtained via setting the scheduled time to the earliest possible arrival time in a way that all the arrivals are exactly on time or later, and none is early.
3. If there is a high penalty for lateness, it means that $c_t \rightarrow 0$ or $c_l \rightarrow \infty$ then $F(T^o) \rightarrow 1$ and $T^o \rightarrow \max(t)$. Given that the lateness burden is very high, the optimum is obtained via setting the schedule time at the latest possible arrival time so all the arrivals are exactly on time or early, and none of them is late.

4. If $c_t = 2.5 c_t$ then $F(T^\circ)=0.60$ or 60%. This means that we are willing to accept only 40% of the trips arriving late.
5. If $c_t = 5.6 c_t$ then $F(T^\circ)=0.82$ or 82% (This is the $\mu+\sigma$ percentile of a normal distribution). This means that we are willing to accept only 18% of the trips arriving late.

The calculation of the lateness is related to the distribution of the arrival time. As described before, and as outlined by Carey (1999), the arrival time is obtained as:

$$T_a = T_d + T_t = T_d + (\bar{t}_{gt} + \bar{t}_{to} + \bar{t}_{tr} + \bar{t}_{ti}) \quad (5.10)$$

Given its distribution, the selection of the penalty determines the percentage of flight allowed to arrive later than the forecasted time. On the other hand, the selection of the percentage of the flights acknowledged to arrive late determines the penalty for every "late minute."

The determination of the arrival time distribution can be done in two ways:

1. **Profile Convolution.** This is a formal way of combining independent distributions and it is given by

$$f_t(t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} t_{gt}(v) t_{to}(x-v) dv \quad t_{tr}(y-x) dx \quad t_{ti}(t-y) dy \quad (5.11)$$

2. **Normality Assumption.** The dominant profile in the flight trip is the flight time, which 80% of the time is a normal distribution. It is arguable then that the summations of the four profiles are going to be very close to a Normal distribution as well. On the other hand, even if the sample is quite small, the *Central Limit Theorem* reinforces this assumption. Under this assumption, the summation of the profiles, then will have a Normal distribution with parameters μ and σ defined as:

$$\mu = E[t_{gt}] + E[t_{to}] + E[t_{tr}] + E[t_{ti}] \quad (5.12)$$

$$\sigma = \sqrt{Var[t_{gt}] + Var[t_{to}] + Var[t_{tr}] + Var[t_{ti}]} \quad (5.13)$$

In the profile example showed in Chapter 3, the above concepts are applied. The graphs in Figure 5.2 and 5.3 show the resulting distributions.

The two distributions have the same mean by construction, but the variance differs; the convolution function is skewed to the left with higher kurtosis than the one obtained with the normality assumption. The important value of the 82nd percentile, that is at $\mu+\sigma$, is off by only 4 minutes.

This is important, in light of the weight given to the lateness time, and the determination of the acceptable level of late flights.

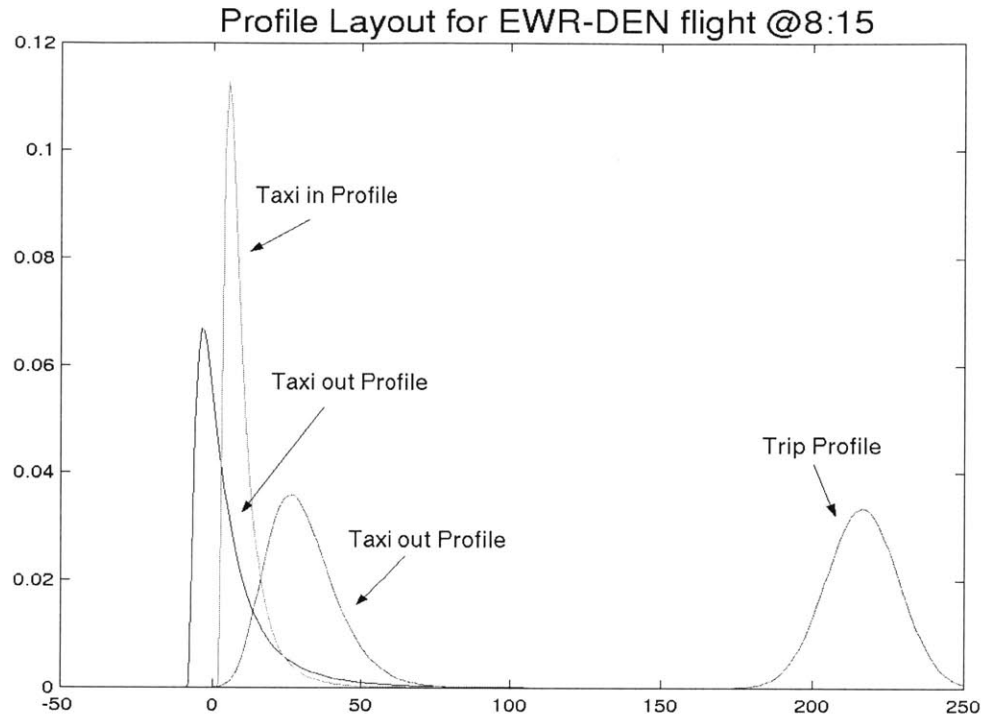


Figure 5.2 Profile Layout EWR-DEN Flight Leg Departing at 8:15

Trip Profile, Profile Convolution and Normality Assumption Comparison

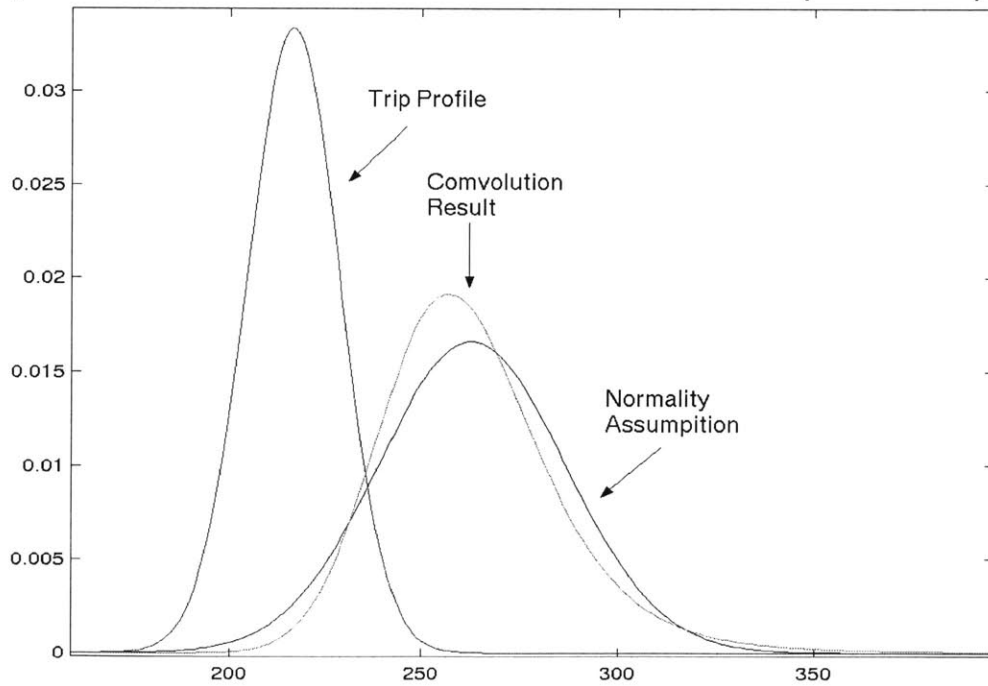


Figure 5.3 Comparison Convolution and Normality Assumption for Arrival Time Profile

5.4 Option Value for a Connection

The approach used for the flight leg can be applied to the station during the connecting time. Given two flight legs, then, it is possible to determine the disruption generated at the connecting station as a function of the arrival and departure profiles for such a station.

The **Option Value for a Connection** given two flights, is defined as

$$OV_{sm}(fl_1, fl_2) = (w_c(t_c + t_v) + w_w t_w) e^{\frac{1-p}{p}} \quad (5.14)$$

where

- t_c is the **minimum connecting time** at the station. This is fixed and determined by the airport authority.
- t_v is the **connecting slack time** added to the minimum connecting time generated by the variability of the arriving profile for flight fl_1 and the departing flight fl_2 .
- $w_c(t_c + t_v)$ is the **effective minimum connecting time** needed for making the connection between fl_1 and fl_2 , with its weight.
- t_w is the extra time spent waiting (*wasted time*) for the departure of the second flight (fl_2).
- p is the probability of making the connection. This is bounded to t_v in terms of how much time is appropriate to insert in the connecting time to allow $p\%$ of the connections to take place.

As said before, the lower the value of $OV(fl_1, fl_2)$, the better the connection between fl_1 and fl_2 .

It is handy to define the **time spent at the station** t_s as the summation of the various times:

$$t_s = t_c + t_v + t_w \quad (5.15)$$

Due to the fact that t_s as well as t_c is fixed during the schedule generation process, then t_v is fixed as the time that gives a percentage of the flights making the connection. It turns out that the value of t_w can be positive or negative. From here on, it is not restrictive to consider $t_w \geq 0$ only. In the case of $t_w < 0$ then t_v is decreased (hence p is decreased as well) and t_w is increased until it reaches 0. Notice also that by definition t_s is greater than or equal to t_c .

There are some interesting relationships between the terms:

1. $t_v \rightarrow \infty \Rightarrow p \rightarrow 1$ This means that for a sufficiently large value of the connecting slack time inserted in the connecting time, the probability of making the connection becomes a certainty.
2. $t_v = 0 \Rightarrow p \rightarrow 0$ This implies that there is an intrinsic minimal chance of losing the connection if relying only on the minimum connecting time t_c .

The calculation of t_v is obtained using the probability distribution of the connecting time.

As described in Carey (1994), this function f_s is obtained as a convolution of the distribution of the trip time of the arriving flight at the station (f_i) and the profile of the gate time of the departure flight (t_{gt}):

$$f_s(t) = \int_{-\infty}^{\infty} t_{gt}(v) f_i(v-t+T+t_s) dv \quad (5.16)$$

where T is the time fixed for such a trip (it determines the probability of arriving at time T).

It is easy then to calculate the probability of making a connection given a certain connecting time t_{min} :

$$F_s(t_{min}) = f_s(t < t_{min}) = \int_{-\infty}^{t_{min}} f_s(v) dv \quad (5.17)$$

As seen before, there is an ideal OV for the connection that can be achieved when there is no wasted time waiting for the connecting flight ($t_w=0$) and there is a certainty of making the connection ($p=0$). This can be achieved when

1. $t_s = t_c + t_v$ This means the waiting time is totally absorbed by the variability of arriving and departing flights. Hence there has been perfect timing between the connecting flights.
2. $t_s = t_c$ and $t_v = 0$. This means that there is no variability in the arriving and departing time of the flights. Hence the connecting time is the only time truly needed to continue the trip.

The definition of the **Ideal Option Value for the Connection** related to the flight considered is obtained as:

$$OV_{sm}^{ideal}(fl_1, fl_2) = w_c(t_c + t_v) \quad (5.18)$$

The I-OV is the best value that can be achieved for the connection between fl_1 and fl_2 .

The OV Ratio for the connection at the station is, like R_{fl} , an indication of the quality of the flight connection (**flight coordination**):

$$R_{con}(fl_1, fl_2) = \frac{OV_{stm}(fl_1, fl_2) - OV_{stm}^{ideal}(fl_1, fl_2)}{OV_{stm}^{ideal}(fl_1, fl_2)} = \frac{OV_{stm}(fl_1, fl_2)}{OV_{stm}^{ideal}(fl_1, fl_2)} - 1 \quad (5.19)$$

The ratio R_{con} can be seen as the percentage of increased burden of a flight in util on a generic passenger while making the connection. This ratio, like the R_{fl} , is a positive number.

Notice that, even if the I-OV is the one that can be best achieved, this should be evaluated in light of the connections with other flights and other competitive consideration.

It is interesting to show that the ratio R_{con} is a way of measuring the coordination of the flight fl_1 and fl_2 as defined in Rietveld (2001).

5.5 Flight Coordination and Rescheduling Costs

The idea used as the basis for the best connecting time between flights is closely related to the coordination of the arrival and departure flight legs composing the flights: a better coordination of arrival and departure schedule improves reliability.

As described in Rietveld (2001), the coordination of flight is based on the frequency of the options. However, frequency alone does not explain total travel costs. Even with high frequencies, bad co-ordination of flight schedules may leave the transfer passenger with long waiting times at airports, or high rescheduling costs at the origin or destination.

Given the frequencies on the two legs of the connections, it may be demonstrated that rescheduling time typically depends on the frequency of the low frequency leg, whereas the high frequency leg ultimately determines the waiting time.

Higher frequencies obviously help to reduce waiting time at airport. However, waiting times do not only depend on frequency per se but also on the way flight departure and arrival times are coordinated, and on the minimum connecting time between them.

When considering a connection with two legs, the *transfer* time depends on the frequency of the *most frequent leg*, whereas the *rescheduling* time depends on the frequency of the *least frequent leg*.

Rescheduling means that passengers cannot realize their desired time of departure and/or arrival because of the low frequencies of the services. The actual time of departure and arrival will differ from the desired levels. This implies that the traveler faces various problems before and after his trip.

Before the trip, the traveler can choose between:

1. **Interrupting an activity.** He cannot complete an activity before the trip at the desired time.
2. **Canceling a desired activity.** Because an activity cannot be completed, it is cancelled leading to a choice of a less desired activity such as waiting.

After the trip, the traveler can choose:

1. **Coming late.** He does not start the activity at the desired time implying a loss of utility of the activity;
2. **Waiting.** He cannot immediately start the activity he most prefers and has to fill his time with another activity.

It is possible to distinguish between *scheduled* and *non-scheduled* activities. For example, a meeting can be conceived as a scheduled activity with fixed times for starting and finishing. Watching a video is a non-scheduled activity that is not linked to any particular time. In the case of non-scheduled activities, a change in their duration only leads to a change in utility related to their duration (for example leisure time is traded off against waiting time). In the case of scheduled activities, it is not so much the duration of the activity that matters but the fraction between actual and scheduled start or end times.

As said before, the frequency of the two flights determines the rescheduling time. In the perfect situation, the rescheduling time is the minimum connecting time t_c' :

$$t_s = t_c' \quad (5.20)$$

whereas in the worst case, in a period T , the time spent at the station is going to be:

$$t_s = t_c' + \frac{T}{F_2} \quad (5.21)$$

More generally this can analytically be expressed as

$$t_s = t_c' + (1 - \alpha) \frac{T}{F_2} \quad (5.22)$$

Where α is called a **timetable coordination coefficient**. If this term is equal to 1, then there is perfect coordination of the schedule. If it is close to 0.5, there is no coordination. If instead, its value is close to 0, then there is counter-productive coordination. Better coordination implies lower waiting time, and hence a better schedule.

Given any movement this factor is calculated as

$$\alpha = 1 - (t_s - t_c') \frac{F_2}{T} \quad (5.23)$$

Considering the connection of two flight legs fl_1 and fl_2 , the frequency of the second flight is 1 ($F_2 = 1$), the t_c' is the minimum connecting time adjusted by the connection slack time ($t_c' = t_c + t_v$), and the time T in the station is t_s , then

$$\alpha = \frac{t_c'}{t_s} = \frac{t_c + t_v}{t_c + t_v + t_w} \quad (5.24)$$

The factor α is therefore a ratio of effective connecting time over time spent at the station. R_{con} is a weighted ratio of the time spent at the station over the effective connecting time:

$$R_{con} = \frac{t_c + t_v + \frac{W_w}{W_c} t_w}{t_c + t_v} e^{\frac{1-p}{p}} - 1 \quad (5.25)$$

Therefore the factor α introduced by Rietveld is another way of calculating the ratio R_{con} . The close relationship between the Option Value at the Connection and R_{con} , indicates that OV is also an indirect way of calculating the schedule coordination.

In Rietveld (2001) it has been shown that the frequency of flights is not the only relevant factor determining average waiting times. Smaller airports can out-compete larger ones by better co-ordination of flights.

In more detail it has also been found that large airports such as London Heathrow and Paris Charles de Gaulle have longer waiting times than the (slightly) smaller airports Frankfurt and Schiphol even though one would expect shorter waiting times given the higher frequencies of service. The reason is that the minimal connecting time is higher and flight co-ordination is less efficient. The latter is clearly reflected by the values of the flight co-ordination coefficient developed above. Obviously there is a trade-off between long minimal connecting times allowing airlines slack in their hub operations and the probability of missed connections.

5.6 Option Value for a Flight

As reported in Carey (1994), a flight can be considered as a transportation chain. Given the profile of the waiting time on the station given the interdependence of the flights, and the profile of the trip time, then the **Option Value for a Flight** is formulated as the costs associated with the travel as a sum of:

1. expected cost of travel time on links and waiting time at stations/stops,
2. expected cost of arriving late or earlier than scheduled at a stop,

3. expected cost of departing later or earlier than scheduled from a stop, and
4. expected cost of having stops.

The final ODV is calculated in this form:

$$OV(f) = \sum_{fl \in f} OV_{leg}(fl) + \sum_{fl \in f} OV_{stn}(fl_i, fl_{i+1}) + w_s |S(f)| \quad (5.26)$$

where

- $S(f)$ is the set of the stations touched by the flight f excluding the origin and the destination.
- w_s is a weight for a stop. This is how much a passenger is willing to pay in utils for an extra stop. Due to the fact that the extra stop is introducing some extra ground time and extra flight time, this penalty results negligible compared to the extra disutility introduced. Hence, it is not reductive to consider $w_s=0$.

As before there is an **Ideal Option Value** associated with it that comes out analytically from:

$$OV^{ideal}(f) = \sum_{fl \in f} OV_{leg}^{ideal}(fl) + \sum_{fl \in f} OV_{stn}^{ideal}(fl_i, fl_{i+1}) + w_s |S(f)| \quad (5.27)$$

And from here the ratio

$$R(f) = \frac{OV(f) - OV^{ideal}(f)}{OV^{ideal}(f_2)} = \frac{OV(f)}{OV^{ideal}(f)} - 1 \quad (5.28)$$

is the final disruption value ratio for the flight f . As before the lower this ratio, the better the flight.

5.7 The Cost Term of a Trip

Unreliability has to be introduced in cost terms if one wants to trade-off the costs of unreliability with other transport costs, such as the cost of scheduled travel time on links and waiting time at stations/stops. Carey (1994) explains that unreliability-related costs are the expected costs of arriving later or earlier than scheduled at a stop, and the expected costs of departing later or earlier than scheduled at a stop.

There are tradeoffs that air travelers make when they choose among different carriers, flights, and fare classes. Previous results provide measures of the premium that business and leisure travelers are willing to pay to avoid schedule delays.

As reported by Prousaloglou (1999), travelers seek to maximize their air travel utility by choosing the air travel option with the highest utility or lowest disutility.

Travelers faced with the joint choice of carrier, flight, and fare class are therefore likely to make tradeoffs among a carrier's overall service, the convenience of its particular flight schedule, and the fare levels and service attributes of each fare class.

The proximity of each scheduled flight to a traveler's preferred departure time is reflected in the "schedule delay" measure that quantifies the convenience of each individual flight to individual travelers. In the air-travel literature, schedule delay has been defined as a measure of convenience related to the difference between preferred and scheduled flight departure times (Douglas and Miller, (1974)). The schedule delay concept has been used to calculate an index that measures the convenience of air service taking into account all "time-related costs" associated with air travel (Bailey et al., 1985). A similar concept is used in empirical studies to estimate travelers' sensitivity to the average time between scheduled departures (Morrison, (1985) and Morrison, (1986)).

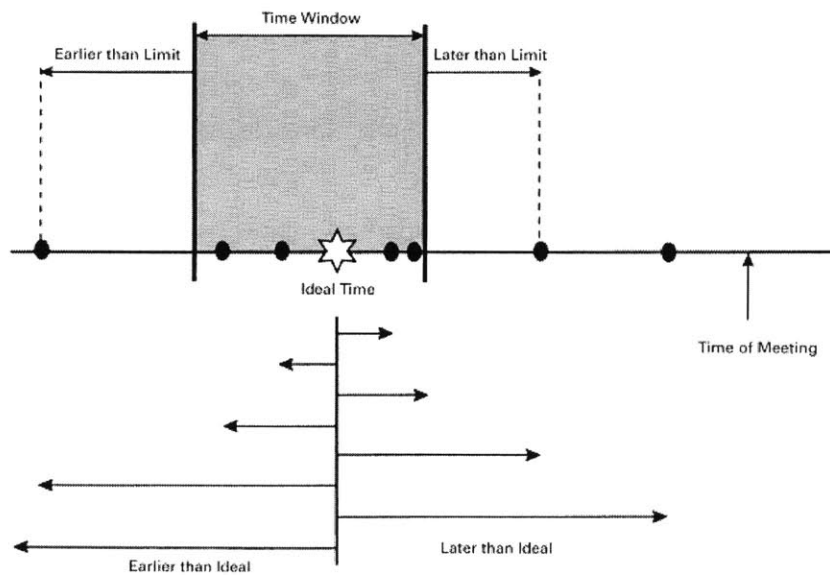


Figure 5.4 Measurement of Schedule Delay: Passenger's Decision-Making Process.

The empirical analysis assesses the relative importance of factors that influence travelers' choice behavior and identifies differences in the relative importance of those factors among segments of the air travel market.

The relative importance that travelers place on the additional time spent unproductively at the destination is equivalent to the costs associated with arriving later than preferred. The difference in the coefficient values by trip purpose reflects business travelers' expected greater sensitivity to schedule delays.

Travelers' sensitivity to schedule delay was further explored by more detailed non-linear formulations of schedule delay. These models indicate a greater sensitivity associated with early departure than with late departure flights. Furthermore, travelers were much more sensitive to schedule delays associated with flights that departed outside their own preferred departure time window, particularly for flights departing before their earliest convenient departure time.

This was accomplished by using a generalized cost of travel, which imposes a constraint on fare sensitivity by using values of schedule delay of \$40 per hour for business travelers and \$10 per hour for leisure travelers. These values are lower than those implied by the fare and schedule delay coefficients in the original model formulation (\$60 and \$17 for business and leisure travelers, respectively). The higher estimated values of schedule delay result from the fare coefficients. These are underestimated as the effect of increased fare is partially offset by un-included service amenities and reduced travel restrictions associated with higher fare classes. The imposition of value of time constraints and the introduction of fare class constants differentiates between the advantages of the higher fare classes and travelers' sensitivity to the cost of air travel.

The result of these papers is that an uncertain minute is worth 2.4 times than the certain one (regular trip time). In addition to that, it is clear that passengers are willing to pay an average of 25% more for a more reliable trip.

Other parts of the trip were considered as well and the results clearly showed that a waiting minute on a platform (airport) is worth 1.5 times more than a minute waiting in-vehicle.

Gate time, taxi-in, taxi-out, and stops were not considered. However we have included them by fixing the value of the minimum connecting time as part of the trip (weight equal to 1), while gate time, taxi in and taxi out can be assumed to be 1.25 because they are the midway between platform and flying time.

The important value is the weight associated with the delay time. As seen before, a value of 2.5 implies a factor of 40% of the trips arriving late. In this case, late means after the schedule arrival time, hence a more relaxed convention can be used in which later means *15 minutes later than the schedule arrival time*.

Another way to manipulate such a value is to improve the on-time punctuality to be 82% of the time; hence that value is equal to 5.6

Nevertheless, when fixing the percentage of late flights (with the appropriate justification for lateness), then the various option values can be calculated.

5.8 Option Disruption Value

The calculation of an option value for every flight gives an indication of the quality of the flight compared to the others. As described in the uniformity characteristic of the reliability measure, the option value allows direct comparisons between flight serving the same markets, and a relative indication of quality for comparisons of flights serving different markets.

When making a decision on which flight to select to go from a to b the selection process is easier and more direct if for every option it is given its option value. However, if there is a need to reschedule a flight, then the situation changes. In such a case, the original plan is disrupted and, therefore it is necessary to find another flight in relationship to the disrupted flight to divert passengers. Considering the disrupted flight relative to any other option can do this. The options in consideration are, very likely, the protecting option put in place during the design process.

The comparison operates based on:

1. The difference in Option Value. The evaluation of the two flights can be done directly because they are serving the same market
2. The difference in departure time. This introduces another level of disutility in the system to the traveling passenger, based on the description outlined in the previous section.
3. The difference in arrival time. Generally, passengers prefer to arrive to destination as close as possible to the original arrival time. The displacement in time is the one that affects them most.

A formal definition of the option disruption value for two flights is as follows:

$$ODV(f_1, f_2) = OV(f_1) - OV(f_2) + w_d |T_d(f_1) - T_d(f_2)| + w_a |T_a(f_1) - T_a(f_2)| \quad (5.29)$$

where:

- $OV(f)$ is the option value of flight f .
- $w_d T_d(f)$ is the scheduled departure time of flight f with its weight.
- $w_a T_a(f)$ is the scheduled arrival time of flight f with its weight.

As described in the cost section, the weight should be identical and set as much as the delay weight of, for example, 2.5.

The lowest level of option disruption value indicates the absolute best choice for travelers. Notice that the option disruption value can be 0, positive or negative. When it is 0 then there is no disruption between the flights in consideration, and therefore they are interchangeable. When it is negative, then f_2 is better than f_1 because it has a lower and better option value that compensates for the displacement in arrival and departure time.

5.9 Analysis of OV and ODV Characteristics

The reliability measurements that will be used from here on are the Option Value and the Option Disruption Value. Both the Option Value and the Option Disruption Value incorporate the characteristics required by any good reliability indicator. In particular these are:

1. **Simple:** The concept behind the OV and ODV is related to the consideration of behavior of a flight related to its “*neighborhood*.” The formula itself is a weighted combination with an exponential penalty that comes into play only when factors like cancellation and load are very high.
2. **Statistical aspect:** The profiles are from historic data. The assumed distribution is corroborated by some logical consideration. An estimate of the probability distribution parameters is done via rigorous methods (Maximum Likelihood Estimation) and formal fit test are used (Kolmogorov-Smirnoff and X^2).
3. **Rules of Thumb.** In addition to the simple weighted combination of the time spent on the various trip/itinerary parts, there is also a combination of cancellation rates (higher the chances of cancellation, higher the value of the disruption cost) and load factor (the higher the load, the less chance to get a seat, and hence the higher the disruption cost).
4. **Market/trip Independence.** Flight considerations are compared to the performance of a collection of other flights leaving or arriving the same station. The large share of operations considered allows giving an average behavior in which the penalty of “very bad” flight, or “extremely good” one is attenuated. In particular when analyzing a new itinerary the flight in consideration does not

necessarily exists: the idea is that if it existed it would follow the profile outlined by its *neighborhood flights*.

5. **Uniformity.** The flights and flight leg performance to be analyzed are compared to the performance of other flights and flight legs that depart and arrive at the same departing and arriving station. Hence direct comparison is possible and meaningful. When comparing different flights, they should have common characteristics such as origin and destination. The normalization of the OV via a ratio such as variation of option value with a base measurement defined Ideal Option Value can give the relative extra burden imposed on passengers on such options. This allows comparisons between flight operations in totally different markets.
6. **Unit Independence.** Due to the characteristic of the formula used the only to quantify the ODV is through a concept of **utility units (utils)**. This is very flexible and can be translated into financial value when the monetary utility value of a trip for a passenger or of the airline is identified.

Other reliability indicators are the ratios R_{con} , R_{leg} , and R . Furthermore they are defined from 0 to $+\infty$: a higher value implies a higher burden imposed on the passengers during the flight in consideration. They can be considered as good reliability sign because they are derived from the Option Value.

5.10 A Posteriori Schedule Effectiveness Measures

When a schedule has been adjusted, it is necessary to determine how well the new schedule will perform. The above reliability indications are operationally good, but they are not as efficient as a posteriori explanation of the event after it has occurred. This kind of analysis should be independent from the process of schedule creation/readjustment, and should be objective and based on characteristics and *output* of the schedule itself. If that is done properly, comparisons between the old and new schedule can become unprejudiced.

Considerations about coverage, robustness, schedule reliability and flexibility are not easily determinable because they will be the main focus only on the new schedule and not on a pre-existing one.

Objective consideration between various schedules given a fixed period of time (like one month) can be obtained as follows:

1. **Cancellation.** This is the calculation of the number of flights and legs cancelled. This can also be translated into broken trips in defined markets. A more objective measure is a percentage of legs/flights/itineraries cancelled in a period of time.
2. **Total passengers' delay.** This is the number of passengers that have been disrupted on the network. They are divided into:
 - a. **Arrived delayed:** number of passengers who arrived at the destination following the itinerary planned.
 - b. **Allocated to other flights.** Number of passengers who arrived at the destination but needed to be relocated to other legs/flights.
 - c. **Spilled to another airline.** Estimation of the number of passengers who arrived at the destination without using the planned legs and then diverted to another airline.
 - d. **Lost passengers.** These are the passengers who decided to give up on their trip and hence did not travel.
3. **Total passenger-minutes delay:** The calculation of the passengers is as described above. For each of them the calculation of the delay should be:

$$D_p = \sum_{p \in Pax} d_p \quad (5.30)$$

where d_p is the delay experienced by every single passenger p . The delay to consider should be only the positive ones, late arrival or departure only.

4. **Total passengers' disruption delay.** This is the offset that the passengers have experienced from their original plan. A measure of this is the passenger delay; however, the negative delay (arriving early) is counted as a disruption as well, even if its disutility value is less than the late arrival one. This can be obtained as follows:

$$D_d = \sum_{p \in Pax} |d_p| \quad (5.31)$$

5. **Total Network Value.** This is the value experienced by every single passenger using every option on the network. It is calculated as follows:

$$D_{OV}(fl) = N(fl) \cdot \left(w_{gt} t_{gt} + w_{to} t_{to} + w_f t_f + w_{ti} t_{ti} + w_s |t_s - t'_s| \right) \quad (5.32)$$

$$D_{OV}(fl_1, fl_2) = N(fl_1, fl_2) \cdot (w_c(t_c + t_v) + w_w t_w) \quad (5.33)$$

$$D_{OV}(f) = \sum_{fl \in f} D_{OV}(fl) + \sum_{fl \in f} D_{OV}(fl_i, fl_{i+1}) + N(f) \cdot w_s s \quad (5.34)$$

$$D_{OV}(S) = \sum_{f \in S} D_{ODV}(f) \quad (5.35)$$

where

- S is a schedule
- f is a flight
- fl is a flight leg
- N is a function that returns the number of passenger using the flight f , flight leg or connection.
- The times used are the real times experienced by the passengers of the relative flight leg, flight and connection, with the proper weight.

The above metrics are called **Schedule Quality Measures (SQM)**. Using the SQMs in combination with the average fare, and/or the fare classes, they can be converted into the financial benefits or the monetary disruption costs of the schedule in consideration. This can be used for assessing cost/benefits analysis.

If, instead, the SQMs are used in combination of monetary costs of adding extra “utility” to the schedule, they can be used to assess the trade-offs analysis of robustness in the schedule by the airline planners.

5.11 Conclusions

Various ways of implementing the reliability measures have been presented. Both a priori and a posteriori valuation methods have been discussed. Furthermore two ways for analyzing the option selection and the option valuation have been outlined.

Considering the reliability measured presented, the next logical step is to identify the actions to perform on the schedule, and where to apply them. This will be discussed in the next chapter.

Chapter 6

Schedule Improvement

The previous chapters presented a generic outline of the process for improving Schedule Reliability. In this chapter there is a more detailed analysis of the operations that can be performed on the schedule. These actions cover the Flight Leg Adjustment Problem, the Multiple Flight Allocation Problem, the Multiple Flight Redundancy Problem and the Station Independence Problem.

The solution of the Critical Station Selection Problem and Critical Flight Selection Problems is performed through a heuristic analysis of the US domestic schedule of July 2000. While this analysis will be valid only for July 2000, can be repeated for any month.

6.1 Operation on the Network

When the various flights and flight legs are graded, it is possible to modify the schedule to achieve a *better level of service*. The level of service offered by the schedule can be quantified by the following measures:

1. **Option Value.** The disutility value of flights and flight legs based on passenger indications. A lower OV implies a less disruptive leg and the service being closer to the passenger's expectations; a lower OV also implies a better schedule.
2. **Option Seats.** The number of seats allocated in a flight leg that accommodates disrupted passengers. The higher the number of seats allocated, the better the schedule obtained.
3. **Coverage.** The number of itineraries that allow the passenger to reach the destination. A higher coverage implies a better chance to recover any disruption and therefore a better schedule.
4. **Robustness.** The same as the coverage but the itineraries considered do not overlap. Higher robustness implies more independent ways to reach the destination, and then a more *fault tolerant* schedule.

5. **Reliability.** The capability of the airline to keep up with the schedule. Reducing the number of cancelled flights and/or increasing the on-time performance affects the reliability of the airline. Higher reliability implies a better schedule.

The operations that are allowed on the network to improve the measures above may be divided into 3 categories:

1. Change a single flight leg.
2. Change a flight.
3. Change the aircraft rotation.

Ultimately, these changes will improve the schedule, which will result in better service to passengers. Better service for passengers implies also a benefit to the airline expressed as:

1. **Image improvement.** The passengers' perception of the carrier as more reliable gives them confidence in the airline and may result in the airline being their airline of preference when selecting their next trip.
2. **Reduction of re-accommodation costs.** During disruptions, passengers need to be relocated to other flights or accommodated for a while when waiting the next available flight. In the worst-case scenario they need to have some incentives for not going away.
3. **Reduction of recovery cost.** If the schedule can better absorb delays and cancellations, then there will be a smaller number of flights to readjust and/or recover, and hence lower costs.
4. **Reduction of revenue lost (spillage).** If passengers are properly routed to the available options when disruptions occur, then they are most likely not to be affected by delays and cancellations; hence they will not look for other flight on other airlines and revenue will not spill to these other airlines.
5. **Reduction of unallocated seats.** Passenger's spillage and missed connections imply empty seats that otherwise could have been sold.
6. **Extra Revenue.** The introduction of new options offers new service that other passengers can take advantage of, and hence the new revenue coming from them.
7. **Fare Changes.** The increased reliability implies better service; hence it justifies a moderate fare increase. In Carey (1999), studies show that the fare increase can be up to 25%.

These benefits do not come for free. Some of the modifications proposed will generate extra costs or some revenue reduction. The main costs involved are the following:

1. **Crew.** Modification of the schedule induces rearrangement of the crew scheduling that might generate extra costs. However, the improved reliability can positively affect these costs.
2. **Operating Costs.** These are the typical costs experienced when flying an aircraft. Variation of the operating costs will occur only if there is a variation of equipment or a modification of the aircraft rotations. This is closely related to the increase or decrease of the aircraft utilization.
3. **Fixed Costs.** Like the previous, these are related to the equipment used. Variation of the equipment type implies a variation of such a cost.
4. **Revenue Reduction.** Revenue comes from flying passengers. Keeping the aircraft size fixed, the increase or decrease of the number of passengers allowed to board will affect the revenue coming from them.
5. **Gate Restrictions.** Changes in the flight schedule time imply a variation of the gates to use, thus the time restrictions and associated costs must be assessed.

Here below follows a more detailed description of the various actions divided by category: there is a highlight of the parameters that will change, the improvement obtained and the costs associated with them.

It is clear that if the modification that will follow will improve the airline schedule reliability, then the passengers will benefit from this process. The considerations about costs and revenue, on the other hand are expressed from the airline perspective: if the net difference between the value of the benefits minus the costs of the changes is positive, then the modifications are benefiting both passenger and airline; hence it is worthwhile activating them as soon as possible.

6.2 Flight Leg Changes

The changes in flight leg are limited to the single flight leg and are made to reduce the OV associated with that leg. They are:

1. **Shift Flight Leg.** Change the scheduled time of the flight leg. This shift is limited to small changes, e.g., ± 20 minutes, so that the demand associated with such flight leg will not change.

2. **Move Flights.** Like the above but the time shift is greater than ± 20 minutes. The demand can change.
3. **Larger Aircraft.** Utilization of different equipment. There is no change in the demand/load. This can affect the fleet assignment and also the rotations but it provides more seat options for the flight to protect.
4. **Protect Seats.** This does not modify maintenance rotation or fleet assignment but limits the load to a portion of the available seats. This reduces the revenue coming from this leg due to seats to protect, but they can be used for accommodating disrupted passengers.
5. **Reduce Slack.** Remove some flight slack time from the flight leg schedule in response to better flight performance. This increases coverage and improves reliability.
6. **Increase Slack.** Add some flight slack time in the schedule for the flight leg in consideration. The time is neither flown nor is it spent waiting. This, however, improves the connection and on-time performance.

The impact of these modifications on the airline operations is reported in Table 6.1. In the table, yes means that it impacts the referred part of the airline operation; and *maybe* that it may impact it.

Table 6.1 Flight Legs Actions: Impact on Schedule Design Process.

Impact	Shift Leg	Move Leg	Protect Seats	Use Larger A/C	Reduce Slack	Increase Slack
Crew	-	Maybe	-	Yes	Maybe	Maybe
Fleet	-	-	Maybe	Yes	-	-
Rotation	-	Yes	-	Yes	-	Maybe
Demand	-	Yes	-	-	-	-
Equipment	-	-	-	Yes	-	-
Load Factor	-	Maybe	Yes	Yes	-	-
Cancellation Out	Yes	Yes	-	-	-	-
Gate Profile	Maybe	Yes	-	-	-	-
Taxi Out Profile	Yes	Yes	-	-	-	-
Fly Time	-	-	-	-	-	-
Cancellation In	Yes	Yes	-	-	Yes	Yes
Taxi In Profile	Yes	Yes	-	-	Yes	Yes
Slack Time	-	-	-	-	Yes	Yes
Connectivity	Maybe	Yes	-	-	-	Maybe
Total Cancellation	N/A	N/A	N/A	N/A	N/A	N/A

Table 6.2 and 6.3 indicate where the modification will result in improvements and increased cost respectively.

Table 6.2 Flight Leg Actions: Schedule Improvements

Improvements	Shift Leg	Move Leg	Protect Seats	Use Larger A/C	Reduce Slack	Increase Slack
ODV	Yes	Yes	Yes	Yes	Yes	Yes
Option Seats	-	-	Yes	Yes	-	-
Coverage	Yes	Yes	-	-	Maybe	Maybe
Robustness	Maybe	Maybe	-	-	-	-
Reliability	Maybe	Maybe	-	-	Maybe	Yes

Table 6.3 Flight Leg Actions: Airline Costs and Restrictions Implications

Costs	Shift Leg	Move Leg	Protect Seats	Use Larger A/C	Reduce Slack	Increase Slack
Crew	-	Maybe	-	Yes	-	-
Operating Cost	-	-	-	Yes	-	-
Fixed Cost	-	-	-	Yes	-	-
Revenue Change	-	-	Yes	-	-	-
Gates Restriction	Maybe	Yes	-	Maybe	Maybe	Maybe

6.3 Flight Changes

The flight changes involve the flight as a set of legs. The flight changes, however, might be applied independently from the flight leg changes described in the previous section. There will be no discussion of changes to the single flight leg because this has been already analyzed: there are no assumptions that the changes highlighted before, have been performed in the schedule for the flight in consideration.

The actions to apply to the flights are:

1. **Ground Time.** Changes in the ground time to allow for better connections and coordination of the flight, and hence an improvement in the on time performance. This can be considered as equivalent to the Leg Shift presented in the Flight Leg changes above.
2. **Stops.** Add or remove an extra leg from the flight. This can imply an introduction of a direct service or a different connection path for the original destination.
3. **Connecting Legs.** Add legs that allow multiple flight connections at different times. This will give more connecting time and allow more itineraries.

4. **Multiple Options.** Add flight legs at the same place but at a different time. This allows continuation of the flight with other legs and a better passenger connection and re-routing option.
5. **Multiple Itineraries.** Rearrange connections through different stations. This implies a non-overlapping path from the origin to the destination of the flight in consideration. Coverage and robustness will be greatly affected.

The impact of these modifications on the airline operations is reported in Table 6.4. In the table, *yes* means that the action impacts the referred part of the airline operation; *no* that it will not impact it; and *maybe* that it may marginally impact it.

Table 6.5 and 6.6 indicate where the modification will result in improvements and increased cost respectively.

Table 6.4 Flight Actions: Impact on Schedule Design Process.

Changes	Ground Time	Stops	Connecting Legs	Multiple Options	Multiple Itineraries
Crew	Maybe	-	-	Yes	Yes
Fleet	-	-	-	Yes	Yes
Rotation	Maybe	-	-	Yes	Yes
Demand	-	Yes	Yes	Yes	Yes
Equipment	N/A	N/A	N/A	N/A	N/A
Load Factor	N/A	N/A	N/A	N/A	N/A
Cancellation Out	N/A	N/A	N/A	N/A	N/A
Gate Profile	N/A	N/A	N/A	N/A	N/A
Taxi Out Profile	N/A	N/A	N/A	N/A	N/A
Fly Time	N/A	N/A	N/A	N/A	N/A
Cancellation In	N/A	N/A	N/A	N/A	N/A
Taxi In Profile	N/A	N/A	N/A	N/A	N/A
Slack Time	N/A	N/A	N/A	N/A	N/A
Connectivity	Yes	Yes	Yes	Yes	Yes
Total Cancellation	Yes	Yes	Yes	Yes	Yes

Table 6.5 Flight Actions: Schedule Improvements

Improvements	Ground Time	Stops	Connecting Legs	Multiple Options	Multiple Itineraries
ODV	Yes	Yes	Yes	Yes	Yes
Option Seats	-	-	-	Yes	Yes
Coverage	-	-	-	Yes	Yes
Robustness	-	-	-	-	Yes
Reliability	Yes	-	Yes	Yes	Yes

Table 6.6 Flight Actions: Airline Cost and Restriction Implications

Costs	Ground Time	Stops	Connection Legs	Multiple Options	Multiple Itineraries
Crew	Yes	-	-	Yes	Yes
Operating Cost	-	-	-	-	-
Fixed Cost	-	-	-	-	-
Revenue Reduction	-	-	-	Maybe	Maybe
Gates Restriction	Maybe	-	-	Maybe	Maybe

6.4 Rotation Changes

Aircraft routing is the less constrained, thus ad hoc re-routing can improve the protection of the flight leg and the flights composing it. The changes in the rotations are based on keeping the schedule fixed but rearranging the aircraft rotation alternating legs as a way of balancing delays in the whole system. These changes can be summarized as:

1. **Leg Alternation.** Alternate tight and *likely to be delayed* flight legs with flight legs that can absorb delay. This will give more breath to the rotations, redistribute delays and then improve reliability.
2. **Leg Padding.** Change the turnaround time between legs of the same rotation to absorb delays. This is a variation of the above but is implemented when there are no legs that can absorb delays.
3. **Leg Isolation.** Isolate the flight leg that generates delays into a restricted market, and keeping the rotation focused on the legs. This will avoid having cancellations or delay in flights that are not directly related to the original delay.
4. **Looping.** Introduce looping aircraft between certain stations with the purpose of protecting sensitive flights.

Every action reported above has the purpose of not propagating delays throughout the aircraft rotation. The *padding, alternation and isolation* are all apt to limit or avoid delay proliferation. The changes affect the leg intermix on a single rotation, constrained only by the fleet assignment and the maintenance: the objective is, therefore, still to cover every leg and to allow every aircraft to go to maintenance after a determined amount of time; however, the extra “constraint” is focused on the statistical delay propagation through the legs composing the rotation. The fact that the rotations will still cover the whole network, and the fleet assignment will be maintained, implies that:

1. There are no changes in the schedule design process impact and benefits described in Table 6.1 or 6.4.
2. It is not possible to describe the impact on the schedule process as done before in Table 6.1 and 6.4 because the rotation actions are done a priori, and they do not affect the reliability metric (OV and ODVs).

On the other hand it is still possible to determine if the changes will result in general system improvements. This is reported in Table 6.7. The areas where added costs associated with the rotation action are likely to occur are reported in Table 6.8.

Table 6.7 Rotation Actions: Schedule Improvements

Improvements	Leg Alternation	Leg Padding	Leg Isolation	Looping
ODV	Yes	Yes	Yes	Yes
Option Seats	-	-	-	Yes
Coverage	-	-	-	Yes
Robustness	-	-	-	Yes
Reliability	Yes	Yes	Yes	Yes

Table 6.8 Rotation Actions: Airline Cost and Restriction Implications

Costs	Leg Alternation	Leg Padding	Leg Isolation	Looping
Crew	Maybe	Maybe	-	Yes
Operating Cost	-	Maybe	-	-
Fixed Cost	-	-	-	-
Revenue Reduction	-	-	-	-
Gates Restriction	-	Maybe	-	Yes

6.5 Action Location

The changes highlighted above can primarily be applied to legs, flights and rotations. The best schedule can be achieved via activating all the changes above to every leg, flight and rotation: this, however, is very impractical.

It is possible, however, to isolate legs, flights and rotations that are more sensitive and more important than others. These are the parts of the schedule that should be protected most. The selection of them is primarily based on internal airline considerations; however, a statistical analysis of the schedule can highlight the parts of the schedule that are most disrupted.

The identification of flight legs, flights, and rotations is the primary output of the *Critical Flight Selection Problem* and *Critical Station Selection Problem* presented in the previous chapter. The results presented below were obtained via statistical analysis. The following paragraphs, therefore, will be focused on elements of the schedule that are impacted most by cancellation and delays.

The most interesting parts of the schedule are:

1. stations that impact the network more than others,
2. legs that are statistically more delayed or cancelled than others,
3. flights that are delayed or cancelled more often,
4. areas or concentrations of station that are impacted most,
5. rotations that consistently have high delay accumulation, and
6. legs that contains higher loads.

The last point comes from the idea that the severity of the disruptions are measured by the number of passengers affected, and therefore it makes more sense to act where most of the passengers are traveling.

The following statistical analysis is made on the whole schedule of all the major airlines flown during the month of July 2000. The same process can easily be extended to any month of any year, and/or focused on any single airline.

6.6 Stations Analysis

The station analysis is done to identify the stations that have experienced more delay and cancellation during the period being examined (July 2000). When such stations have been identified it is necessary to determine if it is necessary to protect them. The protection is done via the station actions, or by considering all the legs going through these stations and applying the flight leg actions.

A total 193 stations and 475,005 flights were analyzed. The stations fall into four categories, based on the severity of the disruption level:

1. Heavy delay or cancellation in highly important and network-influential stations. This is the case of Chicago (ORD), Atlanta (ATL), and La Guardia (LGA).
2. Heavy delay or cancellation on low traffic/importance/influential stations. These stations have a high percentage of cancellation and delay compared to the traffic. These should not be used for protecting other stations, and might not be

worthwhile protecting. Unalaska (AK, symbol DUT), Dillingham (AK, symbol DLG) and King Salmon (AK, symbol AKN) are some example.

3. Important non-influential station with high cancellation or delay ratios. The ratio of cancellation/delay over traffic is high, even if there is substantial traffic. These stations, in most cases are worth protecting. Example Boston (BOS), Seattle (SEA) and Philadelphia (PHL).
4. High traffic stations. These stations are worth protecting for the simple reason that there is significant passenger traffic going through them. If they are protected then most of the flights that will use them will be protected as well. Consequently a lot of passengers will be shielded from disruption. Examples of these stations include Atlanta (ATL), Baltimore (BWI), Boston (BOS), and Charlotte (CHR).

A detailed analysis of the stations is presented below.

6.6.1 Stations: Cancellations

An analysis of station cancellation was performed based on the number of flight legs cancelled at that station in the month of July 2000. The cancellation distribution for each station was identified. A flight that is cancelled will be counted as such in both origin and destination station. A diverted flight is counted as canceled only at the destination station. Figure 6.1 and 6.2 illustrate the cancellation distribution.

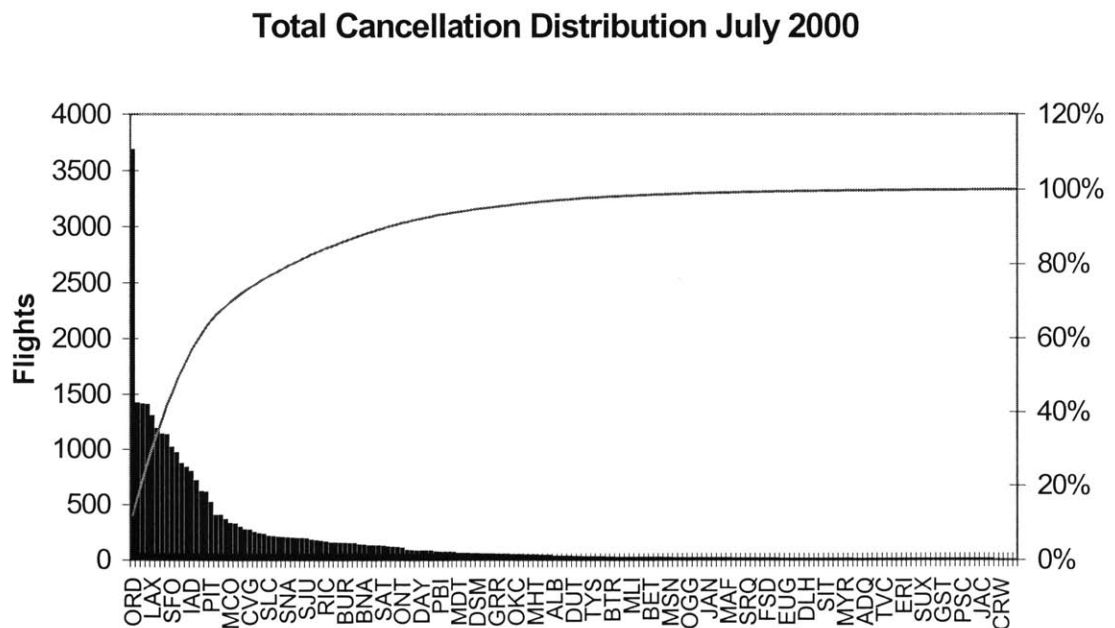


Figure 6.1 Total Cancellation Distribution July 2000

Number of Cancellations at Stations July 2000: First 80%

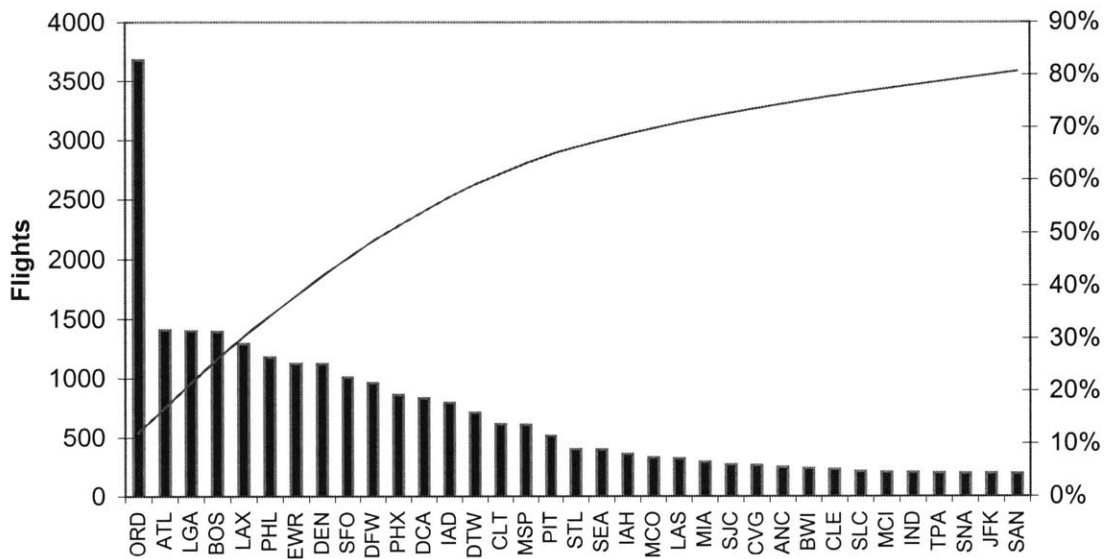


Figure 6.2 Distribution First 80% of the Cancellation in July 2000

These charts highlight the fact that 80% of the cancellations are generated by 18% of the stations. In more detail Chicago (ORD), Atlanta (ATL), New York-La Guardia (LGA), Boston (BOS) and Los Angeles (LAX), that are only the 2.6% of all the stations, are responsible for 30% of the total cancellations.

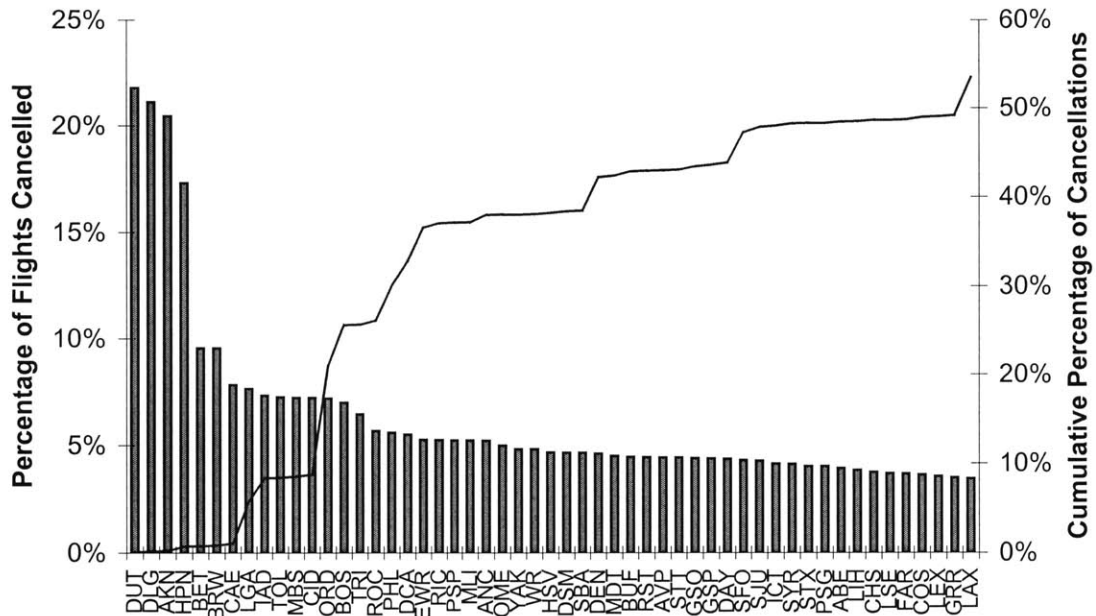
The chart in Figure 6.3 can be obtained by analyzing the cancellations as a function of the total traffic on the single station. The figure shows that some stations have almost 1 flight out of 5 cancelled. This is difficult to interpret because it could be that:

1. The stations are over-served, and then such cancellations are irrelevant.
2. There is not enough demand to serve. Hence canceling any flight is not affecting the flying public.
3. There are only a few flights; hence, only a few cancellations can heavily impact the total operation statistic.

These charts in figures can be read from two perspectives: the first one is that these are the stations that are more unreliable in term of cancellations of the whole network during the month of July. The departure cancellation rate and the arrival cancellation rate are very high, and therefore these are the stations to avoid when rearranging the schedule.

The other perspective in which these data can be read is that these are the stations more prone to cancellation during the month of July, and therefore they are the ones the schedule should protect most.

Cancellation Compared to Station Traffic July 2000



The analysis is done in two ways: the first is to consider the total square flight delay (Figure 6.4). This gives an indication of how consistent are the flights with their scheduled departure and arrival time. The second is done as an average square delay per flight (Figure 6.5). This accounts for the number of movements at the station. Normalizing for the traffic volume allows for uniform comparison of larger and smaller stations.

Departure Schedule Variability Using Delay² in July 2000

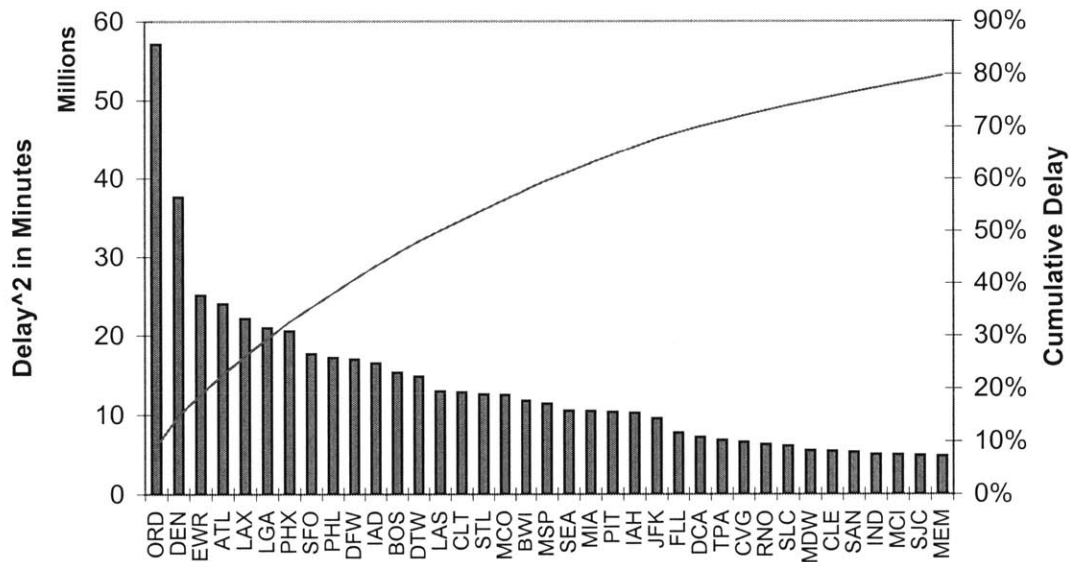


Figure 6.4 Departure Schedule Station Variability

Even in this case, the charts indicate that only a small percentage of stations are responsible for the majority of the variability, and hence, the delays.

The charts in Figure 6.4 and Figure 6.5 report the total departure delay and the average departure delay. In a total of 638M total minutes square delays, 80% of the network delay is attributed to only 18% of the stations (36 stations). Analyzing, instead, the average delay then the picture changes: in this case only 45% of the total delay is attributed to 24% of the stations.

Departure Schedule Variability in Average Delay² in July 2000

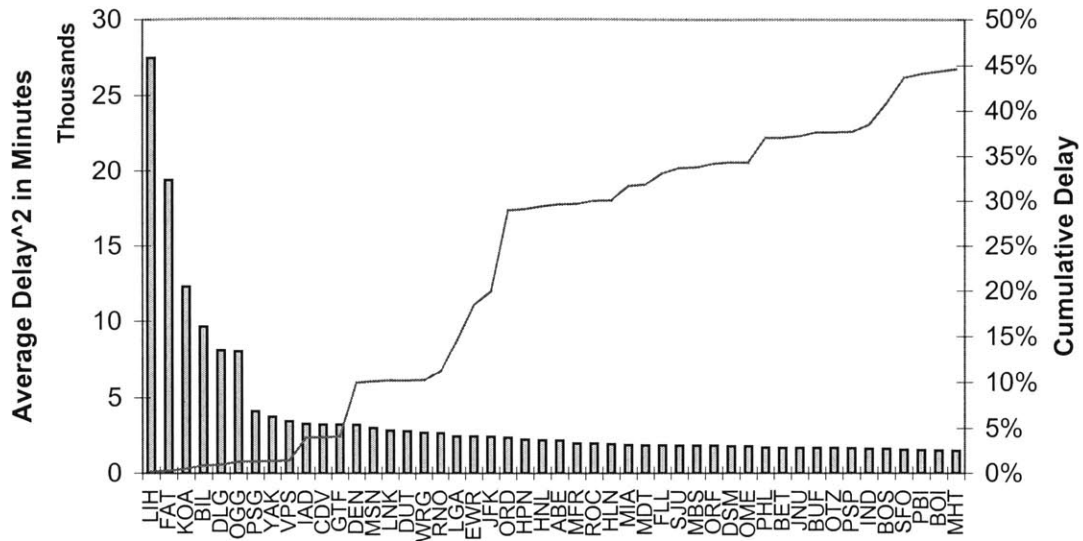


Figure 6.5 Departure Schedule Station Average Variability

Notice that the sets of stations are different: the stations that contribute most on total delay are not the ones that have the highest average delay.

The same can be observed in arrival. Figure 6/6 and 6.7 presents that. The total square delay is of about 761M minutes. Also in this case 80% of the total square delay is attributed to 35 stations (18%). Considering instead the average delay case, we see that 50% of the total delay is made up by 45 stations (23%).

Notice that even in this case the set of stations is different, as is the set of stations driving the arrival and departure delay.

This data indicates that only a few stations need to be avoided or protected to preserve network performance.

Arrival Schedule Variability Using Delay² in July 2000

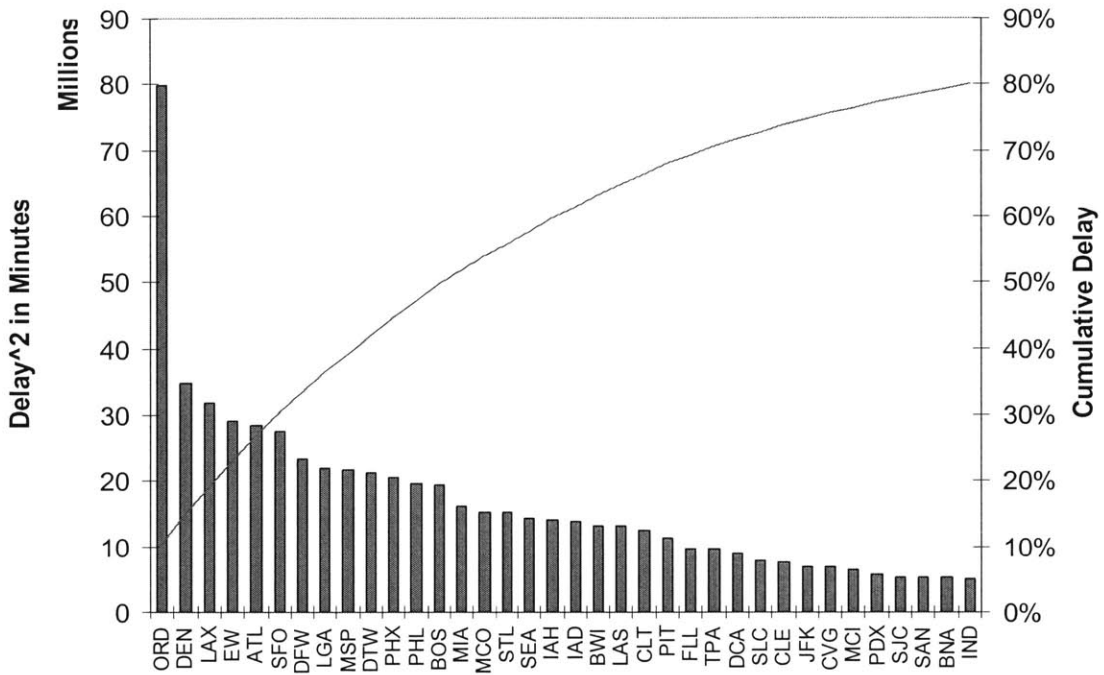


Figure 6.6 Arrival Schedule Station Variability

Arrival Schedule Variability in Average Delay² in July 2000

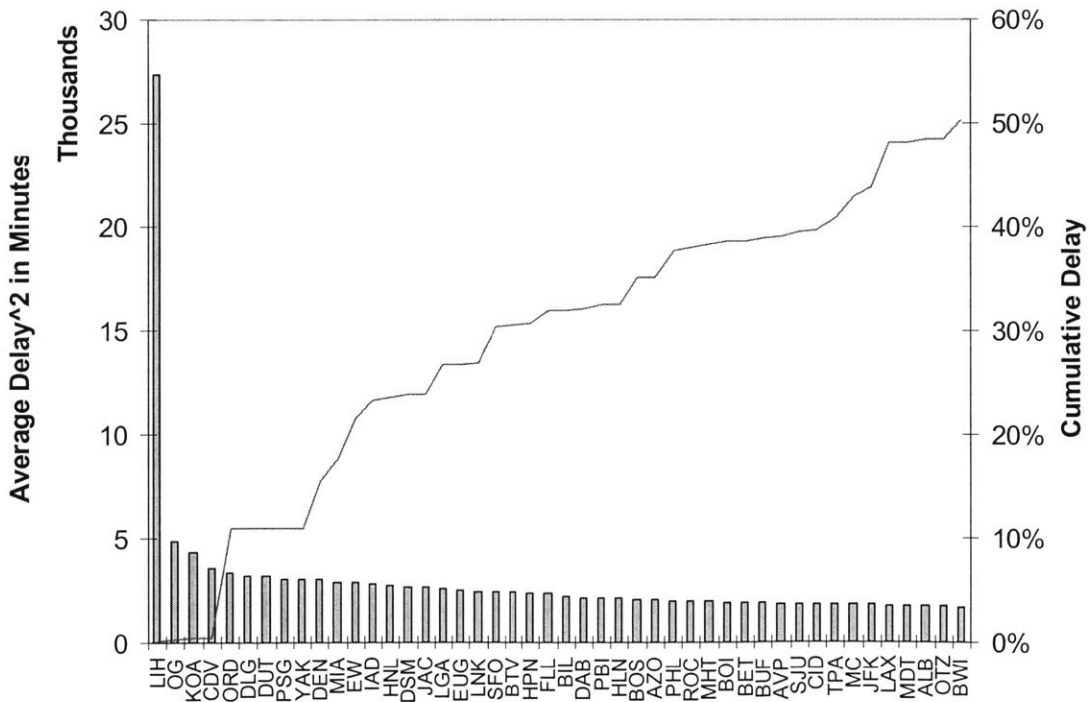


Figure 6.7 Arrival Schedule Station Average Variability

6.7 Flight Leg Analysis

An analysis of all the flight legs in the ASQP database for the month of July 2000 is presented below. For this month the total number of distinct legs (OD) planned is 2124 and the total number of legs scheduled is 423,731.

As with the stations, a small proportion of legs is responsible for most of the cancellations and delays. The selection of the legs to protect is not clear because we have no insight into the airline strategy or the profitability of the legs themselves.

Thus, the analysis focused on the time variability: the higher the variability, the more protection required. At the same time, the very same legs should not be used for protecting any other flights. It is interesting to note that in certain cases there are some very simple adjustments that can be put in place to improve reliability. In particular when there is a flight with high average delay and very low variability, a way of improving its reliability is to increase the scheduled trip time.

6.7.1 Flight Legs: Cancellations

The number of legs that were cancelled or diverted is 16,361, which is 4% of the total number of scheduled legs scheduled. Figure 6.8 and 6.9 shows the OD pairs with the most cancelled legs. Only 2.2% of the OD pairs account for almost 25% of the total legs cancelled. In particular the OD pair Los Angeles – San Francisco (LAX-SFO) is the one showing the most cancellations (340 out of 2,612 legs, equal to 12% of the total traffic).

The chart in Figure 6.9 presents the number of cancellations in relation to the traffic for such legs. An astonishing 1 flight in every 3 flights is cancelled on the Chicago - Milwaukee leg (ORD-MKE), even though there are less than 20 flights during the month. On the other hand, 1 flight in every 5 flights is cancelled on the Boston - Washington leg (BOS-IAD) where there were more than 120 flights (4 a day) during the month of July 2000.

It is not clear if these cancellations are due to lack of demand or to the fact that the market was over-served. In the first case, there is a limited impact on the flying public (likely in ORD-MKE) because, arguably, there is a limited demand; in the second, the high frequency reduces the impact on the flying public too, due to the ready availability of another flight “right after” the one cancelled. Nevertheless, it is clear that these may be the flight legs to protect or to avoid when trying to guarantee connectivity and reliability.

Legs Cancelled July 2000

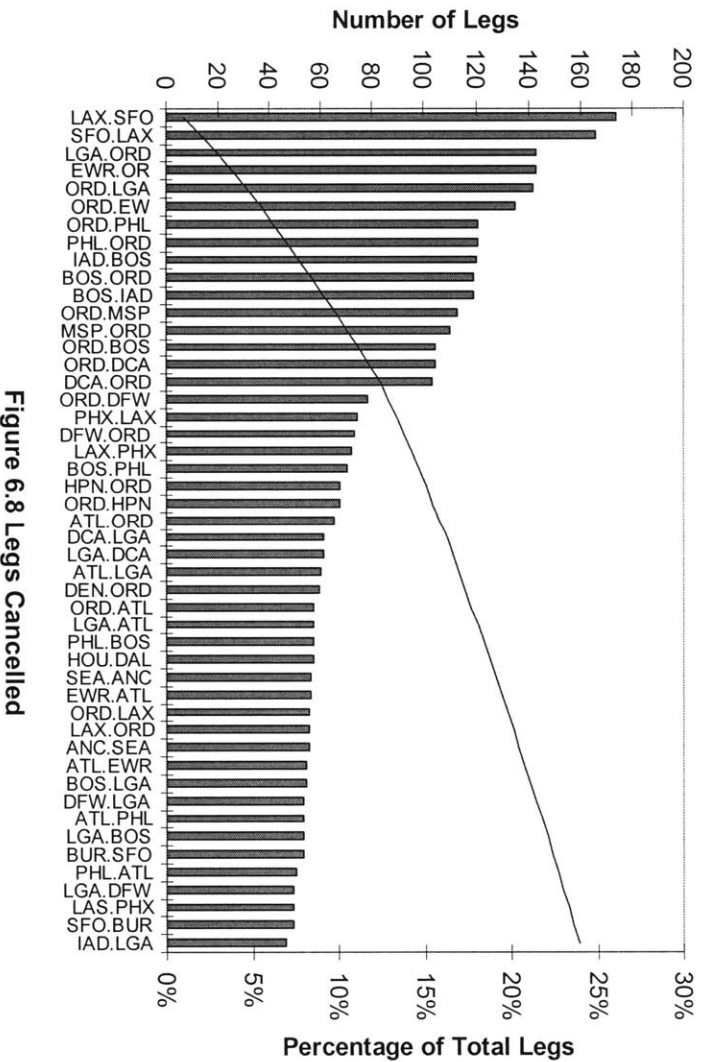


Figure 6.8 Legs Cancelled

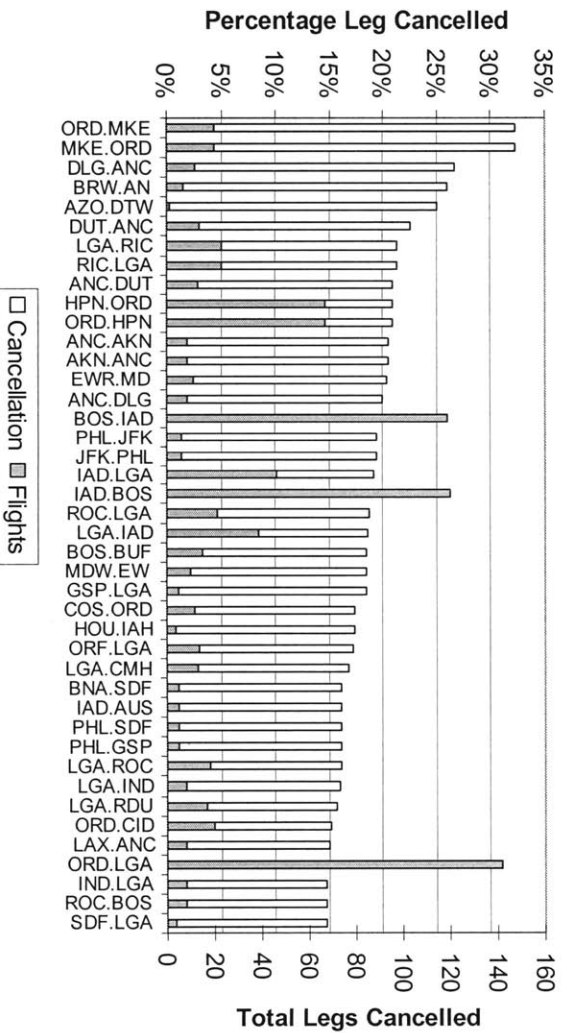


Figure 6.9 Proportion of Legs Cancelled

6.7.2 Flight Legs: Delays

The analysis of the leg delays focuses on:

1. **Departure Delay:** the delay experienced at the gate.
2. **Arrival Delay:** the schedule delay obtained as a difference between the schedule arrival time and the actual arrival time; this is greatly affected by the departure delay.
3. **Schedule Delay:** how the summation of on-gate, taxi-out, fly time and taxi-in time affects the overall flight performance. This is obtained as a delay discounted by the departure delay.
4. **Airborne Variation:** this is the variation of the in-air flight time. This is important for understanding how the flight is disrupted during the trip.

The departure delay is not easy to quantify because, as seen, it does not follow a normal distribution, and hence an average and a standard deviation are very misleading. However, squaring the individual delay, and then comparing the various legs on this basis, gives some measure of the variability. In this case, early departure and late departure are considered equally disruptive for the passenger. This is very true when connections are to be accounted for (see flight analysis). In this case there is a total of 638M square delay minutes. Only 5.7% of all the legs flying on only 1.3% of all the OD pairs account for 12.5% total delay square. Figures 6.10 and 6.11 show the general situation and the highlight of the 41 legs most disrupted in departure, respectively.

Departure Variability in Delay² All Stations July 2000

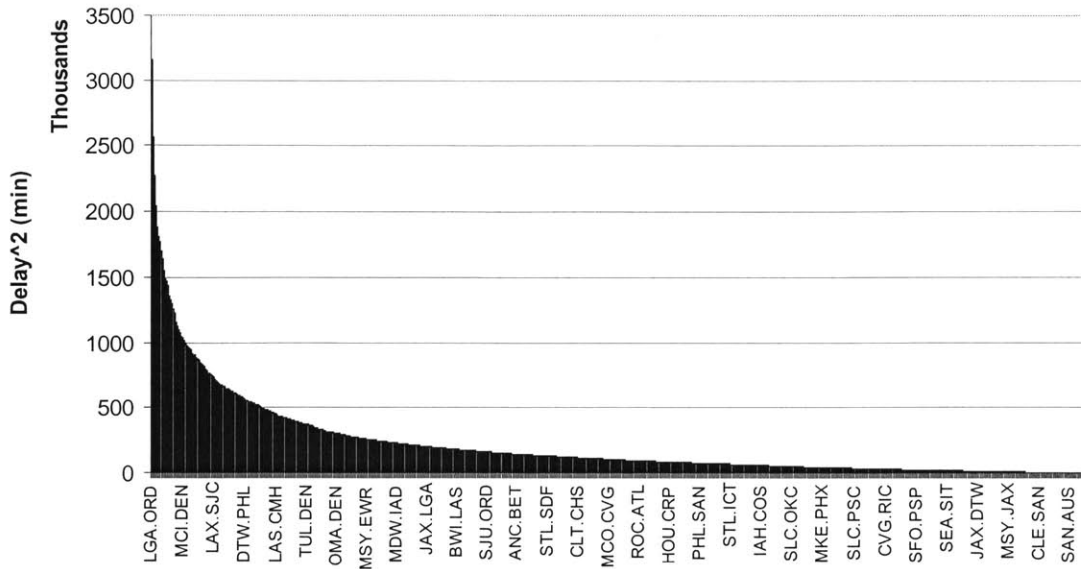


Figure 6.10 Leg Departure Variability. All Stations.

Departure Variability in Delay² July 2000

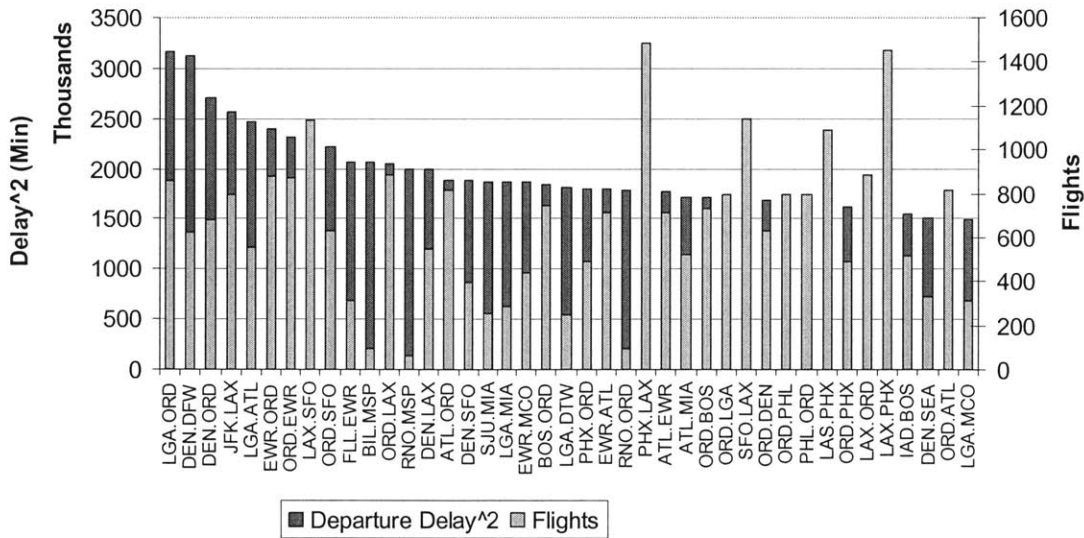


Figure 6.11 Leg Departure Variability. Most Delayed Legs.

The arrival delay is analogous to the departure delay, because the arrival delay is greatly affected by the departure delay. In this case, the total delay square is about 761M minutes. In this situation only 5.8% of the total legs flown (41 legs) accounts for

13.2% of the square delay. The OD pairs involved are only 1.3% of the total number of options. Figure 6.12 shows the general situation. Figure 6.13 shows the top 41 legs with higher variability. Notice that the 41 legs most variable in departure are not the same as those in arrival.

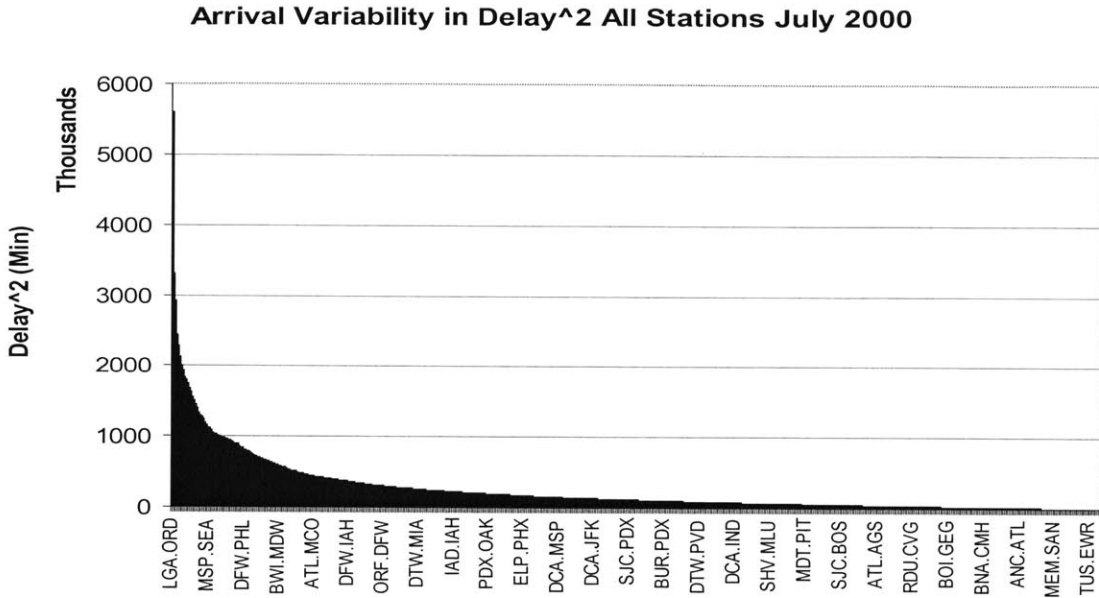


Figure 6.12 Leg Arrival Variability. All Legs

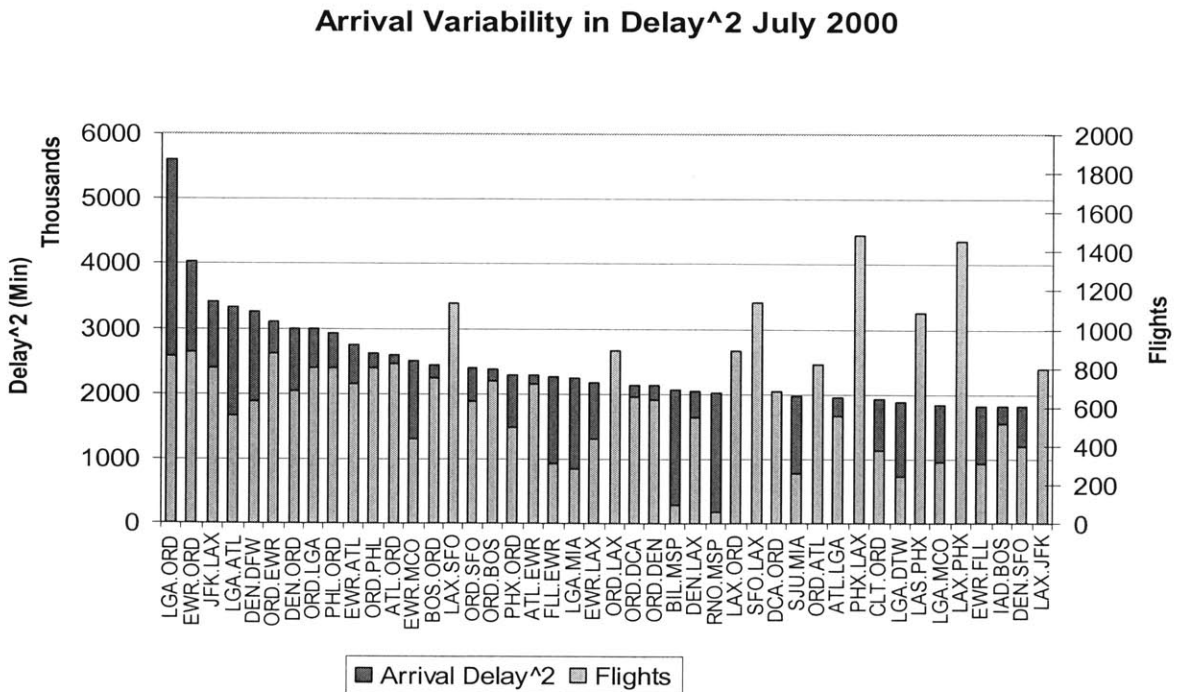


Figure 6.13 Leg Arrival Variability. Most Delayed Legs

The difference between the most delayed in arrival and departure is highlighted in Table 6.9 and Table 6.10. The one on the left shows the top 41 legs that have the highest departure delay square, with their position in the arrival delay performance. The other Table shows the 41 stations that, instead, have the highest arrival delay square, with their relative position on the departure performance.

Table 6.9. Departure Variability

Top Variability - Departure Highlight		
Leg	Departure	Arrival
LGA-ORD	1	1
DEN-DFW	2	5
DEN-ORD	3	7
JFK-LAX	4	
LGA-ATL	5	4
EW R-ORD	6	2
ORD-EW R	7	6
LAX-SFO	8	15
ORD-SFO	9	16
FLL-EW R	10	20
BIL-MSP	11	26
ORD-LAX	12	23
RNO-MSP	13	28
DEN-LAX	14	27
ATL-ORD	15	12
DEN-SFO	16	
SJU-MIA	17	32
LGA-MIA	18	21
EW R-MCO	19	13
BOS-ORD	20	14
LGA-DTW	21	37
PHX-ORD	22	18
EW R-ATL	23	
RNO-ORD	24	
PHX-LAX	25	35
ATL-EW R	26	19
ATL-MIA	27	
ORD-BOS	28	17
ORD-LGA	29	8
SFO-LAX	30	30
ORD-DEN	31	25
ORD-PHL	32	11
PHL-ORD	33	9
LAS-PHX	34	
ORD-PHX	35	
JFK-LAX	36	3
LAS-PHX	37	38
IAD-BOS	38	
DEN-SEA	39	
ORD-ATL	40	33
LGA-MCO	41	39
EW R-ATL		10
EW R-LAX		22
ORD-DCA		24
LAX-ORD		29
DCA-ORD		31
ATL-LGA		34
CLT-ORD		36
LAX-PHX		40
EW R-FLL		41

Table 6.10. Arrival Variability

Top Variability - Arrival Highlight		
Leg	Departure	Arrival
LGA-ORD	1	1
EW R-ORD	6	2
JFK-LAX	36	3
LGA-ATL	5	4
DEN-DFW	2	5
ORD-EW R	7	6
DEN-ORD	3	7
ORD-LGA	29	8
PHL-ORD	33	9
EW R-ATL		10
ORD-PHL	32	11
ATL-ORD	15	12
EW R-MCO	19	13
BOS-ORD	20	14
LAX-SFO	8	15
ORD-SFO	9	16
ORD-BOS	28	17
PHX-ORD	22	18
ATL-EW R	26	19
FLL-EW R	10	20
LGA-MIA	18	21
EW R-LAX		22
ORD-LAX	12	23
ORD-DCA		24
ORD-DEN	31	25
BIL-MSP	11	26
DEN-LAX	14	27
RNO-MSP	13	28
LAX-ORD		29
SFO-LAX	30	30
DCA-ORD		31
SJU-MIA	17	32
ORD-ATL	40	33
ATL-LGA		34
PHX-LAX	25	35
CLT-ORD		36
LGA-DTW	21	37
LAS-PHX	37	38
LGA-MCO	41	39
LAX-PHX		40
EW R-FLL		41
JFK-LAX	4	
DEN-SFO	16	
EW R-ATL	23	
RNO-ORD	24	
ATL-MIA	27	
LAS-PHX	34	
ORD-PHX	35	
IAD-BOS	38	
DEN-SEA	39	

There are various observations:

1. The leg that is experiencing the highest delay is in both arrival and departure LGA-ORD.
2. The legs that have top variability in departure are not the same as the ones that have top variability in arrival. In particular JFK-LAX is very late in departure (4th position) but not in arrival (not in the first 41). The opposite is true of the leg EWR-ATL: 10th position for arrival delay, and no presence in the departure delay.
3. The top variability in one direction is not the same as that in the other direction: for example DEN-DFW is 2nd in departure and 5th in arrival delay, but DFW-DEN does not appear in the top variability at all.

This highlights the capability of certain route and/or station to compensate or expand delays during their trip: this strengthens the initial assumption of a station-independent performance network.

Most Unreliable Legs July 2000 - Delay Variability

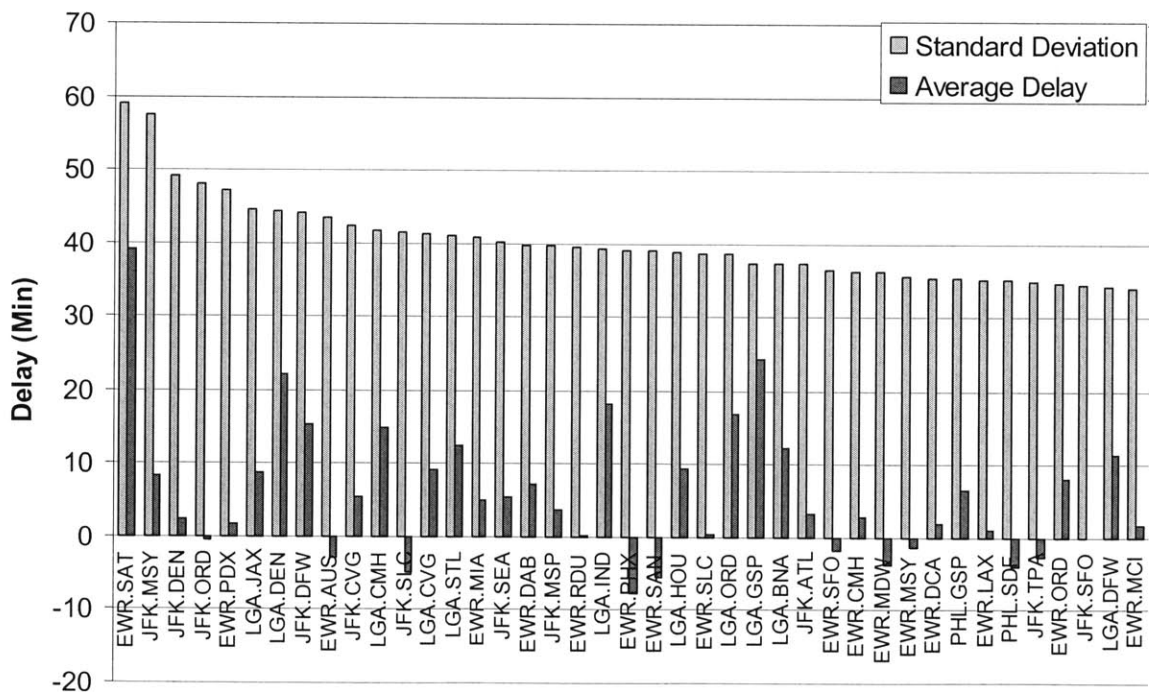


Figure 6.14 Most Unreliable Legs

The schedule delay is the delay experienced by the legs considering the legs shifted by the departure delay.

Most Delayed Legs July 2000 - Average Delay

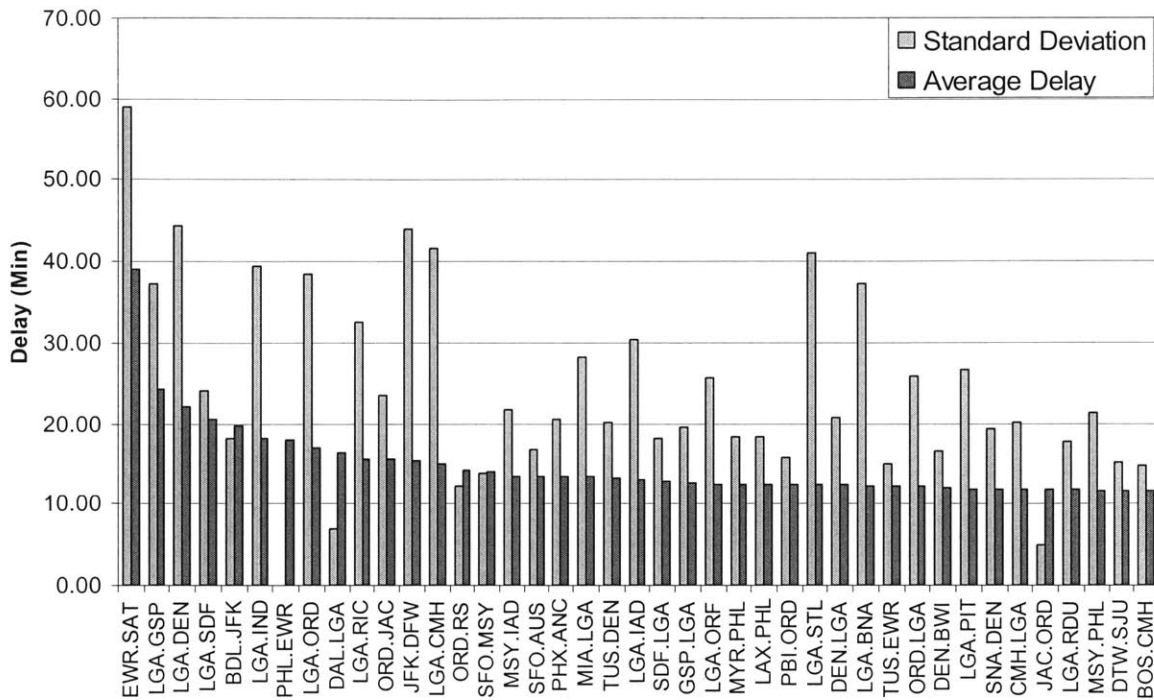


Figure 6.15 Most Delayed Legs

Figures 6.14 and 6.15 highlight the legs that have the highest average delay, and its variability. Notice that if the delay is high, but the variability is low, then the flight could be made more reliable by adding extra slack time. This is the easiest way to improve reliability of such legs as outlined by Carey (1999). On the other hand, if the variability is high but the average delay is low, then it is not possible to improve the reliability of such a flight. Figure 6.14 shows the most unreliable flights: the ones that have the highest delay variability. The high variability of these legs implies that it is difficult to improve their reliability and these legs may not be used to improve the schedule. Figure 6.15, instead, focuses on the most delayed. These legs are the ones that may need to be protected, or may be easily adjusted to improve the reliability of the schedule. Notice that in some cases the delay variability is higher than the delay itself.

Figure 6.16 shows the variation of airborne time of every leg. The flights with higher airborne variations are the ones that could be protected or not used for reliability purposes.

Leg Airborne Variability July 2000

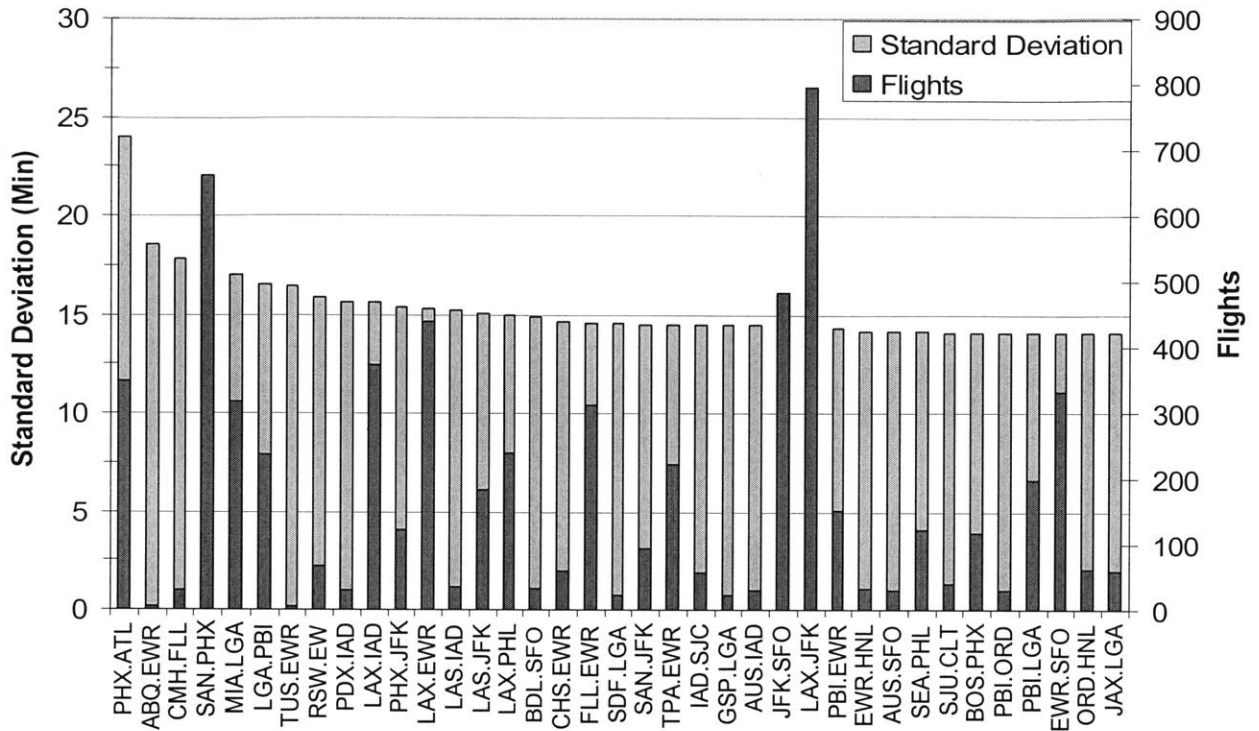


Figure 6.16 Leg Airborne Variability

6.8 Flight Analysis

The flight analysis results are more complicated because the ASQP does not contain any flight indications, but only flight legs. The flight number is not a unique way of identifying a flight: it appears that the same number is used multiple times and in different itineraries. However, it is possible to build up all the available itineraries using the flight legs by fixing the maximum number of stops, the minimum and maximum connecting times, and other restrictions such as fares and classes. Since flights are composed of flight legs, the reliability of the flight legs composing a flight determines the flight reliability. Hence, the reasoning applied before is valid for analyzing flights as well. The last leg of the flight determines its real reliability: if it arrives late, then the flight is late. It could be interesting to analyze the missed connection and the passenger diversion, but this cannot be determined using the ASQP database alone.

Overall, it is clear that the flights to protect are the ones that have the most unreliable legs composing them, with special attention to the last leg. On the other hand, the flights that can give better coverage are the ones that offer better reliability. If there is a need

for a flight that does not exist, then it can be built up following the *flight rules*, using the most reliable legs available.

6.9 Areas of Concentration

Analyzing the various legs, it is possible to identify the area of concentration of cancellations and delays. The origin and destination station defines a leg, and hence a cancelled leg determines the two stations that are affected by it. In the case of leg diversion, only the last station is disrupted. Plots of the cancellation concentration are shown in Figures 6.17 and 6.18. The data is normalized by the total traffic. The stations having less than 600 flights (average 20 flights a day) have been omitted. Because of the limited number of operations, any minor disruption can greatly affect them. In such a case, the inclusion of these “small” stations would make the chart quite unreliable.

6.9.1 Areas of Concentration: Cancellations

The most affected area is the east coast. The top unreliable location is the New York area. There are three major airports in New York: La Guardia (LGA), Newark (EWR) and JFK (JFK). There are also two minor airports: Westchester (HPN) and Islip (ISP). Their cancellation rates are 5.5%, 8.4%, and 2.1% respectively for the majors, and 17% and 0.4% respectively for the minors. The number of cancellations has been normalized by traffic; hence, even if this is a densely populated area (the most profitable markets are New York – Los Angeles and New York - Boston), it is fair to compare it with the rest of the US. Other areas include hubs, such as Atlanta (ATL), Chicago (ORD) and Washington DC (IAH), and important markets such as Boston.

These are the areas to protect and isolate. However, these are the areas to avoid when planning connections and protecting options for other flight legs, flights or stations.

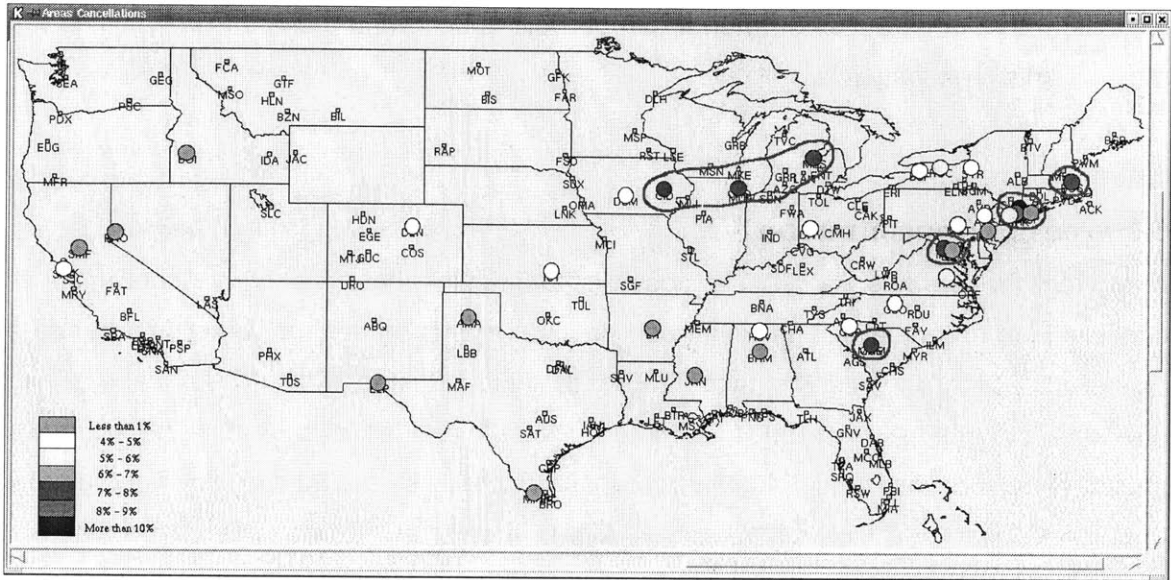


Figure 6.17 Areas of Concentration: Cancellations

6.9.2 Areas of Concentration: Delays

The same analysis as that for the station delays has been performed. In this case, the flight analysis includes the departure and arrival delays at the same time with equal weight. The Figure 6.18 highlights seven other areas. Even in this case the “low-tier” airports are excluded: these are the ones with less than 600 operations in the month of August.

The two most affected areas are the ones centered in Denver and New York. Notice that the areas showing the most cancelled and the most delayed flight ratios are not the same.

Cross-referencing with data for delay and cancellation highlights New York, Washington and Chicago as the most disrupted areas. Interestingly there are areas that are affected by both delay and cancellation (e.g., Chicago), but other areas are more affected by delay than cancellation (e.g., Denver).

The argumentation done for the cancellations areas can apply in this case as well: these are the sensitive locations to protect via the modifications highlighted in Chapter 5. Furthermore, if there is a need of new protecting options, these areas are to be avoided. Notice however, that they contain major hubs like Newark, Chicago, Denver, San Francisco and Washington DC. The implementation of alternative hubs, or large gateways, might be appropriate for lowering the load on these highly disrupted stations

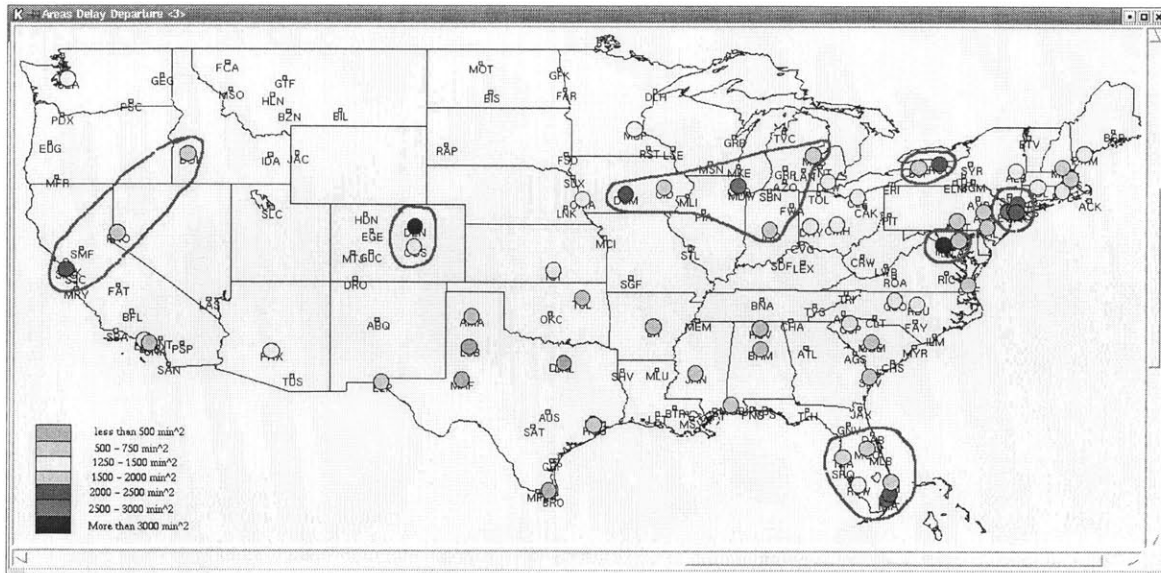


Figure 6.18 Area of Concentration: Delays

6.10 Rotation Analysis

It is not possible to do a rotation analysis for the cancelled legs, because it is not known which aircraft was associated with the cancelled legs. However, the rotations can be analyzed for the delays, which occurred on the legs composing them. It should be noticed that swaps, modifications and adjustments to the original rotations have already taken place. Another problem is that the maintenance stations are not known, and hence we cannot determine where a rotation starts and ends, and also how many rotations happened in the given period. We can assume, however, that the aircraft repeats the same rotation over and over again: this is a strong assumption but it allows comparisons of rotation delay.

A tool for visualizing the rotation was developed as shown in Figure 6.19. This tool has been very useful for highlighting the rotation distribution, the pattern of the movements on the network and the interchange of utilization between hubs.

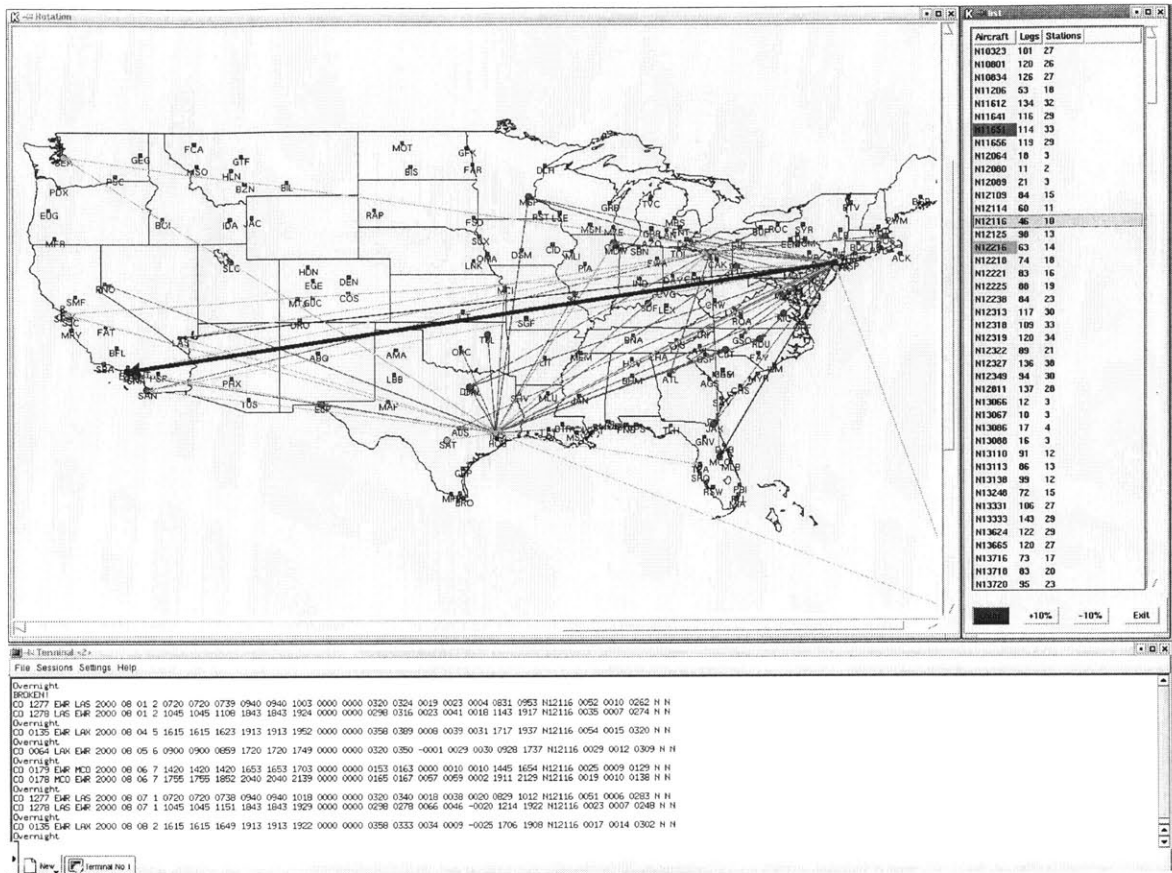


Figure 6.19 Rotation Visualization Tool

In particular the schedules of United, American Airline, Continental and SouthWest have been visualized. It is interesting to notice that for larger airlines (UA, AA) the hub and spoke system works loosely from the rotation perspective, and there is a high interchange between hubs. Part of a single aircraft rotation involves one hub for a limited period, after which another hub is used. This happened multiple times in a rotation during a month. It appears clear that the rotations tend to be spread across the network to have a uniform utilization of the aircrafts.

In the case of Continental (CO) the hub and spoke system is very tight. The same aircraft is used to serve the same hub: there is seldom hub exchange. This can be justified by the small size of the network, and the concentration of the major market in only a subpart of the nation.

The most interesting situation is that of SouthWest case. SouthWest, contrary of any other major airline, does not use a hub and spoke system. The same aircraft tends to take a long loop around the nation and rarely does it serve the same markets in a row. It is also noteworthy that SouthWest is the airline that has the best on-time performance.

6.10.1 Rotation: Delays

Summing up the delay accumulated into a single rotation and then dividing it by the number of legs composing it, highlights the rotations that are more prone to propagating delays through the network. As shown in Figure 6.20, there are only a limited number of rotations that are very unreliable with respect to their average leg delay.

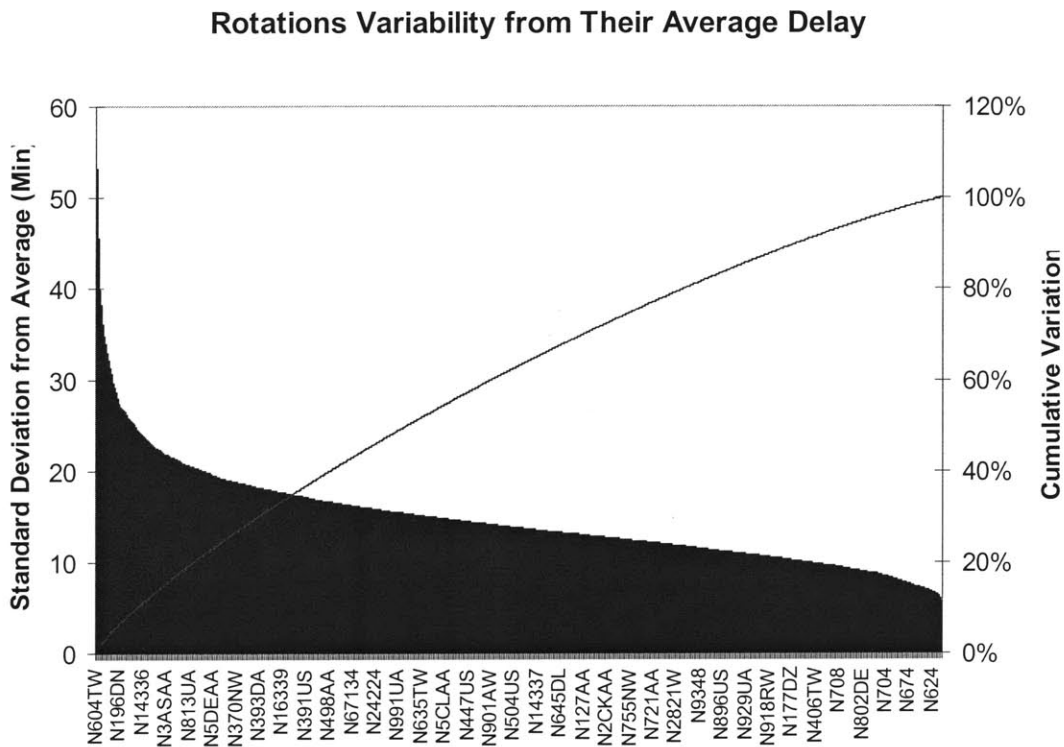


Figure 6.20 Rotation Delay Variability

If instead the rotation variability is considered as the delay squared collected during every leg flown, then it is possible to plot such variability in reference to the various airlines and see how the delay spreads across their network. Figure 6.21 highlights the rotation variability across airlines. Even in this case there is no airline that tries to absorb delay accumulated into the rotation across the network.

Figure 6.22, on the other hand, compares the rotation-cumulated delay squared per single aircraft type. Even here it is noticeable that the same aircraft type is not uniformly used by the same airline. Even comparisons between different airlines indicate a different spread of the delay across the rotation of the same aircraft type.

Rotation Variability in Delay² per Airline

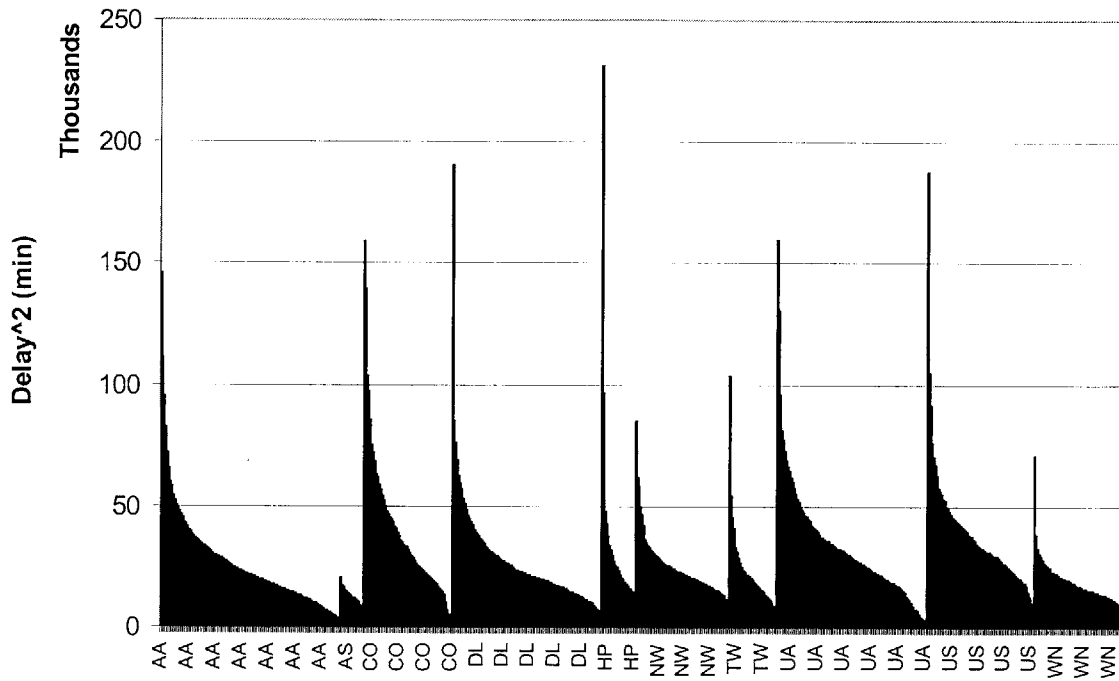


Figure 6.21 Rotation Variability by Carrier

Rotation Variability by Aircraft and Airline

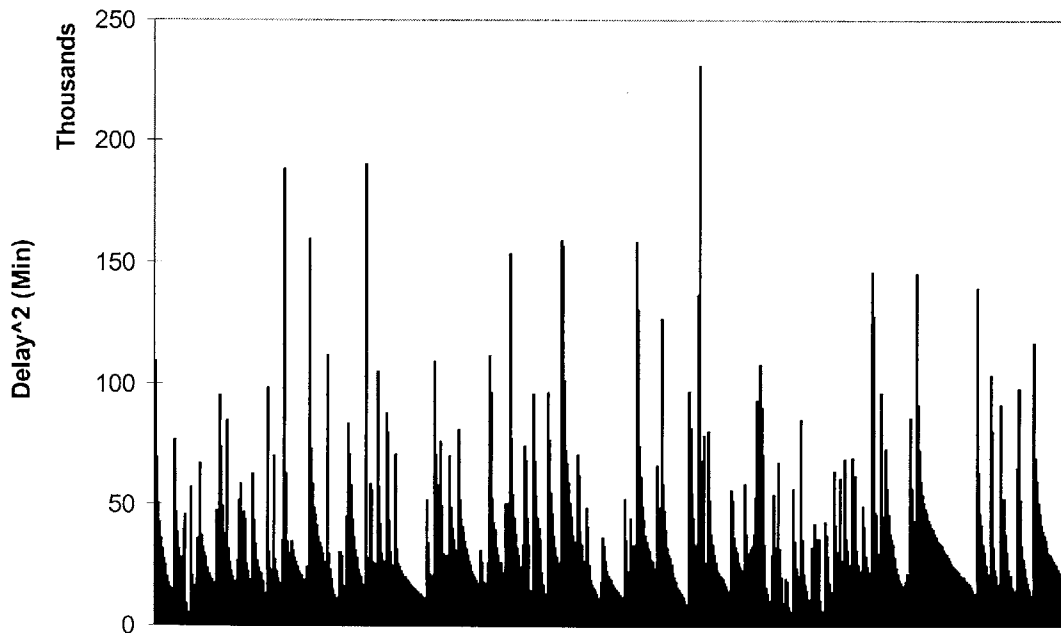


Figure 6.22 Rotation Variability by Equipment Type

A proper balance of the delay across the rotations means that the airline tends to have all the rotation with the same delay variability. On the other hand this imply that the delay (or the most delayed legs) is speeded across multiple rotations. The charts, therefore should not have the peaks highlighted in Figure 6.21 and 6.22.

6.11 Leg Load Analysis

Identification of the legs that have the highest load determines the true delays experienced by the flying public. If these legs have a high delay, then they will affect most of the people. The lower this value, the better the perception of the airline. The following consideration can only be discussed at the theoretical level, because information on load factors was not availability. However, the high load leg could have been determined by using the 10% ticket sample data (Form 41). This was not available at this writing. In the next chapter we will see real load data coming from Continental Airline for July 2000.

The best way of intervening is to isolate the legs with higher demand: this can be determined by analyzing the observed load factors or the demand model results obtained in the initial stage of the schedule planning process (*market identification problem*).

From recent studies on the Fleet Assignment Model, there is only a limited amount of legs that are capacity limited (around 30%). From data analysis the legs with highest load can also be identified. Having identified that, what is clear is:

1. These are the legs should be used as protection legs for other flights, due to their poor seat availability.
2. Reliable rotations should be put in place for them.
3. Extra care in timing and connectivity should be done.
4. Unreliable legs should not be connected with these legs.
5. The connecting time and the padding of these legs should be adjusted to make them more reliable.

The reasoning above will be explored in the model description, in the next chapter.

6.12 Conclusions

This chapter has outlined the various actions that may be performed on the schedule to introduce reliability. In detail, 6 adjustments for flight legs, 5 adjustments for flights and 4 adjustments for rotations have been outlined.

A heuristic analysis of the schedule used in July 2000 was performed to highlight the legs, flights and rotations to protect, and to determine the legs, flights and rotations that should not be used as protection options.

The result of this chapter will be used in the next chapter, where they will be implemented as reliability improvement actions.

Chapter 7

Model Interactions

The previous chapters outlined the details of the scheduling planning process, the reliability measures, and the actions to apply to the schedule. This chapter defines the flight legs, flights and rotations to manipulate to have a significant impact on the schedule.

The first part of this chapter delineates the RSM model, the interactions between its parts, and introduces the input and the output of the full model. In the end of this section, there is a set of techniques for validating the results generated by RSM.

The second part of this chapter highlights one of the actions described in Chapter 6: **Reduce/Increase Slack Time**. Based on the analysis done in Chapter 6, this is followed by an analysis of the schedule flown by Continental Airlines in July 2000, that identifies the legs with the greatest loads. These are the legs that will be adjusted for reliability based on the output of a tool for identifying the proper padding time for these legs.

At the end of the chapter there will be the analysis of the results obtained by this simple sub-model.

7.1 Robust Schedule Model Input, Output, Interactions

The main purpose of the RSM model is to adjust the network and the schedule given as input, with the purpose of improving reliability and reducing passenger disruption.

The full model is composed of two major parts: the **Critical Input Identification** and the **Schedule Modification**.

The process to identify the critical input, such as the critical legs and the critical stations, was outlined in Chapter 6. The results presented below are based on heuristic considerations and statistical analysis. Then considerations should be implemented by airline policies and strategies if available.

The analysis determines of the legs with highest loads for the schedule flown by Continental Airlines in July 2000. These legs are then adjusted with the **Reduce/Increase Slack Time** action. These legs are critical for the system because they are carrying the highest load, and hence most of the passengers. Adjusting the reliability of these legs will impact the passengers who utilize them; and hence, the flying public will have a much greater perception of the carrier is reliability improvements.

The center focus of this chapter is the **Schedule Modification** part of the model. The Schedule Modification is composed of four major problems:

1. Flight Leg Adjustment Problem
2. Multiple Flight Allocation Problem
3. Multiple Station Redundancy Problem
4. Station Isolation Problem

The interaction scheme of the parts was shown in Figure 2.2 and is repeated in Figure 7.1.

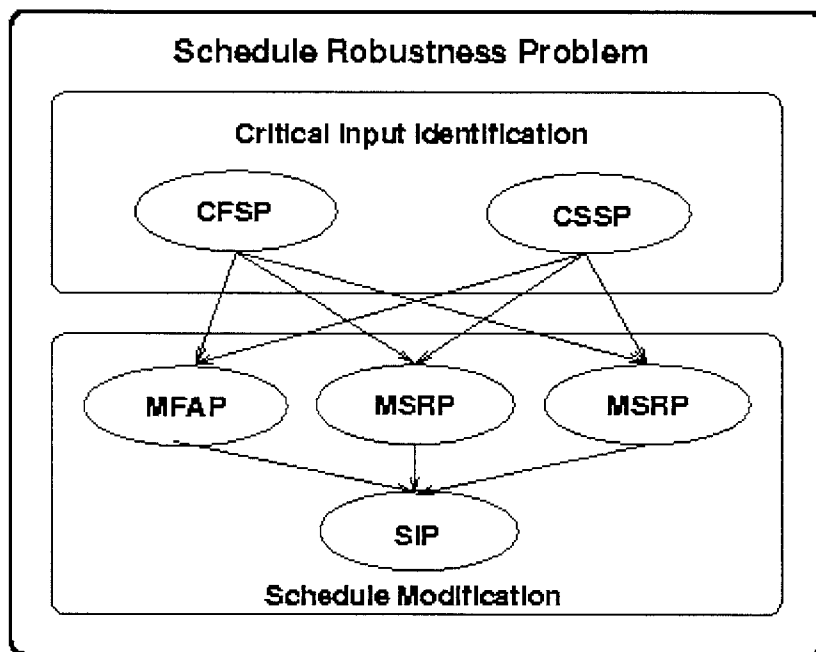


Figure 7.1 The Schedule Robustness Problem

The problems' interactions can be outlined also by functionality and scope. As outlined in Chapter 6, the functionality and scope of the Schedule Modification is:

1. Flight Leg Modifications
2. Flight Modifications

3. Station Modifications

These three blocks are the main part of the RSM: each part has its own mandatory inputs and its optional inputs. The output of every block is cascaded to the next one: each block gives its own improvements to the schedule received as an input, and then transfers the “updated” schedule to the next step.

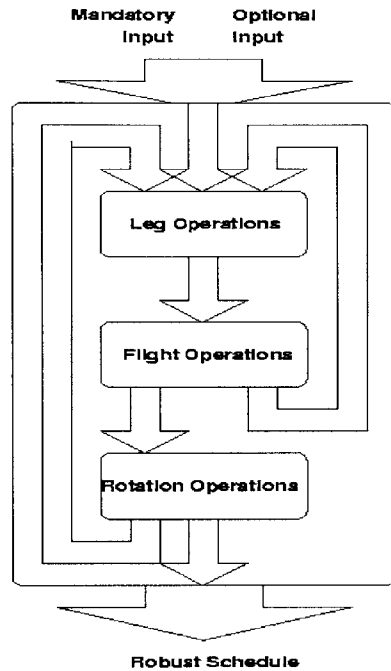


Figure 7.2 Interactions of the Core Parts

There may also be some reiteration of the various steps: the new schedule can be passed again to the previous steps for further improvement. As described before, every step is quite independent and focused on one specific aspect of the problem. It may be that applying the same steps on a slightly different schedule may cause the resulting schedule to be very different. This happens because the “optimization” generated by each step is local to the situation at hand. Given a different situation, the algorithm has different “directions” of improvement.

The interactions of the schedule modification parts are shown in Figure 7.2. As shown, there is a set of fundamental inputs and a set of optional inputs. The fundamental inputs are needed for solving the core problem. The optional inputs are used for giving more insight on how to solve the problem more efficiently.

The RSM fundamental inputs are the following:

1. **Original Schedule.** The schedule that needs to be readjusted. It has a period of validity and all the legs to put in place during such period.

2. **Ground Profile Forecast.** The station profile for gate-out, taxi-out, and taxi-in. The fundamental information is the probabilistic law regulating the profile and its parameters (*shape, location and scale*).
3. **Fly Time Profile.** The connecting profile between city pairs.
4. **Flight Legs to Protect.** Flight legs considered most important for the airline.
5. **Flights to Protect.** Flights considered most important for the airline.
6. **Station to Protect.** Stations considered most important for the airline.
7. **Areas to Protect.** The area that needs to be more reliable. This can be identified as a set of stations.
8. **Rotations.** The aircraft rotations that describe the aircraft cycle of utilization.
9. **Load Forecast.** Load for every flight leg needed for determining the leg to protect based on the passengers carried. This is used also for determining the seat availability.
10. **OD Demand Forecast Model.** The demand model used for generating the legs and the service to protect or the leg to avoid.
11. **Leg Capacity Assignment.** The pool of aircraft assigned to every leg. This is a by-product of the Fleet Assignment.
12. **Level of Protection.** The levels associated with the flights to protect in term of minimum coverage options, robustness options and maximum level of Option Value Ration accepted.

These inputs can be integrated with costs (fixed and variable) and revenues (average fares, class subdivision) that will determine the airline monetary benefit of the new schedule.

There are also other inputs that are limited to the subpart to optimize. In particular

1. **Minimum Turn Around Time.** This is needed for the rotation generation to ensure that aircraft are available for the next leg.
2. **Maximum Turn Around Time.** This is an optimal input and it is an indication of the longest time allowed for the preparation of an aircraft for the next flight.
3. **Flight Validity Rules.** These are the rules necessary to build up the flights' protections.
4. **Overnight Stations.** This is the location in which the aircraft can spend the night.
5. **Hubs and Major Gateways:** These are the stations where connections are permitted.

6. **Minimum and Maximum Connecting Time.** This is part of the rules associated with the flight creation. This time is the minimum and maximum time allowed for passengers to connect on a multi leg flight. The minimum connecting time is fixed by station.

Some input are shared across the blocks, and some others are peculiar to each block; some, on the other hand, are outputs of the previous block. As mentioned before, none of the blocks, or even any sub part of them, are mandatory. Each one of the blocks provides a very specific function and can be treated independently from the others

The problem itself defines which are the blocks that must be activated across the subproblems. The termination conditions, however, depend on the single block characteristics. More generally, the algorithm stops when:

1. No more improvements are possible in any block.
2. The improvements given by each block are satisfactory.

A combination of the conceptual view and the subproblem separation of the RSM is presented in Figure 7.3.

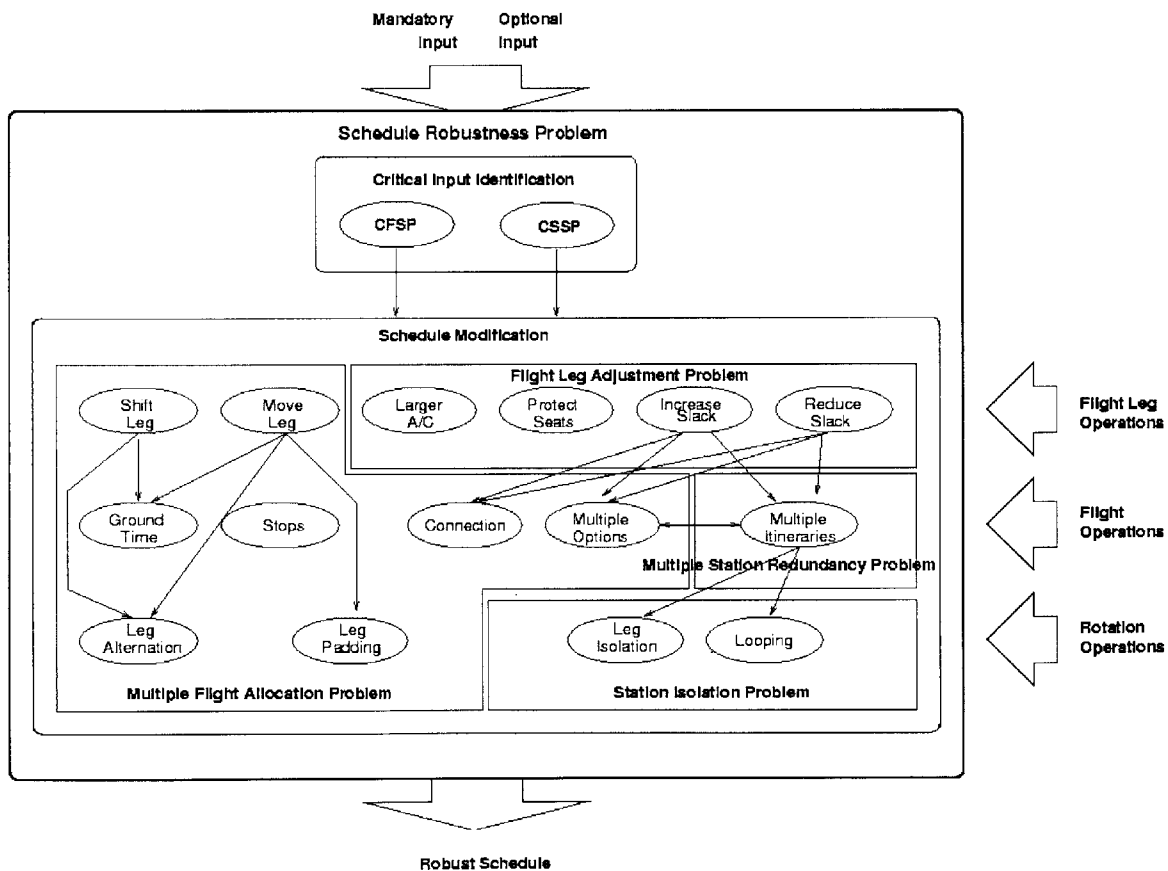


Figure 7.3 RSM with Highlighted Conceptual Division and Sub-Problems

7.2 Model Validation

A *validation model* is a model that objectively quantifies the goodness of a schedule. The measures that should be taking into account are the **Schedule Goodness Measurements** (SGMs) presented in Chapter 6. The validation can be done only on pre-existing schedules: this allows comparisons between an existing situation and a “forecasted optimized” situation. The only way to achieve this is by simulation: most of the models presented below are simulation models.

The main source of information for these validation models is the ASQP database. Schedules, itineraries, delays and cancellations can all be derived from this database divided by month. The load factor, the demand and the passenger capture, recapture and spill, instead, should be derived from other sources.

There are six validation models that can be put in place for the RSM:

1. **Static Allocation Model (SAM)**. It is based on the demand already allocated to the various flights e.g. load factor. The schedule is the real one and it has already been flown; modification, delays, cancellation and recovery have already happened. This gives, therefore, the exact value of the SGMs for the schedule considered.
2. **Network Simulation Model (NSM)**. This is obtained through simulation of every flight using the same demand found on the original schedule. What is simulated is the departure time, taxi-out, fly time, taxi-in time and the cancellation rate. These distributions and rates are the ones determined by the various profiles.
3. **Network Simulation Model with Perturbation (NSM-P)**. This model is analogous to the previous; however, two new *agents* are added to the simulation: **weather** and **traffic**. The weather condition is a random continuous event that increases or decreases in intensity in certain areas. The areas are determined a priori, but the intensity, the duration and the cycle is randomly generated. Traffic conditions (congestion) behave as the weather but it is limited to a day span and it has faster changes. Notice that these perturbations are applied to a schedule that has already been affected by the real weather and traffic conditions.
4. **Robust Network Simulation Model (RNSM)**. This is the same as the NSM but the schedule used is the one coming out of the RSM model. Notice that for the flights that are not present in the previous schedule, the demand needs to be determined by inference from the existing data.

5. **Dynamic Robust Network Simulation Model (DRNSM).** This uses the simulation as above but the decision for a passenger to take a flight, to be rerouted to another or not to travel is based on the disruption forecast for such a flight. This is limited by the network configuration, the flight options and the aircraft capacity available. There are no considerations about weather and traffic congestion. Notice that the distribution of the various flights is known a priori. This is defined by their profile.
6. **Dynamic Robust Network Simulation Model with Perturbation (DRNSM-P).** This is like the above but the two agents, weather and traffic, are introduced in the simulation.

The SAM is the baseline to compare the various schedules: there is no simulation involved, but a simple calculation of the various SGMs.

The first simulation model is the NSM: the simulation is done via keeping fixed the passengers allocated for every flight, and by reproducing the behaviors of the trips in a random way. The RNSM is exactly the same but the schedule used is the one generated by the RSM. DRNSM, instead, will use some *station agents* that will decide, flight by flight, how many passengers (proportion of the allocated load) to send to such a flight, and where to send the remaining ones.

7.3 Validation Models: Expectations

Given the validation models presented in the previous section, it is possible to forecast what should be the expected results on a regular and on a robust schedule. The characteristics of the output of the simulations are:

1. **Cancellation.** All the models should generate approximately the same number of flights cancelled. The model handling perturbation should have a higher number of flights cancelled than the others due to the introduction of weather and traffic.
2. **Total passengers' delay.** The robust schedule model is focused on creating a schedule that limits the passengers' delay, and therefore, in the following analysis, it is natural to expect that the new schedule is, overall, better than the existing one. In detail here is the relationship of the various SGMs obtained from the validation models proposed.

- a. **Arrived Delayed:**

$$\text{SAM} \sim \text{NSM} < \text{NSM-P}$$

> RNSM > DRNSM ~ DRNSM-P

Both the static and the basic simulation will return approximately the same number of passengers arrived delayed. The model with perturbations will increase this quantity. The RNSM will improve the situation and will reduce the number of passengers arriving late; however, it is not clear how it will compare with the original network simulation with perturbations. In general the schedule with perturbations will behave worse than the one without perturbations. It is definitely possible to assume that

NSM-P >> DRNSM-P

b. **Allocated to Other Flights:**

0 = SAM < NSM < NSM-P

~ RNSM << DRNSM ~ DRNSM-P

The simple simulation should give approximately the same results as the static allocation, even if the robust schedule should not relocate more than the normal schedule. This happens because the robust schedule has already allocated part of the traffic across the protecting options.

The robust schedule in the dynamic condition tends to allocate more passengers to other flights. This reduces the disruption that passengers would experience if they remain in the assigned flight. Even in this case it is not clear how to compare the perturbed simulation with the not perturbed one. It is instead clear the relationship between them is

NSM-P << DRNSM-P

c. **Spilled to Another Airline:**

SAM ~ NSM < NSM-P

> RNSM >> DRNSM ~ DRNSM-P

In this case the number of spilled passengers tends to increase with the original schedule. The utilization of the robust schedule, instead, tends to reduce this amount.

d. **Lost:**

SAM ~ NSM < NSM-P > RNSM >> DRNSM ~ DRNSM-P

Another strength of the robust schedule is that it limits the number of passengers who will not reach their destination. In this case, then, it is natural to see some increasing benefits in using it. It is possible to safely

compare the perturbed schedule with the non-perturbed one because the intrinsic benefit of the robust schedule should outweigh the absolute differences between NSM and NSM-P.

3. **Total passenger-minutes delay:**

$$\text{SAM} \sim \text{NSM} < \text{NSM-P} < \text{RNSM} < \text{DRNSM} \sim \text{DRNSM-P}$$

The passenger's delay is another stronghold of the robust schedule; therefore there will be an improvement in this area as well.

4. **Total passengers' disruption delay:**

$$\text{SAM} \sim \text{NSM} < \text{NSM-P} > \text{RNSM} \gg \text{DRNSM} \sim \text{DRNSM-P}$$

This quantity is basically a reliability indication. Moving passengers to less disrupted options tends to positively reduce the number of passengers experiencing delays.

5. **Total disruption value:**

$$\text{SAM} \sim \text{NSM} \sim \text{NSM-P} > \text{RNSM} \gg \text{DRNSM} \sim \text{DRNSM-P}$$

This value will be reduced as well, as an intrinsic benefit of the schedule.

Notice that most of the time it is expected that the dynamic perturbed robust schedule and the non-perturbed model will return similar values: this is because the robust schedule is designed to absorb the variability in the system as much as possible.

In theory we can determine a *factor of improvements* for every SGM. This factor can be calculated as a ratio with the value returned by the SAM.

The outcome of these validation processes is that if there is an improvement of the SGMs (a greater than one improvement factor), the new schedule is better than the existing one. This means that the original schedule has already been flown and all the optimizations, recovery procedures, and recommendations have already happened; therefore, the proposed one can be considered even *better* than an adjusted one. Hence, if the model works, the benefit observed will be only a lower bound of the possible benefits achievable.

7.4 Increase/Decrease Flight Slack Time Problem

The following model illustrates the effectiveness of the actions highlighted in the previous section. The objective of the **Increase/Decrease Flight Slack Time Problem** is to determine the proper flight slack time for the flight leg given as input so that only a

limited number of flights are late. The definition of “late” will follow the standard recognized by the industry that is “later than 15 minutes from the scheduled arrival time.”

The calculation of the proper slack time is computed for each of the legs in input. The formula used will be described later on. The schedule used is the one flown by Continental in July 2000.

The total legs scheduled by Continental Airline in the domestic US markets by jet aircraft are 32,798. The total independent leg flown divided on a week period is 4,616.

Given this data, the process followed for analyzing it is:

1. Isolate the legs with higher load.
2. Generate the optimal slack time.
3. Define the level of OV required.
4. Compute the Option Value for the two schedules

7.4.1 High Load Legs

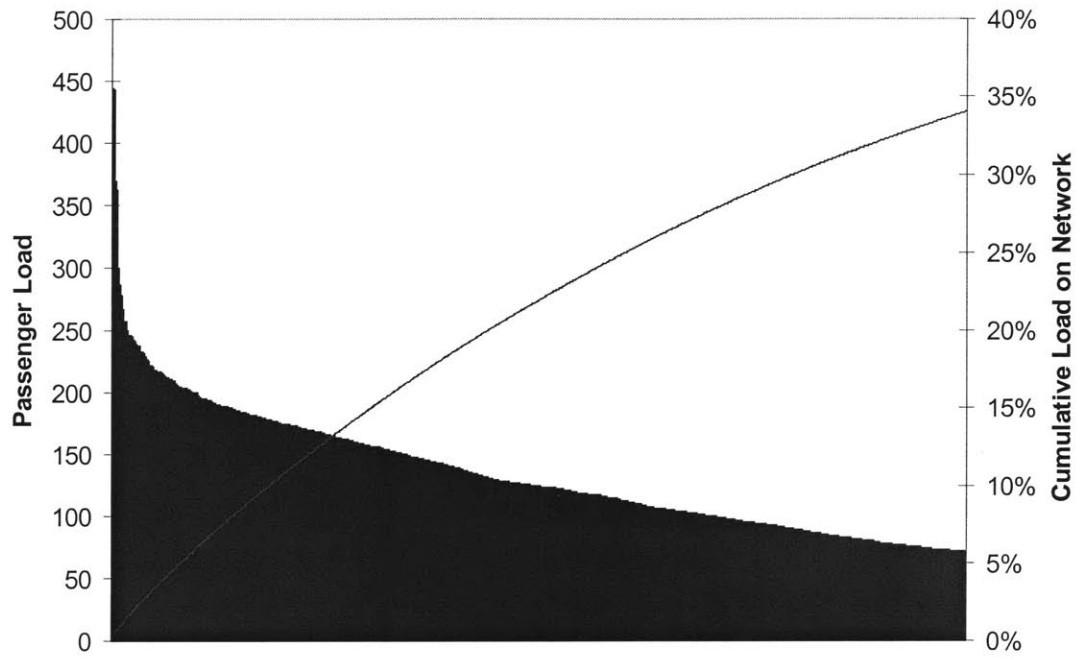
An analysis tool has been developed isolate the legs with highest load. The load has been computed for each of them divided by day of the week. The Figure 7.4 shows the situation: 20% of the total legs flown in July 2000 are responsible for 60% of the load.

The legs represented in Figure 7.5 are all the leg flown by jet and turbo prop aircrafts domestically and internationally.

The ASQP database contains only the jet aircraft for the US domestic markets; therefore, some of the legs presented cannot be manipulated by the profile analysis.

Having eliminated the international legs and the legs flown by turbo-prop aircraft, the 1000 legs with the higher load have been selected. As shown in Figure 7.5, these legs account for 35% of the network load occurring in the considered period.

Load On 1000 Most Used Legs



7.4 Legs Load on Continental Airline Network July 2000

Leg Loads

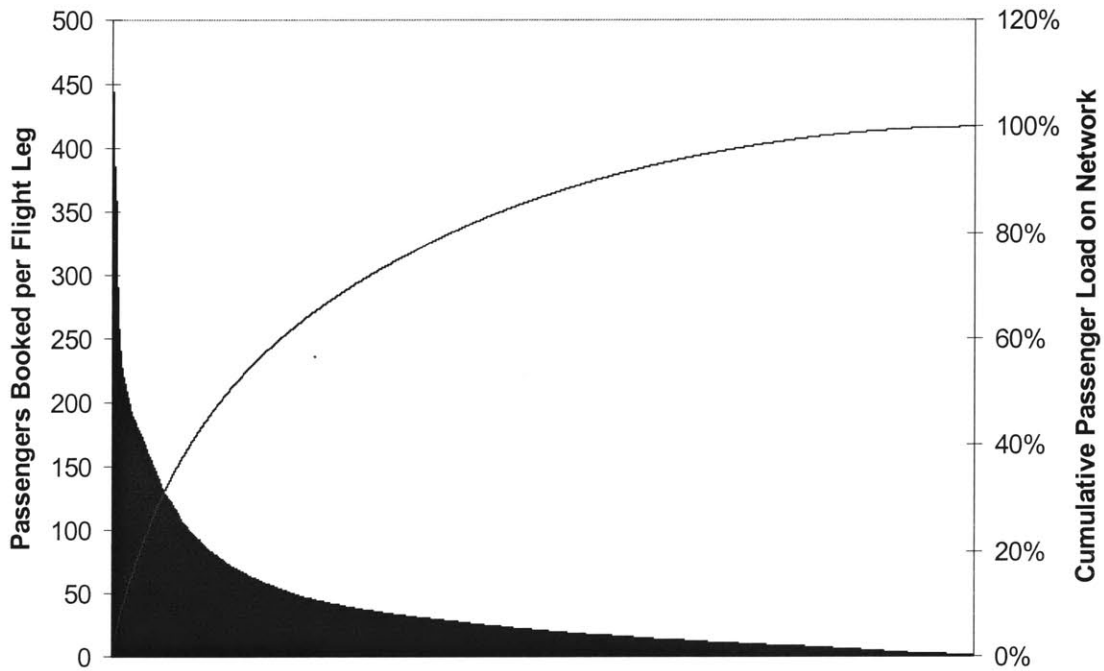


Figure 7.5 Load Distributions on the 1000 Flight Legs with Higher Load.

7.4.2 Profile Analysis

The calculation of the Optimal Slack is obtained using the time profiles generated for the US domestic market during July 2000. The formula used for the OV is the following:

$$OV_{leg}(fl) = (w_{gt}\bar{t}_{gt} + w_{to}\bar{t}_{to} + w_{tr}\bar{t}_{tr} + w_{ti}\bar{t}_{ti} + w_s \max(|t_s - t'_s| - 15, 0)) e^{\frac{p_c}{1-p_c}} e^{\frac{l^3}{1-l^3}} \quad (7.1)$$

In this case a more relaxed situation is used in which a flight is considered late only if it is more than 15 minutes later than the schedule arrival time.

The profile aggregation is obtained through the normality assumption (the trip time is following a normal distribution). Considering the 15 minutes of leeway, the difference between the profile convolution and the normality assumption is negligible.

The weights for the terms in the formula are the ones outlined in Chapter 5. In particular:

- $w_{gt} = 1.5$ The weight for the gate time. Considered as waiting on the platform.
- $w_{to} = 1.25$. The weight for taxi-out time considered as beginning of the trip.
- $w_{tr} = 1$ The weight for the trip time. This is the basic reference.
- $w_{ti} = 1.25$. The weight for taxi-in time. As in the taxi-out case, this is considered the ending part of the trip.
- $w_s = 5.6$. The weight for the slack time. This value implies that only 82% of the flights will arrive before their schedule arrival time plus 15 minutes.

No information about load factor is available; hence, it is assumed that there are always seats available for accommodating passengers ($l=0$).

The calculation of the slack time is obtained from the formula:

$$t'_s = \sigma = \sqrt{Var[t_{gt}] + Var[t_{to}] + Var[t_{tr}] + Var[t_{ti}]} \quad (7.2)$$

The optimum slack time will be the obtained by:

$$t_s^o = \max(|t_s - t'_s| - 15, 0) \quad (7.3)$$

7.4.3 Slack Time Calculation

The calculation of the slack time was performed and the results are shown in Figure 7.6. The results indicate that more than 50% of the legs considered do not need any variation in their Slack Time. Their performances are in line with the expected one.

Flight Slack Time Distribution

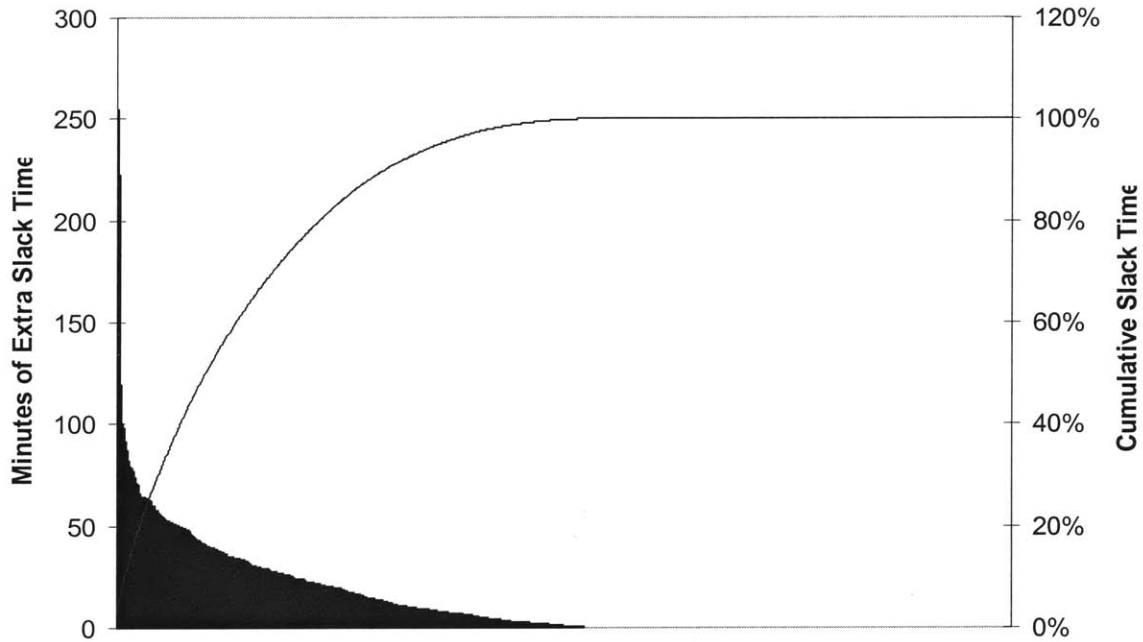


Figure 7.6 Slack Time Distributions Across the 1000 Most Loaded Legs

The reliability of 30% of the remaining legs may be improved through a slack time variation limited to no more than 50 minutes.

The remaining 20% of the legs analyzed require more attention. The resulting slack time is particularly high: they would require more than 100 minutes of slack time on a scheduled gate-to-gate time of 100 minutes (increase of total trip time of 100%). It is noteworthy that the extra time is coming from a high delay at the gate.

The top 10 legs with highest slack time variation, in fact, involve the stations of Boston (BOS), Cleveland (CLE), Newark (EWR) and Houston (IAH) on flights departing between 5pm and 9pm, mostly on Sundays. This is a typical evening peak time.

The average gate time is around 30 minutes with a variability that ranges from 65 minutes to over 200 minutes. Furthermore, it appears that all the legs are under-scheduled by a factor of 20 to 40 minutes.

Notice that the profiles have been tested with the Kolmogorov-Smirnoff Test: the hypothesis of lognormal distribution for gate, taxi-out and taxi-in time, and normal distribution for trip time cannot be rejected with a confidence level of 1% in all the cases.

7.4.4 Option Value

The total **Option Value** calculated across the 1000 legs is around 252,000 utils. The total **Ideal Option Value** is 157,000. This implies that the 1000 legs inflict an extra 60% on average of disutility on the flying public. There are only a few legs that do not impose any extra disutility. Sixty percent of the legs give less than 40% disutility. The top most disruptive legs have a disutility ranging from 500% to 800%. These legs, when the slack time is adjusted, will lower their Option Value Ratio (OV-R) below 30%.

These data indicate that any passenger flying on these legs is going to have their trip disrupted, and the impact in terms of utility value is an extra 60%.

The factor that most affects the option value, after the incorrect slack time is the cancellation rate. The cancellation rate in the top ten legs varies from 2% to 20%. These, however, are not the legs that are most affected by cancellations. In fact, some of the remaining legs have a cancellation rate as high as 45%: it appears that they have lower OV because they have a more reasonable planned slack time.

On average, if the slack time is adjusted properly the total utility is lowered to 169,000 units. In general, the extra disutility is limited to an increased 8% on average across the legs. Nevertheless, some legs still have a high disruption burden: the maximum Option Value Ratio (OV-R) is 116%.

It could be better to limit the OV-R to a rate not higher than 20%. Even in this case, however, some legs do not match the best level. Therefore, other means for improving the schedule need to be found.

7.5 Conclusions

The interaction of the two parts of the RSM model has been presented. In addition, the details of the core parts have been introduced, and a definition of the input manipulated and output generated has been outlined.

A formal way to compare different schedules has also outlined for use in the validation of the model.

One of the methods for improving the schedule has been applied on the Continental schedule flown in July 2000.

The results of the **Increase/Decrease Flight Slack Time Model** indicate that some legs can be improved in reliability through modest modifications of the slack time. These

improvements can be done on 40% of the legs. The variation of the total trip time can be limited to less than 30 minutes.

In some cases, this is not enough: the sub-part of the model presented does not solve the whole RSM. Nevertheless, its impact is significant.

Chapter 8

Conclusions

The previous chapter outlined one of the actions to apply to the schedule of Continental Airlines. The model has been run and the results are encouraging, but further analysis is required.

This chapter contains a summary of the hypothesis, methodology and findings obtained during this research. After that, a list of drawbacks and suggested improvements to the model presented. Finally, there is a discussion of future directions for further improvements and research.

8.1 Summary

This paper was structured into 3 major sections:

1. The **Schedule Design Process**.
2. The idea of **time profile** and profile analysis.
3. The **reliability measures** and how to put reliability into the schedule.

Chapter 1 introduced the schedule planning process. The process itself was initially conceptualized and presented in its canonical form. This is the process followed by all the airlines considering an “ideal scenario” in which the environment where the airline operates is not influenced by any external events. However, many disruptions occur and the schedule used is far different from the one desired. The major drawback of the initial process is that there are no opportunities to incorporate any new objectives or constraints into the schedule, particularly reliability. The process is fairly simple and linear; however, flexibility and adaptability are missing.

Chapter 2 outlined the core concepts of **Robustness** and **Coverage** for a schedule. These concepts were based on analysis of the airline recovery procedures. These procedures are effective but they only take place after the disruption has occurred. It was clear, therefore, that a characterization of an *a priori property* of the schedule was

needed. This property, **reliability**, marginally affects the recovery procedures: its central focus is reducing the impact of any disruption *before* the system reaches a breakdown situation and there is consequent need for a recovery action. A new process was then described. Terms and concepts for expressing different schedule planning processes were presented and used to introduce the reliability characteristics.

Chapter 3 presented the core idea of the profile analysis. The network was seen as a source of disruptions: a flight was then fully characterized by its departure performance and arrival performance. When these were fully described by statistical laws, then inference was used to determine the expected trip time, delays and cancellation chances.

Chapter 4 described formalism for the parts used in the definition of schedule reliability measures presented in Chapter 5.

In Chapter 5 reliability was initially qualitatively described; furthermore, a description of “*goodness*” property of an effective reliability measures are presented. Based on the information available and the profile analysis, a final quantitative measure of reliability was introduced: the **Option Value**.

This value was expressed for a single leg, a connection and a flight. The Option Value was not enough for comparing different flights; the **Ideal Option Value (I-OV)**, and the **Option Value Ratio (OV-R)** were consequently defined. These terms are very useful for identifying the disutility imposed on passengers due to a poor flight schedule. By comparing I-OVs and OV-Rs, it was possible to determine the burden that a hypothetical passenger experiences when flying a certain flight leg or flight. It is actually possible to define such burden for every leg and every stop of the passenger’s trip. The burden, or **disutility**, is defined in terms of difference between the expected and the statistical inferred trip performance. The difference in utility values determines the level of disutility accepted during the schedule design and the reliability incorporation: it is possible, therefore to limit the disutility to certain levels (**reliability levels**) and then plan the schedule accordingly. This information is needed for evaluating a static schedule; however, for practical purposes, another way of comparing options when rerouting passenger is needed. This is achieved with the **Option Disruption Value**. This is the value of disutility experience by a traveler when he is rerouted from one flight to another.

Chapter 6 introduced 15 suggested ways to improve the schedule. An analysis aimed at identifying of the parts of the schedule that need to be improved was also

presented. With the same analysis it was possible also to identify the parts of the schedule that cannot be used to ensure reliability.

Finally, in Chapter 7 one of the methods described in Chapter 6 was applied to the schedule flown by Continental Airlines in July 2000. The method implemented was the **Reduce/Increase Flight Slack Time (R/IFST)**. This method gave encouraging results but it is clear that it is not enough to ensure full coverage, robustness and reliability of the schedule that was processed.

8.2 Model Results

The model presented in Chapter 7 showed that it is possible to improve flight reliability by changing the Flight Slack Time. The model works fairly well in 70% of the cases.

In the remaining 30%, the extra slack time is higher than would be acceptable for any airline: the flight slack time would exceed 100% of the actual schedule time. The time profile analysis indicates, however, that correct accounting for variability pushes this limit even farther: the highest variation obtained is on a Cleveland-Boston flight, departing at 9:30pm on Sundays. This flight is scheduled to take 103 minutes, but it is under-scheduled by 44 minutes. This means that the summation of the average part of this flight is about 147 minutes. The introduction of the slack time due to variability in the profiles increases the trip time to 224 minutes. This value is more than 2 times the scheduled trip time!

It is clear that other means of improvement should be used. Integration of the R/IFST step in a much larger scheme should limit the impact of the slack time on the total reliability of the schedule. The R/IFST was activated only to increase reliability for the flight leg: no actions were taken to increase coverage or robustness and no consideration on flight and rotation were operated. This operation does not involve any modification of the demand forecast, the equipment used, or the crew flying the plane. There are, however, some restrictions that should be checked. In particular:

1. **Slot.** The flight is not arriving at the current arrival time: if a slot is allocated to the flight then it has to be shifted by the new slack time. This shift sometimes is not available, especially if it is late at night.
2. **Crew.** The crew does not change; hence, there is no need to find a new crew to fly the new modified flight. However, no crew can fly more than a fixed

amount of time. Furthermore, after a certain period any crew has a mandatory rest period. The crew flying time is considered from gate to gate; hence the introduction of a larger slack time implies a longer *crew flying time* as well. This implies that the crew scheduling problem need to be resolved for the new schedule. The introduction of extra time in the flight imposes costs and constraints on the crew flying the aircraft, which need to be considered.

3. **Rotation.** As described for the crew, an aircraft cannot fly more than a fixed number of hours. The introduction of extra slack time can push the limit of the maximum flying time of the specific plane, especially if the aircraft rotation is very tight.

The previous model has other drawbacks: crew connection, passenger connection and aircraft connections. The connectivity concerns arise because the variation of the trip time impacts the system in which the leg is placed. The increase in Flight Slack Time increases the on-time performance of every leg, but it is not guaranteed that the connections are going to be preserved. It can be assumed that a modification of less than 30 minutes in increased leg time is not going to impact too many connections: the tightest connections should be re-planned. However this can be considered the limit that distinguishes the applicability of the method.

Another problem is related to market competitiveness: the modification of the flight schedule time can put the airline at a disadvantage during the passenger trip planning process. Passengers are very risk-averse when planning a trip; however, they tend to select the shortest travel time when planning an itinerary. Passengers want to dedicate as much time as possible to the scheduled activity, and hence they want to spend as little time as possible on the travel portion of the trip.

These are, therefore, two conflicting objectives: being reliable and being fast. The main benefit of the extra slack time is increased reliability; however, the trip time is also increased. How much slack time is acceptable to introduce in the schedule for having the trip reliable enough but not too long such as passenger do not opt for other competing options and modes. This consideration needs more investigation: a deeper understanding of the trade-offs that passengers make during their decision-making process needs to be assessed.

Another important aspect that is missing in the previous modeling process is a validation of the results. The validation could have been done, as reported in Chapter 6,

via a static method and via a simulation model. Simulations, even if time consuming and input-intensive, can give great insight into the problem.

The static analysis of the new schedules indicates that, based on the **Schedule Quality Measurements**, the new schedule is much better than the original one. The quantification of this improvement is reported at the end of Chapter 6.

These are strong indications that the model is effective; however, a deeper validation through simulation could have further corroborated the work done so far.

8.3 Additional Research

The purpose of this paper was to show the possibility of putting in place simple actions targeted to specific aspects of the revised schedule planning process that are centered on introducing reliability into the schedule. The reliability has been summarized into a model: the **Robust Schedule Model**. This model can be integrated effectively into the original schedule design process.

There are certain aspects that need more investigation. The first is the implementation of the other parts of the model. In particular, the **Multiple Itineraries Problem** is the most promising for increasing the robustness of the system.

Furthermore, there is a validation model to implement. It could be beneficial to create the simulation model for the new schedule. Usually the creation of such tools gives great insight into the problem of interest.

In particular the dynamic simulation model can show the benefits of multiple protecting options: the selection of the option with which to reroute the passenger traffic based on weather and congestion information can provide a full understanding of the potential of a robust schedule. This is the kind of process that has been outlined in Chapter 4. The RSP generates a more reliable schedule; however, the multiple opportunities for keeping the system flexible and resistant to disruptions need to be used. Some reliability modifications are already in place and are already fully exploited, such as the flight slack time; others, like the protecting options, robustness and coverage, need to be actively exploited.

The exploitation is done through the **dispatcher**: this central authority and point of interaction is the one that is receiving all the information about the traffic, weather and equipment functionality. Based on such information he/she has to decide which

passengers and how many need to be rerouted to the available options. Proper criteria, methodology and algorithms should be put in place to generate the best *solution* to the passenger flow rerouting. This process is presented in Figure 8.1.

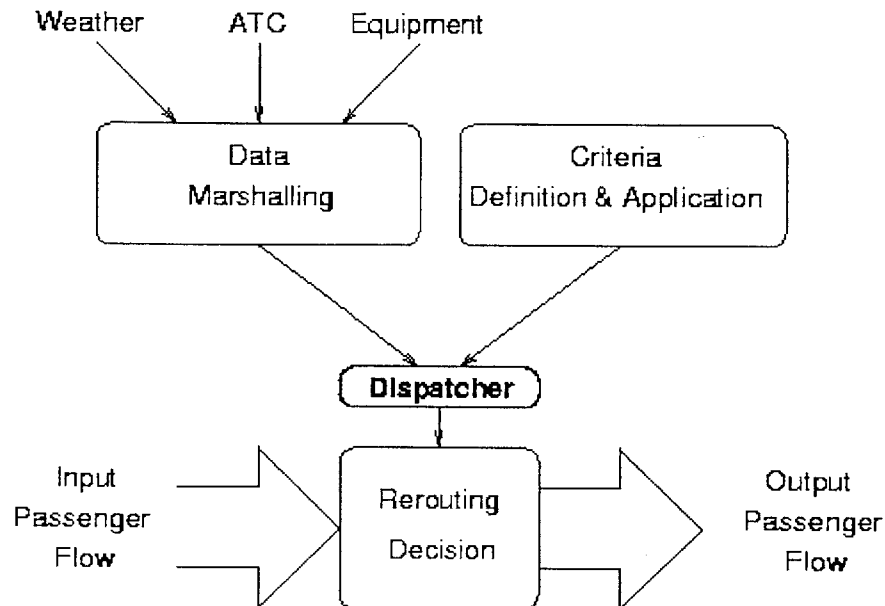


Figure 8.1 Dispatcher Interactions and Rerouting Activity

There are, however, other parts of the schedule design process and of the airline operations in general, that can give better improvement to the current approach, and in particular:

1. The rules used for generating flights. How to automate the selection and construction of more reliable flights.
2. Who to divert. How to select the portion of passenger flow to move to the options available. "How to" gives more freedom to the dispatcher but at the same time can create inconvenience to the flying public. There are *fairness* issues involved that need to be clarified. However, the prospect of being less disrupted during the trip can be used as an *extra* benefit for a different class of passengers.
3. How to automate the selection of the flight to protect. The analysis of the network operations and the selection of the critical flights and stations need to be addressed. The approach used here is a statistical analysis. This kind of proceeding is resource intensive and needs to use no quantifiable property

of the airline business such as strategy and competitiveness. A deeper understanding of these interactions with the selection of the best market to protect is required.

4. Interaction of the robust schedule with the recovery operations that the airline has in place. The utilization of the robust schedule implies that the portion of passengers affected by delay and cancellation is reduced; however, the problem of the irregular operation has not been eliminated. After all, only passenger flow is rerouted: no modification to aircraft or flight routing has been decided. The different cost/benefit trade-offs when comparing a robust and a non-robust schedule from the recovery perspective should be interesting to analyze.

All the issues above are left to future research.

8.4 Conclusions

The dissertation focused on the core part of the airline business: the schedule planning process. While integration of all the existing stages in airline schedule development may produce a global optimal solution to the schedule design process, the computational burden and the complexity generated has been proven to be too difficult. Furthermore, introduction of additional objectives such as robustness is also difficult.

The approach presented here, instead, subdivides the schedule design problem into smaller and more tractable sub-problems; introduces other properties such as robustness; formulates additional sub problems; and then optimally integrates them a new Schedule Planning Process.

We have shown that the schedule robustness problem can generate beneficial results for the schedule and deliver a better experience to the flying public. Further research is needed to exploit other aspects of the problem and generate even more *building blocks* with a much more limited focus but with greater improvement potential.

Areas with the potential for most improvement have been highlighted. Other areas can be identified during the development of the subparts that were presented. The current approach, however, is far from optimal but gives good insight into the Schedule Planning Process and can lead to new and exciting opportunities for further research.

Bibliography

Antes, J. [1997]. Structuring the process of airline scheduling. *WINFORS Working Paper*, University of Cologne, Germany. URL: <http://www.informatik.uni-koeln.de/winfors/pub/ja-sor97-1.ps>

Arguello, Michael, Jonathan Bard, and Gang Yu (1997) "An Optimization Model for Aircraft Routing in Response to Groundings and Delays", Working paper, University of Texas - Austin.

Bailey, E., Graham, D. and Kaplan, D., 1985. *Deregulating the Airlines*. The MIT Press, Cambridge, MA.

Barnhart, Cynthia et. al. (1997) "Flight String Models for Aircraft Fleeting and Routing", Working paper, MIT Center for Transportation Studies.

Berge, Matthew and Craig Hopperstand. (1993) "Demand Driven Dispatch: A Method for Dynamic Aircraft Capacity Assignment, Models and Algorithms", *Operations Research* 41 (1), pp. 153-168.

Bruinsma, F.R., Rietveld, P. and van Vuuren, D.J., 1999. Unreliability in public transport chains. In: Meerman, H., van de Voorde, E. and Winkelmanns, W., Editors, 1999. *World Transport Research*. Vol. 1: Transport Modes and Systems, Elsevier, Amsterdam.

Cao, Jia-Ming and Adib Kanafani. (1997) "Real-time Decision Support for Integration of Airline Flight Cancellations and Delays, Part I: Mathematical Formulation", *Transportation Planning and Technology* 20, pp. 183-199.

Cao, Jia-Ming and Adib Kanafani. (1997) "Real-time Decision Support for Integration of Airline Flight Cancellations and Delays, Part II: Algorithm and Computational Experiments", *Transportation Planning and Technology* 20, pp. 201-217.

Carey, Malachy, Reliability of interconnected scheduled services, *European Journal of Operational Research*, Volume 79, Issue 1, 24 November 1994, Pages 51-72.
(<http://www.sciencedirect.com/science/article/B6WPV-3TJ9W0W-1P1/1/1d7d868e1384947a2c8d8b17ab91824e>)

Carey, Malachy, Optimizing scheduled times, allowing for behavioral response, *Transportation Research Part B: Methodological*, Volume 32, Issue 5, June 1998, Pages 329-342.
(<http://www.sciencedirect.com/science/article/B6V99-3VR0F32-4/1/3b365b346dcedac3aeaa4316ef325fc4>)

Carey, Malachy, Ex ante heuristic measures of schedule reliability, *Transportation Research Part B: Methodological*, Volume 33, Issue 7, September 1999, Pages 473-494.
(<http://www.sciencedirect.com/science/article/B6V99-3WXP2SV-2/1/4b459881455488510659e4834c1c81f1>)

Clarke, M. "Solving the Problem of Irregular Airline Operations", presented at the INFORMS New Orleans Fall 1995 Meeting, October 30, 1995

Clarke, Michael (1997). "The Airline Schedule Recovery Problem", working paper, MIT International Center for Air Transportation.

Clarke, Michael (1997). "Irregular Airline Operations: A Review of the State-of-the-Practice in Airline Operations Control Centers", Working paper, MIT International Center for Air Transportation. Submitted to Journal of Air Transportation Management

Douglas, G.W. and Miller, J.C., 1974. Economic Regulation of Domestic Air Transport: Theory and Policy. The Brookings Institution, Washington, D.C.

Engineering Statistics Handbook: <http://www.itl.nist.gov/div898/handbook/>

Eschmaier, M.M. & Mathaisel, D. F. (1985). Airline Scheduling: An Overview, *Transportation Science* 19(2): 127-138

Ghobrial, A., Balakrishnan, N. & Kanafani, A (1992). A heuristic model form frequency planning and aircraft routing in small size airlines, *Transportation Planning and Technology* 16: 235-249

Grandeau, Seth. (1995) The Processes of Airline Operation Control. MIT Flight Transportation Laboratory Report R95-2.

Jarrah, Ahmad et. al. (1993) "A Decision Support Framework for Airline Flight Cancellations and Delays", *Transportation Science* 27 (3), pp. 266-280.

Lohatepanont, M. Airline Fleet Assignment and Schedule Design: Integrated Models and Algorithms, PhD Thesis. Massachusetts Institute of Technology 2001

Mathaisel, Dennis F. X. Decision support for airline system operations control and irregular operations, *Computers & Operations Research*, Volume 23, Issue 11, November 1996, Pages 1083-1098.

(<http://www.sciencedirect.com/science/article/B6VC5-3VV40WT-5/1/3e23baebe3b02a54bcb25231a6188de6>)

Morrison, S.A. and Winston, C., 1985. An econometric analysis of the demand for intercity passenger transportation. *Research in Transport Economics* 2, pp. 213-237. EconLit

Morrison, S.A. and Winston, C., 1986. The Economic Effects of Airline Deregulation. Brookings Institution, Washington, D.C.

Prousaloglou, K.E. and Koppelman, F.S., 1995. Air carrier demand: an analysis of market share determinants. *Transportation* 22, pp. 371-388.

Prousaloglou, Kimon, Koppelman, Frank S. The choice of air carrier, flight, and fare class, *Journal of Air Transport Management*, Volume 5, Issue 4, October 1999, Pages 193-201.

<http://www.sciencedirect.com/science/article/B6VGP-3XX69PR-3/1/2d9d4783022b5717c578aff330f15862>)

Rakshit, Ananda, Nirup Krishnamurthy, and Gang Yu (1996) "Systems Operations Advisor: A Real-time Decision Support System for Managing Airline Operations at United Airlines", *Interfaces* 26 (2), pp. 50-58.

Rushmeier, R.A. & Kontogiorgis, S.A. (1996). Advances in the optimization of airline fleet assignment, Technical report, USAir Operation Research Group, published on *Transportation Science*, 31:159--169, 1997.

Rietveld, Piet, Brons, Martijn. Quality of hub-and-spoke networks; the effects of timetable co-ordination on waiting time and rescheduling time, *Journal of Air Transport Management*, Volume 7, Issue 4, July 2001, Pages 241-249.

<http://www.sciencedirect.com/science/article/B6VGP-43G38G6-5/1/514e78ff6515099a7c39768ed5a931ba>)

P. Rietveld, F. R. Bruinsma and D. J. van Vuuren, Coping with unreliability in public transport chains: A case study for Netherlands, *Transportation Research Part A: Policy and Practice*, Volume 35, Issue 6, July 2001, Pages 539-559.

<http://www.sciencedirect.com/science/article/B6VG7-42Y1490-4/1/58f7555a2ff4141ca19c4c8e37fd9d6f>

Suhl, Leena (1995), *Computer Aided Scheduling: an Airline Perspective*, *Deutscher Universitäts Verlag*, Wiesbaden.

Talluri, Kalyan (1996). "Swapping Applications in a Daily Airline Fleet Assignment", *Transportation Science* 30 (3), pp. 237-248.

Teodorovic, Dusan and Slobodan Guberinic (1984) "Optimal dispatching strategy on an airline network after a schedule perturbation", *European Journal of Operational Research* 15, pp. 178-182.

Teodorovic, Dusan (1985) "A Model for Designing the Meteorological most reliable Airline Schedule", *European Journal of Operational Research* 21, pp. 156-164.

Teodorovic, Dusan and Goran Stojkovic (1990) "Model for Operational Daily Airline Scheduling", *Transportation Planning and Technology* 14, pp. 273-285.

Teodorovic, Dusan and Goran Stojkovic (1995) "Model to Reduce Airline Schedule Disturbances", *Journal of Transportation Engineering* 121 (4), pp. 324-331.

Teodorovic, D. and G. Stojkovic, "Model to Reduce Airline Schedule Disturbances," *ASCE Journal of Transportation Engineering*, Vol. 124, No. 4, July/Aug 1995.

Yan, Shangy and Dah-Hwei Yang. (1996) "A Decision Support Framework for Handling Schedule Perturbation", *Transportation Research: PartB* 30 (6), pp. 405-419.

Zils, M. (1996). Genetische Algorithmen zur strategischen Flotteneinsatzplanung in der Luftfrachtindustrie, Diplomarbeit, Universität zu Köln. *Seminar für Wirtschaftsinformatik und Operations Research*.

3231-29