Title

Real-time Inter-modal Strategies for Airline Schedule Perturbation Recovery and Airport Congestion Mitigation under Collaborative Decision Making (CDM)

Author

Zhang, Yu

Publication Date

2008-10-01

University of California Transportation Center
UCTC Dissertation No. 154

Real-time Inter-modal Strategies for Airline Schedule Perturbation Recovery and Airport Congestion Mitigation under Collaborative Decision Making (CDM)

Yu Zhang
University of California, Berkeley
2008

Real-time Inter-modal Strategies for Airline Schedule Perturbation Recovery and Airport Congestion Mitigation under Collaborative Decision Making (CDM)

by

Yu Zhang

Real-time Inter-modal Strategies for Airline Schedule Perturbation Recovery and Airport Congestion Mitigation under Collaborative Decision Making (CDM)

by

Yu Zhang

Real-time Inter-modal Strategies for Airline Schedule Perturbation Recovery and Airport Congestion Mitigation under Collaborative Decision Making (CDM)

by

Yu Zhang

B.S. (Southeast University) 1997

M.S. (University of California, Berkeley) 2003

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering-Civil and Environmental Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Mark M. Hansen, Chair
Professor Dorit S. Hochbaum
Professor Martin Wachs

Fall 2008

The dissertation of Yu Zhang is approved:

Chair ______________________________________ Date _______________

______________________________________ Date _______________

______________________________________ Date _______________

University of California, Berkeley

Fall 2008

Real-time Inter-modal Strategies for Airline Schedule Perturbation Recovery and Airport Congestion Mitigation under Collaborative Decision Making (CDM)

by

Yu Zhang

Abstract

Real-time Inter-modal Strategies for Airline Schedule Perturbation Recovery and Airport Congestion Mitigation under Collaborative Decision Making (CDM)

by

Yu Zhang

Doctor of Philosophy in Civil and Environmental Engineering

University of California, Berkeley

Professor Mark M. Hansen, Chair

The main goal of this dissertation is to propose a new analytical framework and supporting optimization models that will encourage the aviation industry to incorporate alternative transportation modes when major airports in the system encounter temporary closures or severe capacity deficiencies. This framework can provide a way to reduce passenger disutility due to delay and misconnection, to help airlines reduce operating cost and recover schedule more promptly, and to assist air traffic flow managers to utilize and distribute scarce resources more efficiently and equitably.

Airline delays cost billions of dollars each year in the U.S. Most of the delays occur on days when airlines' planned schedules are disrupted and off-schedule operations (OSO) have to be performed. One main reason for the disruptions is airspace capacity shortfall caused by adverse weather or other temporary events. It is suggested in this study that when there is a significant capacity shortfall, airlines with hub-and-spoke networks could incorporate ground transport modes into their operations. Real-time inter-modalism

includes the substitution of flights by surface vehicle trips and, when the hub is part of a regional airport system, the use of inter-airport ground transport to enable diversion of flights to alternate hubs. Real-time inter-modalism is different from the air-rail cooperation currently practiced in Europe because it is only triggered by severe demand and supply imbalance at major hub airports and emphasizes operational integration of existing airside and ground transport capabilities rather than major capital investment.

In the first strategy, Real-time Inter-Modal Substitution (RTIMS), airlines substitute short-haul flights with ground transport modes during severe disruptions. In this way, scarce arrival and departure slots can be used by long-haul, large jets with more passengers or other high-priority flights. As a first step, a deterministic queuing analysis is used to identify flights whose substitution by ground transport would result in net time savings, based on a comparison of the reduced flight delay and increased line-haul time that would result. Then for arrivals and departures of one airline, a mathematical programming model is constructed to help the airline make decisions on whether to cancel flights, substitute them with motor coaches, or assign them delays commensurate with airport capacity constraints. An approximation algorithm is proposed to reduce the substantial computation time required to solve large-scale non-linear integer programming problems. A set of experiments are designed to assess potential savings from ground transport substitution. Sensitivity analyses are conducted to investigate the effects of severities of capacity shortfalls, passenger value of time, distances of short-haul spoke airports, load factor of the flights, schedule peaking, and connecting patterns of transfer passengers on the optimization results and the savings from inter-modal substitution.

Many major metropolitan areas are served by one or two major hub airports and several surrounding regional airports, which collectively form a regional airport system. Some of the regional airports are underutilized and contain runways long enough to serve most of the commercial aircraft types. Hence, the second strategy of inter-modalism is to divert hub-bound flights affected by capacity shortfall at the hub airport to regional airports (called alternative hubs). The strategy is termed real-time inter-modal diversion (RTIMD). In contrast to diversions currently implemented in airline operation, this strategy moves passengers between the airports as necessary to access diverted flights or make inter-airport connections. Toward this end, dedicated ground transportation services will be available to transport passengers between a primary hub and alternative hub. Like RTIMS, RTIMD also allows short-haul flights to be substituted by motor coaches. To implement this strategy, the alternative hub is selected based on evaluation of maximum runway length, driving distance, and correlation of weather impact and demand profiles. An extension of the mathematical programming model of RTIMS is proposed and similar approximation algorithm is used to obtain optimum results for a case study. Considering the possibility of the alternative hub being overwhelmed by diverted flights, it is suggested to enhance the current Ground Delay Program (GDP) to determine the Controlled Time of Arrivals (CTAs) at both the major hub airport and the alternative hub airports. That proposed enhanced GDP is termed Regional GDP.

There are fundamental issues that need to be considered while implementing the strategies, such as motor coach service provision, passenger, security, and airport facility issues. This study identifies and assesses these issues, suggests solutions based on preliminary investigation, and highlights needs for further research and policy decision making.

The inter-modal framework proposed in this thesis will substantially reduce the costs of recovering from major hub capacity shortfalls by providing alternatives to simple cancellation in airlines' schedule perturbation recovery, thus reducing the number of disrupted passengers and the delay propagated to later flights and other parts of the network.

Professor Mark M. Hansen

Dissertation Committee Chair

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This research was conducted under the supervision of Professor Mark Hansen, to whom I am deeply indebted for his continuous guidance, inspiration, and encouragement through my graduate studies at the University of California, Berkeley. It was Professor Mark Hansen who guided me into the aviation area and gave me constant support for me to carry out this research. It is a privilege for me to have had the opportunity to work with him and I will continuously benefit from the technical knowledge and working style that I have learned from him. Professor Mark Hansen is always reachable and dedicating his time to research discussion with students, which made my graduate study an affirmative and inspirational experience.

I would like to thank two other members of my committee, Professor Martin Wachs and Professor Dorit Hochbaum. I got to know Professor Martin Wachs when I was auditing his class before officially enrolled into UC Berkeley. I am grateful to Professor Martin Wachs for his encouragement of pursuing higher degrees and bigger dreams. I sincerely appreciate the guidance and support from both of my committee members.

I would also like to acknowledge Professor Adib Kanafani, Professor Carlos Daganzo, Professor Michael Cassidy, Professor Samer Madanat, Professor Elizabeth Deakin, Professor Alexander Skabardonis, Professor Raja Sengupta, Professor Shmuel Oren, and Professor Botond Koszegi. What I learned from their courses have strengthened my background and helped me to conduct the research. I also sincerely appreciate many valuable comments and suggestions from Professor Michael Ball of the University of

Maryland, Dr. Robert Hoffman from Metron Aviation, and Mr. Scott Nelson from American Airlines. I would like to thank the customer service managers at United Airlines and the representatives from charter service companies that I have interviewed.

My graduate study has been much delightful because of many fellow Berkeley students. I would like to express my appreciation to them, particularly to Monica Menendez, Avijit Mukherjee, Megan Smirti, Anne Goodchild, Wanjira Jirajaruporn, Gautam Gupta, Helen Yin, Meng Li, Yanfeng Ouyang, Yafeng Yin, Jing Xiong, Amy Kim, Yoonjin Yoon, for their friendship and academic help. Moreover, I would like to thank Pei-Chen Liu and Chieh-Yu Hsiao for allowing me to benefit from their experiences and broad technical knowledge in air traffic management.

I would like to thank my mother, Yuying Chen, and brother, Yong Zhang, for their constant encouragement throughout my graduate study far away from home. And I want to dedicate this work to my dear father, Mingben Zhang, who passed away during the period when I was working on this dissertation. I know that he is always watching me and blessing me no matter where he is.

Finally, I would like to thank my husband, Qing Lu. It would not have been possible for me to finish the study without his love and long-lasting support.

This work was supported by dissertation grant from the University of California Transportation Center.

CHAPTER 1: INTRODUCTION

"Transportation is key to the productivity, and therefore the success, of virtually every business in America. Congestion and delay not only waste our time as individuals, they also burden our businesses and our entire economy with inefficiency and higher costs."

- Secretary of Transportation Normal Y. Mineta

January 2001

The main goal of this dissertation is to propose a new analytical framework and supporting optimization models that will encourage the aviation industry to incorporate alternative transportation modes when major airports in the system encounter temporary closures or severe capacity deficiencies. This framework can provide a way to reduce passenger disutility due to delay and misconnection, to help airlines reduce operating cost and recover schedule more promptly, and to assist air traffic flow managers to utilize and distribute scarce resources more efficiently and equitably.

Transportation congestion has been identified as one of the fundamental impediments to economic growth. Airline delays have been estimated to waste $9.4 billion a year. [1] In May 2006, a *National Strategy to Reduce Congestion on America's Transportation Network*[2] was launched. The initiative furnishes a comprehensive blueprint for addressing transportation congestion at all levels of government and the private sector and is

---

[1] http://www.dot.gov/affairs/dot5706.htm, Office of Public Affairs, U.S. DOT, accessed at Feb. 26, 2007

[2] National Strategy to Reduce Congestion on America's Transportation Network http://isddc.dot.gov/OLPFiles/OST/012988.pdf, accessed at Feb. 26, 2007

considered by the U.S. Department of Transportation (USDOT) as the most important new undertaking of in many years [Shane 2006]. USDOT and the Federal Aviation Administration (FAA) have been working for the past three decades on a fundamental overhaul in the management of air traffic. As part of the new national strategy, the Next Generation Air Transportation System (NextGen) projects air traffic demand in the next 20 years and recommends possible ways to accommodate the growing demand without any sacrifice of safety or reliability. One of the objectives in the newly released concept of operations of NextGen is to minimize the impact of weather and other disruptions.[3]

Airports are pivots in the national transportation network. When demand at a major U.S. airport starts to reach its capacity, the airport becomes more vulnerable to disruptions such as adverse weather, equipment outages, security threats, and other transient events. For hub-and-spoke networks operated by most airlines, capacity shortfalls at hub airports cause enormous aircraft delays and numerous passenger misconnections, making delay reduction one of the most critical issues facing the air transportation system. Volume-related delay (delay caused by excessive demand) can be addressed by controlling demand of scheduled flights, through either market or administrative mechanisms. These measures, however, are not appropriate for temporary-event-related delay because limiting schedules to the level that can be accommodated on those days will unnecessarily constrain airline operations on normal days. Instead, a real-time or

---

[3] Concept of Operations for the Next Generation of Air Transportation System, Joint Planning and Development Office, Draft 5, Version 1.2. Feb. 28, 2007, p1-3.

emergency-response congestion relief strategy is needed to solve the temporary-event-related airport capacity reduction problem.

The inspiration for this thesis comes from the following comments. “The transportation chaos in the aftermath of 9/11 tells Americans that the US is excessively reliant upon a single mode of transportation, inter-modal connectivity is poor in many parts of the country, and intercity commercial passenger transportation alternatives are poor or nonexistent…There are urgent needs to develop an integrated transportation system in the U.S. …”[4] A full-scale inter-modal system requires comprehensive planning, nation-wide coordination, much time, and enormous capital investment; but applying it as a relief strategy for congestion due to temporary airport capacity reduction is much easier to implement. It may also be a natural stepping stone to a full-scale inter-modal system.

In the remainder of this chapter, we first provides background for our subject. Topics addressed include airport and airspace delay, air traffic flow management, airline operations, inter-modal transportation, and regional airport systems. We next present the objectives, approaches, and expected contributions of this research.

---

[4] National Center for Inter-modal Transportation. “A New Agenda for America: in the Aftermath of 11 September 2001”. November 2001.

1.1. Airport Capacity Reduction

1.1.1. Adverse Weather and Airport Capacity

Airports are acknowledged to be bottlenecks in the National Airspace System (NAS). Most delays do not originate from chronic over-scheduling but from airport capacity reduction due to the adverse weather [Nilim et al. 2002; Kasper 2004]. Statistics obtained by processing data downloaded from the FAA Air Traffic Operations Network (OPSNET) database show that in the calendar year 2007, the percentage of weather-caused delayed flights are more than 50 percent at 36 of 45 OPSNET airports[5] and more than 70 percent at 19 of the 45 OPSNET airports (see Figure 1-1).

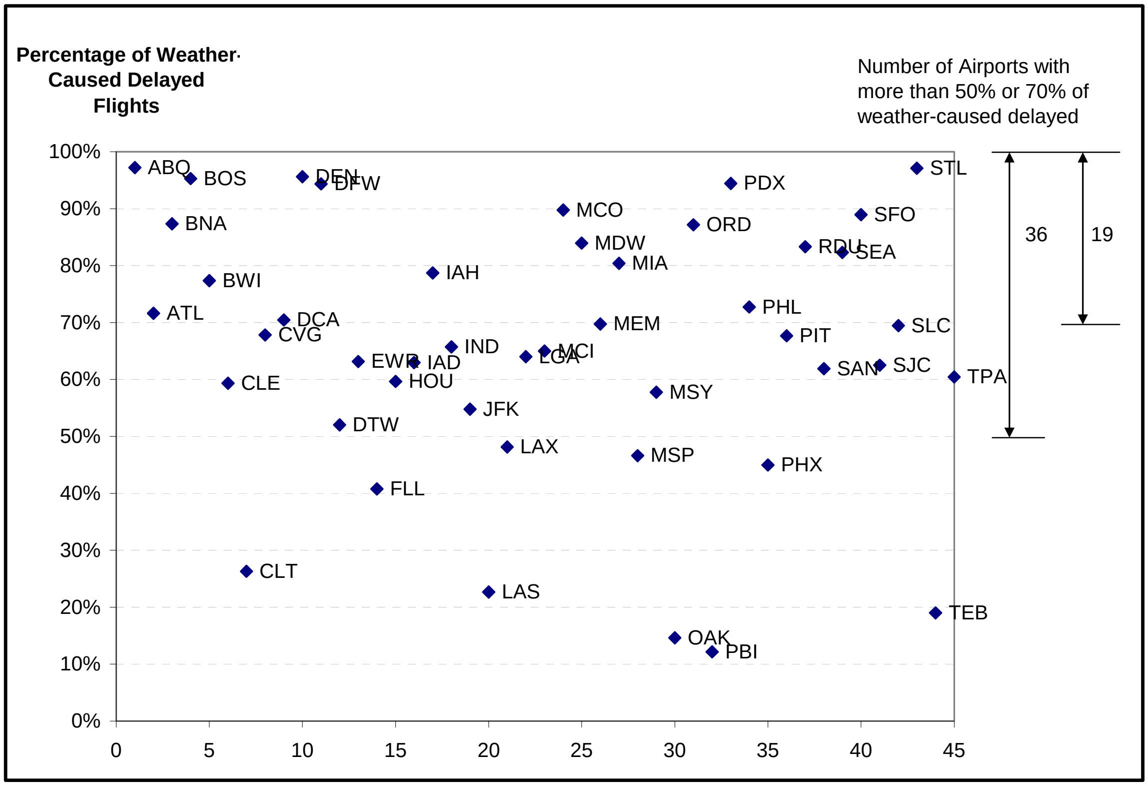

Figure 1- 1 Percentages of Weather-caused Delayed Flights at OPSNET Airports

---

[5] The airports selected in the ATC Daily Report that is produced each morning for the FAA's executive managers.

Weather conditions at terminals, such as low ceiling and visibility, a thunder storm, snow, and ice, reduce the airport's capacity. Convective en-route weather conditions reduce not only sector capacities but also acceptance rates at nearby airports. In the summer, almost half of serious delays at major airports are caused by thunderstorms [Forman et al. 1999]. In addition, undesirable headwinds, tailwinds, or crosswinds at terminals force the utilization of particular runway configurations with suboptimum capacities.

Weather-related delay may propagate from individual airports to the entire aviation network. To utilize resources efficiently, flights are dispatched in profitable markets at desirable times with limited time for turnaround. Hence, even a small perturbation may cause a large chain reaction. For example, the delay of an arrival flight at a hub airport may cause transfer passengers to miss their connecting flights. At the same time, a departure flight will also be affected if it uses the same aircraft or flight crew as the arrival flight. In a recent study of the spillover effect at LaGuardia airport (LGA) in New York city, estimation results from a simultaneous equation system showed that one minute of delay at LGA would cause three minutes of delay system-wide in the NAS [Hansen and Zhang 2005].

1.1.2. Outages and Airport Capacity

Outages of equipment located at airports or in their vicinity will increase aircraft spacing or even cause runway closure. The ground-based airport equipment includes navigation aids, Instrument Landing Systems (ILSs), and the tower [Nolan 1999]. ILS systems used under Instrument Flight Rule (IFR) conditions consists of localizer (LOC), glide slope

(GS), distance measuring equipment (DME), marker beacons (middle marker, outer marker, inner marker), and approach lighting systems (ALSs). Failure of any one of these components will lead to malfunction of ILS. Rakas and Schonfeld analyzed the NAS outages during the first 10 months of 1996 using the cause code in the National Airspace Reporting System (NAPRS) [Rakas and Schonfeld 1998]. Their results showed that the majority of outages involve the ILS. They constructed the distribution of total outage times for the top 17 NAS equipment types, and found that the top 13 types of equipment are related to ILS, navigation systems, and lighting. Although airport capacity during equipment outages depends on a variety of factors such as airport automation level, equipment redundancy, maintenance, air traffic control skills, and airport operational procedures, a numerical example with a deterministic queuing analysis showed that the capacity may decrease to only one third for a single-runway airport if ILS fails [Rakas and Schonfeld 2004].

1.1.3. Runway Incursions, Collisions and Airport Capacity

Runway incursions and consequent collisions may cause temporary closure of runways and lead to extensive flight delays. A runway incursion is "any occurrence at aerodrome involving the incorrect presence of an aircraft, vehicle, or person on the protected area of a surface designated for the landing and take off of aircraft."[6] A runway incursion can be caused by different personnel, such as air traffic controllers (OE/D), pilots (PD), and ground vehicle drivers (VPD). The historical trend of runway incursion from Year 2003

---

[6] The definition used by FAA Office of Runway Safety. http://www.faa.gov/runwaysafety/, accessed on January 12, 2008.

to Year 2007 is shown in Figure 1-2. A catastrophic runway collision may occur as a consequence of runway incursion. In 2000, after reviewing and analyzing the collisions at towered airports over 1983-1998, Barnett and Paull claimed that "recent patterns indicate roughly 15 fatal runway collisions over 2003-2022 at towered US airports. Most of these accidents would involve at least one large jet plane." [Barnett and Paull 2000] The main reason for such a large number of runway collisions is a projected increase of air traffic over the period. Strong evidence demonstrated that the risk of runway collisions varies with the square of the amount of traffic. However, the authors admitted that the claim is pessimistic because it did not consider the benefits of technological and various other initiatives underway in the US that aim to prevent runway collisions.

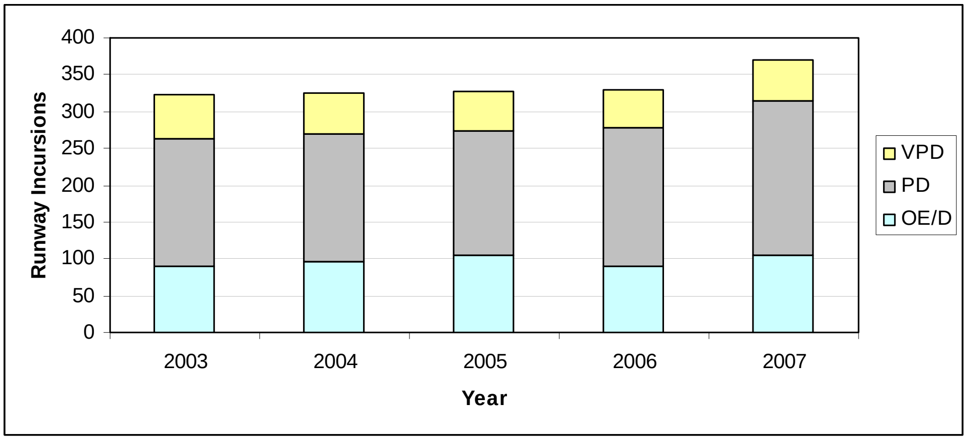

Figure 1- 2 Historical Trend of Runway Incursion[7]

1.1.4. Terrorist Threat and Airport Capacity

The terrorist threat to the US remains real and serious. Airports and government buildings are two major types of targets. On February 24th, 2006, a written bomb threat discovered

---

[7] Source: FAA Office of Runway Safety. http://www.faa.gov/runwaysafety/, accessed on January 12, 2008.

on a flight from Seattle to Chicago prompted aviation officials to temporarily close Chicago Midway Airport (MDW). A four-hour investigation which included a 20-minute airport shut-down caused “minimal delays” according to the airport spokesman.[8] The immediate threats, however, sometimes are hard to verify, so that airport closure could be much longer and consequent flight delay much more extensive. According to the Airport International web site, there have been worldwide a total of 827 terrorist incidents reported by airports from 1968 to 2006.[9]

1.2. Deterioration of Airspace System Performance with Increase of Traffic Demand in the Near Future

Because delays increase sharply when demand approaches capacity, demand increases predicted for the near future may significantly reduce airspace performance and make airports more vulnerable to the disruptions identified above. According to a recent demand forecast prepared for the Joint Program Development Office, air transportation demand will double in the next 20 years [Knorr 2006]. With the possible down-gauging of aircraft sizes, the flight demand at airports may even triple. Assuming that weather conditions in the years of 2014 and 2025 are the same as in 2004, the projected delays in these two years compared to the 2004 level are as shown in Figure 1-3. The higher delays still concentrate in the summer and in the winter. But due to the increased demand, the

---

[8] Source: Airlines Plagued by Recent Security Incidents, Fox News, August 25, 2006, http://www.foxnews.com/story/0,2933,210471,00.html, accessed in May 2008.

[9] Source: http://www.airport-int.com/categories/bomb-containment-and-removal/bomb-containment-and-removal.asp , accessed on January 12, 2008.

seasonal delay fluctuation is accentuated. The delay peak is predicted to reach six times that of its 2004 level within the same time frame. The longest average delay is predicted to occur in the winter when adverse weather reduces capacity at airports.

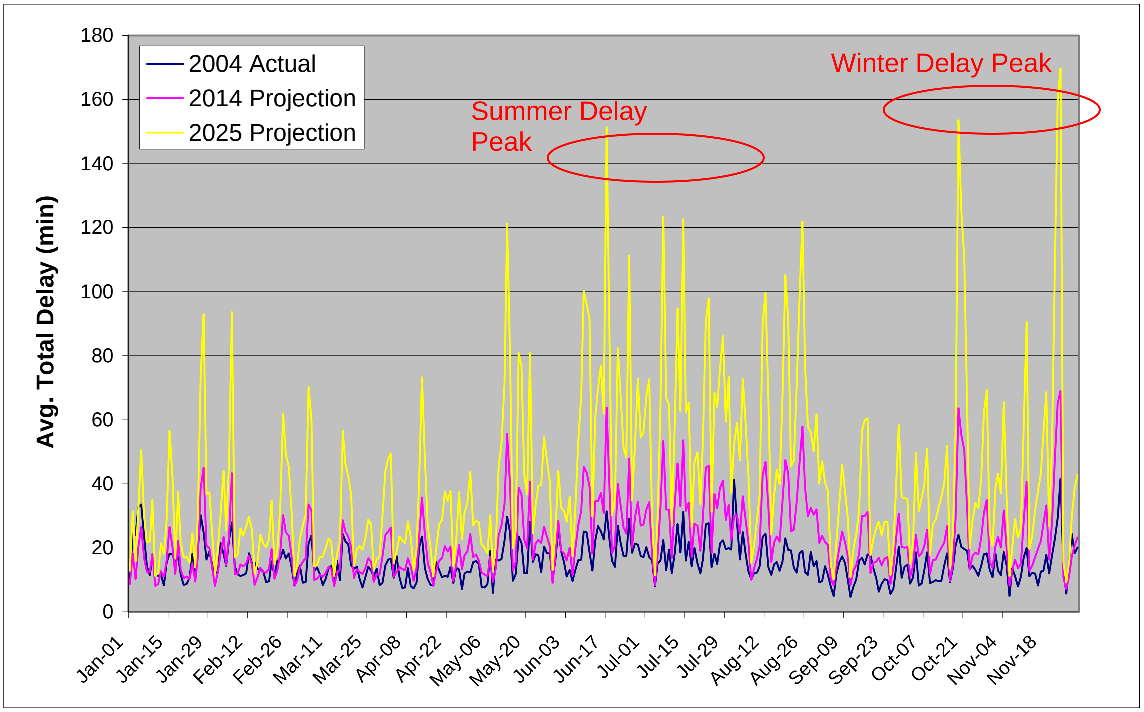

Figure 1- 3 Projected Delays in 2014 and 2025 [Knorr 2006]

The OPSNET database list delay causes including weather, terminal volume, center volume, equipment, runway, and other. The volume related delay, or called recurrent-delay, occurs when scheduled demand exceeds the capacities of the facilities. This kind of delay could be alleviated by long-term approaches such as expanding airspace capacities, or by shorter-term strategies, such as limiting the number of arrival and departures that can be scheduled into busy airports, using either market-based or administrative mechanisms. These measures, however, are not appropriate for solving the delay problems caused by adverse weather, equipment outages, and other temporary events, which can be called as incident-delay, because limiting schedules to the level that can be accommodated on days with unusual circumstances will reduce airline profit and level of service to passengers on normal days.

As will be described in the next section, airlines operating hub-and-spoke systems suffer the most from airport capacity reduction at major hub airports. Fortunately, due to the characteristics of their route network structure and the flexibility of the existing air traffic flow management system, there are potential ways to help mitigate the airport congestion and reduce delays during disruptions caused by adverse weather or other temporary events.

1.3. Airline Operations

There are two common structures in the airline route network; hub-and-spoke, and point-to-point. Since the 1980s most major US airlines have adopted the hub-and-spoke route structure, and selected one or more airports (hubs; see Figure 1-4) as connection points for the majority of their flights. At these connection points, passengers from different origins can transfer to the same destination flight, and passengers from the same origin can transfer onto flights with different destinations. Airlines exploit the economies of scope and density associated with this structure, offering frequent service on economically-sized aircraft in a large number of low-density markets [Zhang et al. 2004]. This structure, however, makes the operation at hub airports critical to the entire network. Airport capacity reduction and resulting delays cause extra operating costs for airlines and also inconvenience passengers. The loss of passenger goodwill may cause passengers to shift their business to other airlines, or dissuade them from traveling altogether.

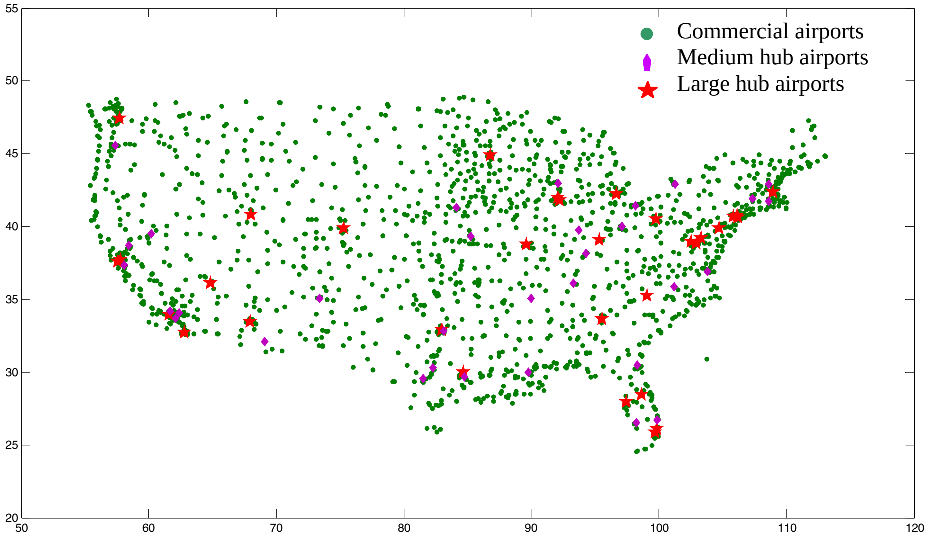

Figure 1- 4 Geographic Distributions of Airports in the Continental US

Capacity reduction at hub airports is a severe problem for airlines operating a hub-and-spoke system because it causes passenger misconnections, which account for to a large percentage of total passenger delay. Bratu and Barnhart [2006] calculated passenger delay using historical monthly data from a major airline operating a hub-and-spoke network with three hubs in the U.S. The data used in the study included passenger booking data, realized airline flight schedules, and passenger recovery priorities and policies. In their study, the authors categorized passengers into disrupted passengers and non-disrupted passengers. Disrupted passengers are those whose itineraries have been interrupted because of capacity reduction. Results showed that disrupted passengers encountered a delay of 303 minutes in average, about 20 times as large as that of non-disrupted passengers. Although the disrupted passengers were only three percent of the total passengers, they suffered 39 percent of the total passenger delay.

Currently, airlines are not responsible for passengers' extra costs resulting from weather-caused delay. However, the debacle of JetBlue airlines on Valentine's Day in 2007 led to a review of airlines operations during weather-related disruption and a call for Congress to pass a passenger bill of rights. One Congressman[10] proposes to restrict no-penalty weather related delay to three hours. For a longer delay, airlines would have to compensate passengers through cash refunds and voucher certificates. Although airlines and representative associations, such as the Air Transport Association, oppose such legislation by arguing that "inflexible standards could easily have the unintended effect of inconveniencing customers more in some situations," the serious consequences of weather-related delay has made major U.S. airlines review and update their customer service plans.[11] The airlines also called for a government review of airline and airport preparations for handling weather-related problem.

1.4. Background of Air Traffic Flow Management (ATFM)

Air traffic flow management (ATFM) was developed in the U.S. and in Europe in the 1980s and 1990s to facilitate air traffic management (ATM) and airport operations. The objectives of ATFM are to prevent overloading of the airspace system, to minimize the

---

[10] Democratic Rep. Mike Thompson

[11] Source: U.S. Airlines Pledge to Review Service Plans, USA Today, posted February 23, 2007, http://www.usatoday.com/travel/flights/2007-02-23-airlines-delays-response_x.htm, accessed in May 2008.

economic impact of air traffic congestion, and to avoid situations that might compromise safety [de Neufville and Odoni 2003]. One of the most important strategies that ATFM deploys to deal with overloading at destination airports is the ground delay program (GDP). If delay is estimated to be severe, a GDP is initiated to hold inbound aircraft at their original airports until there is reasonable assurance that, after departure, they will be able to proceed to their destination with a minimum amount of delay in the air [de Neufville and Odoni 2003]. GDPs are widely accepted because they reduce expensive and controller-workload-intensive airborne delays and save on fuel consumption. The main inputs for GDP planning are the weather forecasts, capacity forecasts derived from them, and demand forecasts.

Centralized ATFM may cause inefficient usage of arrival capacity. Under centralized ATFM, FAA assigns a ground holding time to each flight. Arrival slots, time windows with rights for landing, are allocated to airlines based on their original flight schedules and the first-scheduled/first-served principle, which is known as Ration-By-Schedule (RBS). Airlines, however, may want to assign slots differently once a GDP is in place because some flights are more important and should thus be given higher priorities than others. Furthermore, airlines may cancel a flight if its ground holding time is so long that it makes little sense to proceed with it. If the airline makes this known to ATFM soon enough, the flights after the cancelled one can be moved forward and the delay of those flights can be sequentially reduced. The airline that cancels the flight, however, may get less benefit than its competitors. In addition, announcing the cancellation too early may cause competitors to withdraw their intended cancellation to attract passengers from the cancelled flight. Thus, under centralized ATM, it is sometimes in the interest of airlines to hold the slot and not inform control centers until it is too late to take advantage of the gap created by the cancellation [de Neufville and Odoni 2003]. As a result, the slot is wasted.

To address these problems, the FAA implemented the Collaborative Decision Making (CDM) approach in 1998. CDM is an effort to improve GDP planning through "information exchange, procedural improvements, tool development, and common situational awareness" [Ball et al. 2000]. The process of GDP planning under CDM [de Neufville and Odoni 2003] is as follows: First, air traffic control system command center (ATCSCC) estimates the imbalance of arrival demand and capacity. Once delay is predicted to be severe, a GDP advisory is issued. The advisory gives a chance to airlines to cancel flights and thus prevent the GDP from being implemented. If the GDP goes forward, the first step is for FAA to implement Ration-by-Schedule (RBS). Airlines can freely cancel flights and reassign flights among the slots they receive via RBS. They transfer their response back to the FAA. Finally, a compression program is performed to take advantage of any empty slots.

Based on GDP records from 1999 to 2004, the distribution of GDP duration is shown in Figure 1-5. The mean is about 7.3 hours, and a standard deviation is about 4 hours. Figure 1-6 shows the annual number of days between 1998 and 2004 with GDPs for 38 airports in the U.S. Each circle in the figure represents the annual number of GDP days for one airport for one year between 1998 and 2004. Peaks in the figure correspond to airports with the more GDP days, which include Atlanta, Boston, Newark, LaGuardia, Chicago, Philadelphia, and San Francisco airports.

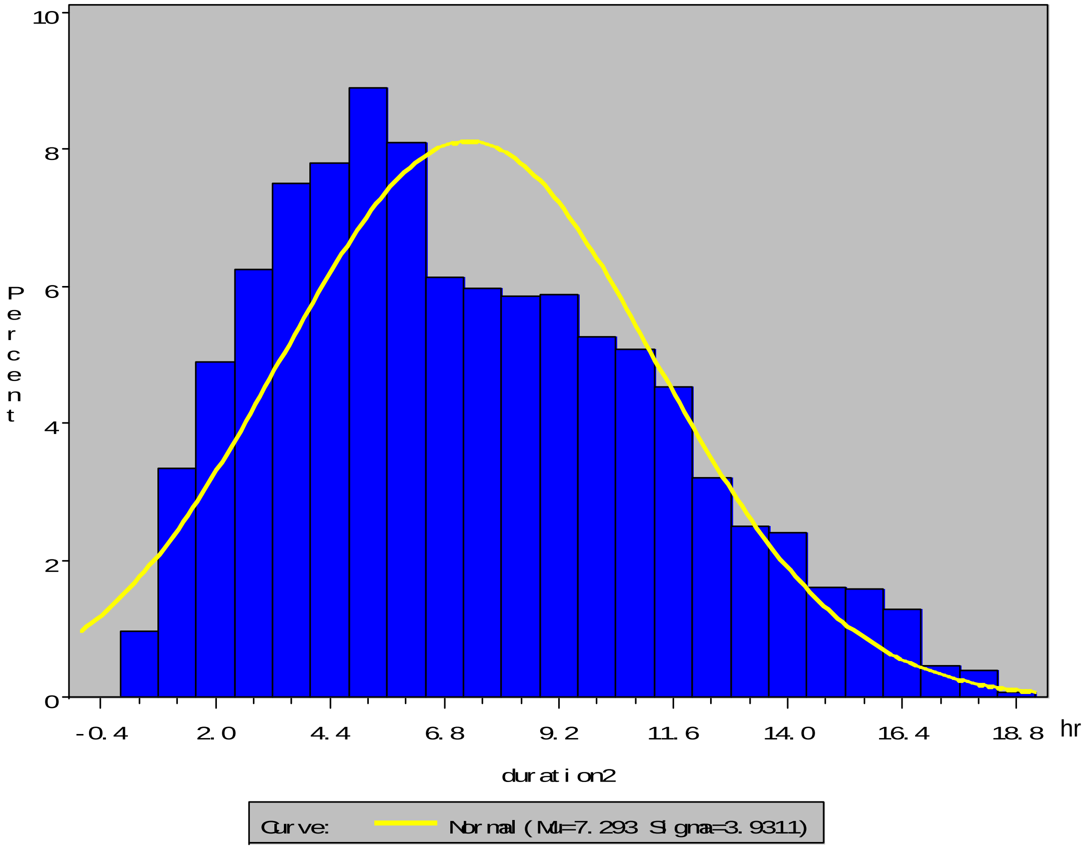

Figure 1- 5 Distribution of GDP Length from 1998-2004

Since 2003, a new mechanism named "slot credit substitution" (SCS) was introduced into CDM [Ball et al. 2005]. Compared to the compression program, a so-called batch-oriented periodic process, the SCS is a fast-response asynchronous process. The conditional language that airlines use to request slot adjustment in SCS is essentially: "we would like to cancel flight i if flight j can be advanced to slot k." SCS provides a near-real-time response to each such request and rearranges the flight sequence if the cancellation is confirmed. The enhanced GDP based on CDM yields large savings in flight delay by inducing airlines to announce their modifications of the schedule much earlier than before [Ball et al. 1998, 2005]. It also reduces the cost of delay by allowing airlines to allocate delay based on their internal business objectives. This encourages the airlines engaged in CDM to use more sophisticated decision-support tools to take advantage of its flexibility, and also provides opportunities for airlines to develop new strategies for managing disruptions caused by airport capacity reduction. Integrating ground transport mode with flight operations, called inter-modalism, is one such approach.

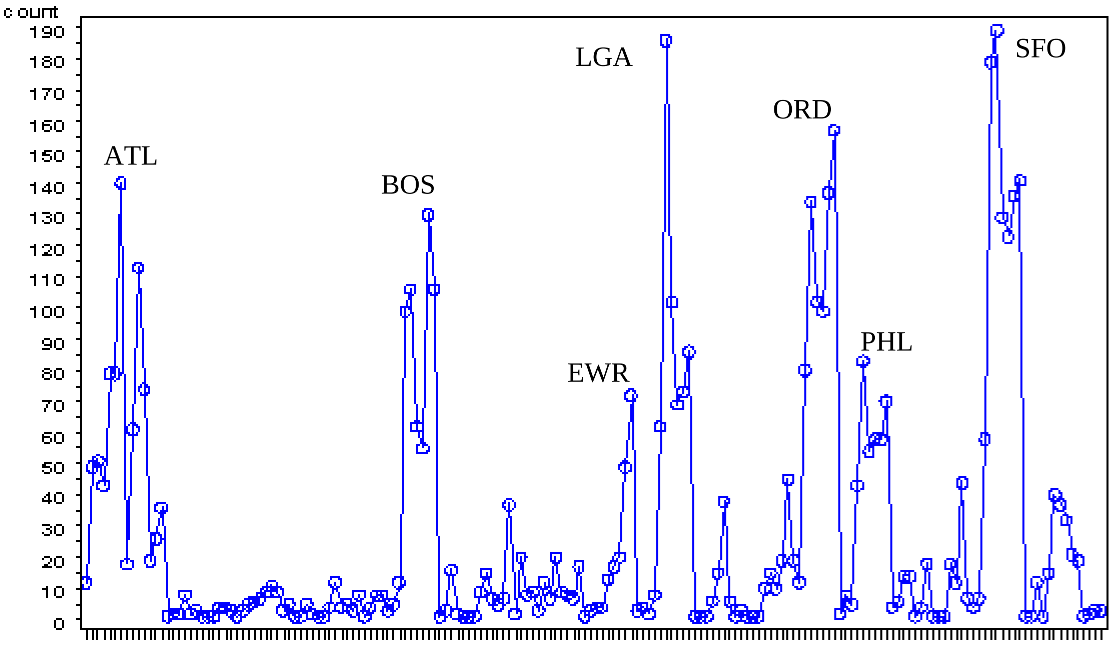

Figure 1- 6 Annual GDP Days at 38 Airports in the U.S. from 1998-2004

1.5. The Airport as an Inter-Modal Facility

Although there are different definitions of inter-modal transportation, a well-accepted one is “the concept of transporting passengers and freight on two or more different modes in such a way that all parts of the transportation process, including the exchange of information, are efficiently connected and coordinated.” [Muller 1999] Inter-modal systems for airline daily operation have been implemented in different countries to reduce short-haul flight frequencies. Air-Rail inter-modal systems have been applied since 1982 when Lufthansa airlines started their “Airport Express” program [Vetrovsky and Kanafani 1994]. Easy-jet, a low cost carrier in the EU, cooperates with charter coach service providers so that they can match their schedules and transport passengers between downtown and further/less-popular airports. High-speed rail in New York State and along the Northeast Corridor allow Continental Airlines (CO) at Newark Liberty (EWR) to cooperate with Amtrak and reduce their flight frequencies in several markets. For instance, passengers from Philadelphia to San Francisco (SFO) or London Heathrow through EWR can book a combination of train and air tickets through the CO reservation system. This cooperation gives CO a means of passenger recovery if they have to cancel remaining flights in those markets. According to the station manager at EWR, more passengers with trip destinations in Europe chose the combination of train and air modes.

Compared to freight transportation or passenger transportation in the European Union (EU), the development of passenger inter-modal transportation in the U.S. has lagged. This is especially true for inter-modal services involving aviation. The integration of transportation modes requires well-designed connection facilities. For example, the

European air-rail inter-modal system features intercity rail stations on the lower levels of major airports [Goetz and Vowles 2006]. In the U.S., lagging passenger rail system development is one of major factors limiting coordination between air and rail modes.

1.6. Current Inter-modal Practice in U.S. Airlines' Disruption Management

U.S. airlines, on occasion, use inter-modal substitution to cope with major disruptions in their operations. To investigate current practice we conferred with representatives at United (UA) and American Airlines (AA). UA operates with five hubs. Through conversations with customer managers at Chicago O'Hare (ORD) and Los Angeles (LAX), it was found that if flights with destinations close to ORD are cancelled, passengers will be re-accommodated on buses to avoid staying at the terminal overnight, or even longer when snowstorms are severe at ORD. The direct cost of hiring buses, as estimated by the customer relation division, is equivalent to the cost of providing discount vouchers for hotel accommodations and reassigning passengers on later flights. At UA's LAX hub, when thunderstorms lead to cancellations of flights out of the airport, the station occasionally hires buses to transport passengers to San Diego (SAN). The driving distance between LAX and SAN is only about two hours. As in the case of ORD, buses are used so that passengers do not have to stay at the terminal overnight.

In March 2007, due to a severe thunderstorm passing through Dallas Fort Worth International Airport (DFW), American Airlines (AA) cancelled flights from DFW to airports such as Will Rogers (OKC), Austin-Bergstrom (AUS), and San Antonio (SAT). Because of the cancellation, there would be no protected space for several days for those markets, and according to airline personnel, “buses became the best option.” The biggest problem that AA encountered was communicating to passengers about who would be re-accommodated on the buses and where to go. AA set up pseudo flight numbers for the buses and the passengers were booked according to first-landing-in-first-out (FLIFO). AA admits that the “BUS OPS” did not run perfectly, but it beat spending a night at the airport. As a result of this experience, they intend to make it a more formal part of the off-schedule operations (OSO) planning package with detailed procedures on how to handle similar situations. Customers’ feedback on this experiment was mixed. AA said that some of them were upset, stating that “if they wanted to ride the bus, they would have bought a bus ticket”, while others were just “happy to get out of town.”

It is intuitively reasonable to make the decision to use buses in the situations described above. Nevertheless, airlines are performing reactively in these cases, without a systematic way to collaborate airfield and ground operations or incorporating network-wide considerations in adjusting original schedules to reduce disruption costs and take maximum advantage of using motor coaches. This is the gap that we would like to fill with this study.

There are other transport modes that can be used in inter-modality in American air transportation, such as rail. The rail system in the U.S, however, does not have direct access to most major airports and the situation cannot be improved in the near future. According to AA and UA, it is fairly easy to obtain buses from their contacts at a local tour company. The customer manager at ORD said that they can get buses within one hour. We randomly surveyed motor coach charter companies and found similar quotes for lead times for such services. Thus our study will use motor coach charter service as the ground transport mode because (1) it does not require massive capital investment for facility development; (2) it is sufficiently flexible for real-time operations; and (3) it can provide terminal-to-terminal service without additional connections. However, the framework and methodologies described in this study can be extended to other modes fairly easily.

1.7. Regional Airport Systems

Many major metropolitan areas are served by one or two major hub airports and several surrounding regional airports, which collectively form a regional airport system. Some of the regional airports contain runways long enough to serve most commercial aircraft types. If these airports are underutilized, it is possible to mitigate disruption congestion at a major hub airport by using excess capacities at regional airports. We will refer to this idea as the alternative hub concept. To implement it, we need to quantify the capacity reduction at the major hub airport that would make utilization of other airports economical. In addition, dedicated, fast, and convenient ground transportation between airports can be used to transport connecting passengers or passengers with destinations at or originating from the major hub airport. This service can be provided by individual airlines or airport authorities. A special process also needs to be set up in the air traffic management system so that decisions to divert flights can be made collaboratively, as scheduling and cancellation decisions are under CDM.

Increasing the utilization of regional airports is an active research area in the aviation community since the ability to expand capacity at major airports is limited, so that, with growing traffic, existing flow patterns in the NAS would result in high delays. A National Aeronautics and Space Administration (NASA) report, *Virtual Airspace Modeling and Simulation System-wide Concept Report*, states that utilizing a better balance between hub airports and regional airports in close proximity to the hub will enable greater capacity in the NAS. Although the report emphasizes long-term planning, it provides some insights for our study on disruption response. In the report, the terminal airspace of large hub airports and smaller regional airports in combination is termed metro-plex. Integrated metro-plex planning and air traffic control are introduced to ensure that incoming and departing flows from both hub airports and regional airports do not interfere with each other, especially when they are reconfigured due to weather and traffic. The report proposes administrative polices about how to divert excess demand from congested airports to surrounding regional airports. To be selected to serve diversions, a regional airport has to be a public-use airport within 30 nautical miles of each hub airport and with asphalt runways in fair or better condition. In the report, flights with a great circle length of 1000 nautical miles or less are considered for diversion.

1.8. Research Objective, Approaches, and Contributions

Responding to the urgent need to mitigate delays on abnormal days with adverse weather or other temporary events, this study proposes a strategy to alleviate airport congestion and airline disruption cost in situations when the capacity of a hub airport is severely reduced or the airport is closed. The strategy does not require costly runway expansion;

rather, it suggests inter-modal substitution of short-haul flights for airlines to utilize scarce air field resources optimally. We call this real-time inter-modal substitution (RTIMS). We also propose an extension for this concept to regional airport systems, which we call real-time inter-modal diversion (RTIMD). Both concepts involve and integrate multimodal scheduling.

For RTIMS, ground transportation is available for substituting cancelled flights. Given slots assigned from CDM on abnormal days, airlines make decisions about delaying, canceling, and substituting scheduled flights with ground transport. For RTIMD, dedicated ground transportation between primary and alternate hubs is provided during disruption periods, allowing airlines to divert flights from the capacity-constrained major hub to alternates with surplus capacity. FAA, meanwhile, guides diversion decisions by providing airlines with controlled arrival times at alternate hubs as well as the major one. The objective of both FAA and airlines is to minimize the disruption cost which includes flight delay, passenger delay, and operating cost.

The remainder of this dissertation is organized as follows. In Chapter 2, existing literature in airport capacity analysis, airline disruption management, and inter-modal transportation is briefly reviewed. Chapter 3 presents the RTIMS concept and proposes a mathematical programming model to support airline cancellation, inter-modal substitution, aircraft assignment, and passenger re-assignment decisions. In Chapter 4, real-time inter-modalism is extended to regional airport systems in which airlines may consider flight diversion and inter-hub passenger transport as well as the options included in RTIMS. Implementation issues for RTIMS and RTIMD are discussed in Chapter 5, followed by conclusions and recommendations for future research in Chapter 6.

CHAPTER 2: LITERATURE REVIEW

In this chapter, we review several streams of literature of relevance to our topic. As we are concerned with use of inter-modal strategies to improve the ability of airlines to recover from disruptions, we first consider past work on *airline disruption recovery*. Next, since our work is intended to reduce the problem of airport congestion, we review literature on *airport congestion management*.

2.1 Airline Disruption Recovery

Airline recovery from disruption is a complicated problem due to the interactions between multiple resources, network propagation effects, and non-linear characteristics of the cost function. Exploration of this problem can be traced back to the 1980s. In 1984, Teodorovic and Gubernic developed a network model to minimize overall passenger delay in the circumstances where aircraft unexpectedly become unavailable. They used a branch and bound procedure to determine the least expensive set of aircraft routings and schedule plan [Teodorovic and Gubernic 1984]. The authors assumed a single fleet and applied their methodology to a small-sized network with eight flights. The methodology, however, is cumbersome for solving realistic-scale problems. Teodorovic extended his research later and coauthored a paper with Stojkovic in 1990, in which they addressed the airline scheduling and routing problems with heuristic algorithms based on dynamic programming [Teodorovic and Stojkovic 1990]. They constructed a two-stage optimization: first minimizing the total number of cancelled flights and then minimizing the total passenger delay on flights that are not cancelled.

The problem of recovering from unpredicted aircraft shortage and maintenance requirement attracted much attention in the 1990's. Jarrah, et al. presented a decision support framework for airline flight delay or cancellation implemented at United Airlines [Jarrah et al. 1993]. They modeled the schedule recovery as a minimum-cost network flow problem and considered delaying or canceling flights separately. Yan and Yang proposed four models for dealing with temporary unavailability of a single aircraft. The objective functions were to minimize the duration of disturbance and determine the most profitable schedule in the perturbation period [Yan and Yang 1996]. The four models handle different combinations of delaying, canceling, or ferrying[12] flights. Lagrangian relaxation and subgradient methods are used to find near-optimal solutions and good bounds. Their models are realized for a relatively small airline, China Airlines. Cao and Kanafani introduced an integrated delay and cancellation model at multiple airports [Cao and Kanafani 1997a, 1997b]. They presented a detailed calculation of downstream delay cost in the quadratic 0-1 programming model. Their model can be extended to formulate some special cases such as the ferrying of surplus aircraft, the replacement of different type of aircraft, and others.

GDP is a key component of Air Traffic Flow Management (ATFM). If traffic demand is expected to exceed the capacity at airports (or sometimes en route airspace), a GDP is

---

[12] Ferrying is defined as transporting an aircraft without passengers.

launched to hold flights at their origin airports unless there is reasonable assurance that, after departure, they will be able to proceed to their destination with a minimum delay in the air. In 1998, Songjun Luo and Gang Yu addressed airline recovery from a schedule perturbation stemming from a GDP. Their first paper is theoretical and methodological, investigating a landing assignment problem with different objectives by assuming that aircraft and crew are “unsplitable [Luo and Yu 1998a].” In their second paper, the unsplitable assumption was relaxed, which leads to a much more complex optimization problem for several objectives listed in the first paper [Luo and Yu 1998b]. With an objective to minimize the maximum delay of out-flights, the problem is NP-hard. As an alternative, they transformed the problem and proposed a heuristic so that real-life problem instances could be solved in a very short time period. Milner was the first to consider the connection dependencies of hub operations in the decision support model for resolving airline schedule disruption [Milner 1995]. He proposed an optimization model to minimize the costs of bank spread and flight cancellation for arrival operations at a hub airport. Carlson argued that the Milner model did not use scarce arrival slots efficiently [Carlson 2000]. He refined Milner’s model and provided a different formulation to test the computational efficiency for a large-scale problem. He assumed that the Collaborative Decision Making (CDM) procedure is in place and addressed a scenario in which the arrival capacity at an airline’s hub airport is significantly reduced and the hubbing airline must make tactical decisions with limited slots assigned by the FAA. Neither the Milner nor the Carlson model considers outbound operations or the connection between arrival and departure flights and banks.

Most recently, Bratu and Barnhart have been working on understanding passenger delays in legacy airlines hub-and-spoke networks [Bratu and Barnhart 2006]. They proposed a model for schedule recovery considering passenger delay. In their model, they took into consideration the crew cost due to insufficient crew needs caused by canceling or delaying flights. There are two papers reviewing the literature in this field: one by Filar et al. and the other by Kohl et al. Readers who are interested in more details are referred to these references [Filar et al. 2001; Kohl et al. 2004]. Nevertheless, there is no published literature discussing the integration of other modes into the air transportation network for airline recovery during disruptions caused by resource shortages.

The literature discussed above focuses on reassigning flights, crews, and passengers for operations between a fixed set of airports. More recently, researchers have considered options involving the use of additional airports to augment system capacity. Meyer et al. evaluated reliever and floating hub concepts for Dallas Fort-Worth Airport (DFW) [Meyer et al. 1997]. A reliever airport is a commercial service airport designated by the FAA to relieve congestion at primary airports and provide other general aviation services. Meyer et al. studied the feasibility and profitability of utilizing airports at cities such as Austin, San Antonio, Waco, and Abilene as reliever hubs for DFW, assuming that DFW is temporarily closed for two hours due to adverse weather. At selected reliever hubs, boarding gates and departure lounges are added to terminals to accommodate diverted passengers. In addition, if two or more airports are selected as reliever hubs, inter-hub shuttle flights with a capacity of 140 seats are provided to transport connecting passengers whose arrival and departure flights have not landed at the same reliever hubs.

Aircraft left over at the reliever hubs and passengers with destinations at DFW will be ferried back to DFW when the weather clears. The cost of implementing the reliever hub concept include the cost of geographic locations of candidate airports, the operating cost of flights between origin airports and potential reliever hubs, the cost of using an inter-hub shuttle aircraft, congestion cost at candidate airports, and infrastructure investment. American Airlines was chosen as a representative airline, but only 30 cities served by this airline were chosen in their case study in order to avoid the computational complexity of a large network. A simple mathematical programming model was used to minimize all annual costs resulting from reliever hub usage. According to the evaluation, Austin-Bergstrom International Airport (AUS) was chosen as the reliever hub. Furthermore, a flight schedule included 197 flights serving the top 13 markets was optimized with a genetic algorithm for three scenarios: normal operation, canceling the banks of flights, and diverting the banks of flights to the reliever hub. Frontier curves of net revenue versus average passenger delay were plotted. The authors concluded that developing a reliever hub system is an effective way of reducing airline schedule disturbances caused by adverse weather; however, high capital investment is needed to construct necessary ground facilities at the reliever hub(s), which may not be justified by the delay savings.

The virtual hub concept is defined by Karow and Clarke. In their study, virtual hub candidates are determined in a two-step process [Karow and Clarke 2002]. First, geographical locations and average delay status are considered. Then excess capacity, including runway capacity and airline gate utilization, is taken into account. A virtual hub network is implemented in two phases in the hours before the weather is predicted to impact the operations at the original hub. In Phase I, passenger flow is maximized iteratively over connecting bank time-windows until the weather is cleared. Disrupted passengers from implementing a virtual hub network in Phase I are reassigned to flights in Phase II by utilizing a Passenger Re-accommodation Module (PRM). Eventually, passengers that cannot be accommodated by the PRM are added to the next time window. The study demonstrated the application of the virtual hub network to a major US carrier. Their results showed a 94 percent reduction of number of passengers who are delayed over two hours compared to actual recovery.

These two papers studied the situation when major hub airport closure or capacity reduction is predicted hours in advance. In Meyer et al.'s paper, although possible inter-hub shuttle flights are mentioned, their operational impact was not explored. In Karow and Clarke's study, passenger flow is maximized without considering airlines' operating cost.

2.2 Airport Congestion Management

There is extensive literature on congestion alleviation from the airspace operators' point of view. A major focus is to find the optimal tradeoff between ground delays, which are less expensive and hazardous but must be assigned based on uncertain information about future capacity, and airborne delays, which are more expensive and potentially unsafe but are less affected by forecasting errors. Some of the studies considered decisions for individual flights [Hoffman and Ball 2000]; others modeled operations in terms of aggregate demand and supply [Richetta and Odoni 1993; Vranas et al. 1994; Rifkin 1998;

Ball et al. 2003; Mukherjee and Hansen 2006]. Dynamic versions of optimization models have also been proposed [Richetta and Odoni 1993; Vranas et al. 1994; Mukherjee and Hansen 2007]. With the introduction and development of CDM, many of the models can be applied to individual airline operations as well.

There is wide-ranging research on airline schedule problems under perturbations; however, previous work has focused more on perturbations resulting from unexpected aircraft shortage and crew shortage than on perturbations due to airport capacity reduction. There are very few papers addressing airline disruption recovery due to schedule perturbation and some of them have considered banking operations at hub airports. Studies have been conducted from the air traffic management point of view as well. Nevertheless, there is no literature has examined integrating airside and ground operations, and thereby exploiting the airport's capability to serve as an inter-modal facility, as a response to disruptions caused by temporary airside capacity shortfall.

CHAPTER 3: REAL-TIME INTER-MODAL SUBSTITUTION (RTIMS)

3.1. Research Motivations and Contributions

3.1.1. Research Motivations

Shortages of airline resources and airspace capacities are two major causes of airline operation disruptions which cost the airline industry billions of dollars annually. Shortages of airline resources occur for a variety of reasons such as crew sick-leave, aircraft mechanical failures, and lack of ground personnel or gates. Airspace resources include en-route (center) capacities and terminal (airport) capacities (For a detailed description of these two types of disruption see Bratu and Barnhart [2006]). Shortages of airspace capacities are caused by adverse weather, air traffic control facility malfunctions, security threats, and other factors.

Under the current CDM system, once a severe imbalance of traffic demand and terminal capacity supply is detected, a GDP advisory is issued, as illustrated in Figure 3-1. This advisory assigns scheduled flights Estimated Arrival Times (EATs), most of which are later than the original Scheduled Arrival Time (SAT). Airlines respond to this advisory by canceling flights. If the imbalance is resolved by these adjustments, the GDP is cancelled; otherwise the air traffic control system command center (ATCSCC) issues each remaining scheduled flight an Expected Departure Clearance Time (EDCT) and a Controlled Time of Arrival (CTA). Airlines manage their CTAs (or arrival slots) in their best internal business interests. An airline operation center (AOC)[13] may cancel more flights, re-order flight sequences, and re-assign flights to CTAs to utilize the arrival slots of delayed or cancelled flights. This flexibility of employing slots under CDM encourages airlines to cancel and re-order flights, and thereby to reduce air traffic demand, passenger delay, and disruption cost.

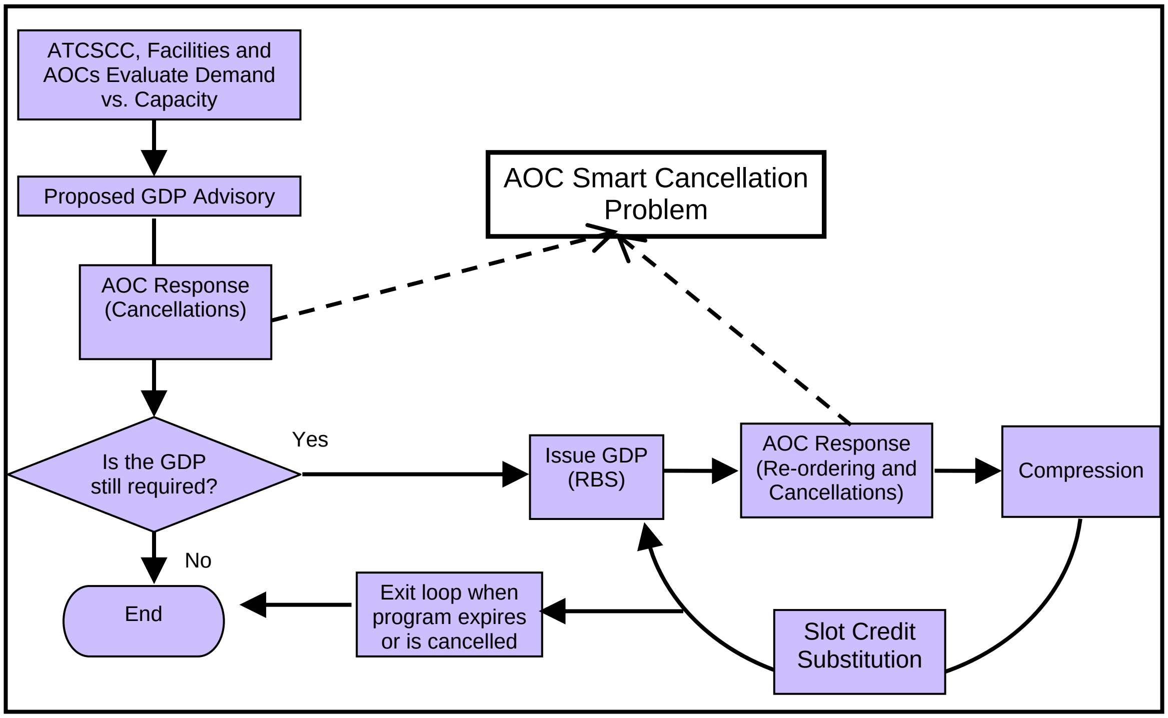

Figure 3- 1 Flow Chart of Collaborative Decision Making

Compared to airline resource shortages, disruptions caused by capacity deficiencies affect more flights and thus have more network-wide effects. Airlines' recovery from these kinds of schedule perturbations is therefore more important from the point of view of

---

[13] Airline Operation Centers (AOC) is a division to centrally manage operations of airline resources, monitoring the safety of operations, and exchange critical information with governmental authorities and other airlines.

both their own operations and the performance of the NAS as a whole. This study deals with these kinds of disruptions, especially when they involve a capacity reduction or a closure at a major hub airport. Such airports with their large numbers of flights and volumes of connecting traffic, serve as pivots of the NAS. We will now approach for mitigating the effects of temporary capacity reductions at these airports. Under this new strategy, when faced with severe capacity deficiencies at major hubs, the airline responds by integrating ground transportation modes into its hub-and-spoke network, implementing multi-modal scheduling to help reduce disruption costs. We will name this strategy real-time inter-modal substitution (RTIMS).

The intuition for using ground transport modes to substitute for short-haul flights is demonstrated in Figure 3-2. In the figure, based on San Francisco International Airport (SFO), the horizontal axis is the time of day and the vertical is the cumulative number of arrivals. The curve on the left is cumulative scheduled arrivals. The piecewise linear curve on the right is cumulative completed arrivals when the capacity of the airport is reduced due to temporary events. Completed arrival times are latter than schedule times because of the reduced airport capacity. Each triangle in the figure indicates when a “flight” could arrive at the airport if it were served by ground transportation. For Flight i in Figure 3-2, the expected delay if served by an aircraft is time t, i.e. the horizontal difference between scheduled and completed arrival curves. Flight i is a short-haul flight, thus the travel time difference between flying and using ground transport, t’, is smaller than the expected delay time t. In short, the delay saving of Flight i is (t-t’) if it is substituted with ground transportation.

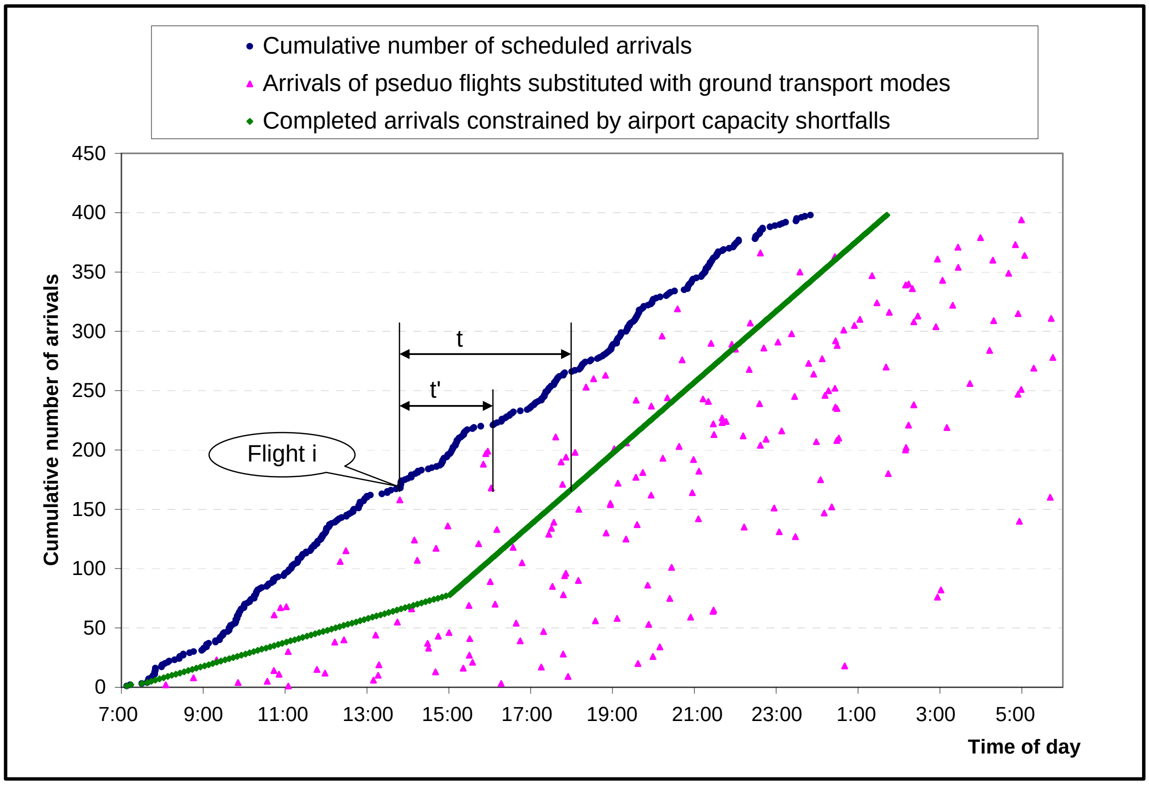

Figure 3- 2 Time Saving from Using Ground Transport Modes

This observation motivated an investigation of the U.S. airports to see if there are significant percentages of short-haul flights in the air transportation network that can be potentially substituted during disruptions. The results, as shown in Figure 3-3, revealed that there are nine large-hub airports in the continental U.S. with about 15 percent or more of their traffic consisting of short-haul flights with stage length of 160 miles or less.

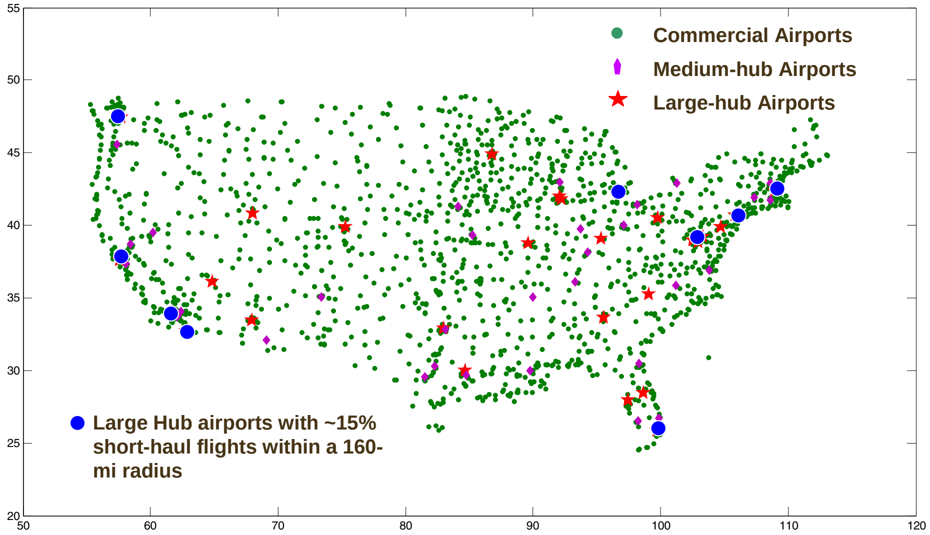

Figure 3- 3 Short-Haul Flight Statistics for Large-hub Airports in the Continental U.S.

3.1.2. Research Contributions

Airlines' currently make little use of inter-modal substitution in disruption recovery, and in the rare exceptions, decisions are made reactively, case by case, and unsystematically. As traffic increases, operations at major airports become more vulnerable to capacity reductions due to adverse weather and other temporary events. This suggests a need for airlines to develop a procedure to utilize other transport modes in disruption management by coordinating airside and groundside operations proactively. Benefits that airlines can achieve from implementing such a procedure include:

- More efficient use of scarce airspace resources and reduced aircraft delay
- Expedited passenger reassignment, especially when there is insufficient aircraft seating capacity in the markets where flight legs are cancelled
- Reduced passenger misconnection and delay cost
- Increased flexibility in crew recovery

The rest of this chapter is organized as follows. Section 3.2 investigates short-haul flight cancellations and substitutions with deterministic queuing analysis. A mathematical model is proposed in Section 3.3 to implement RTIMS. This is done for a one-stage hub-and-spoke network, in which passengers take either a direct flight or connect at most once at the hub airport. The model helps airlines coordinate airside and ground transportation to facilitate airline recovery from temporary airport capacity reductions. Complexity of this model is discussed and an approximation solution algorithm is presented in Section 3.4. The model is evaluated with experimental data in Section 3.5. In Section 3.6, sensitivity analysis is conducted to better understand the properties of the model. Section 3.7 gives a summary of this chapter.

3.2. Investigating Short-haul Flight Cancellations and Substitutions

The previous section briefly discussed the time savings for short-haul flights if they are substituted with ground transportation during disruptions. In this section, a deterministic queuing analysis will be conducted using real schedule data from SFO. The questions to be answered include:

- What are the numbers of short-haul flights that can save time if substituted with ground transportation?
- If those short-haul flights are substituted, and consequent vacant slots are made available to the remaining flights, what are the flight delay savings the result?
- How do the numbers of substitutions and amount of flight delay savings vary with the severity and timing of capacity shortfalls and the ground transportation speed?

3.2.1. Severity of Airport Capacity Shortfall

For a typical Thursday in November in 2005, there were around 400 flights scheduled to arrive at SFO between 7:00 am to midnight. It is assumed that the average speed of flying is 450 nautical miles per hour, the terminal time[14] is half an hour, and the average speed of ground transportation is 50 miles per hour. It is assumed that passengers' boarding and disembarking times are the same for the air and ground mode. For capacity shortfalls, three scenarios are considered:

I. Airport capacity drops to 10 arrivals per hour from 7:00 am to 12:00 pm and resumes 30 arrivals per hour afterwards.
II. Airport capacity drops to 10 arrivals per hour from 7:00 am to 3:00 pm and resumes 30 arrivals per hour afterwards.
III. Airport is closed from 7:00 am to 12:00 pm and resumes a capacity of 30 arrivals per hour afterwards.

FAA's airport capacity benchmark report shows that the medium Visual Meteorological Condition (VMC) capacity of SFO is about 60 arrivals per hour and Instrument Meteorological Condition (IMC) capacity is about 30 arrivals per hour. The comparatively low capacity levels in the scenarios are chosen to illustrate what would happen in the near future if traffic grew relative to capacity. To represent this situation, it is easier to reduce capacity than increase demand.

---

[14] Terminal time, also known as loss time, is the additional time required for a flight for to taxi, climb, and descend, and is assumed to be independent of flight distance.

Figure 3-4 shows the cumulative number of scheduled arrivals, and three hypothetical arrival curves under the assumed airport capacity shortfall scenarios. Each triangle in the figure indicates the time a pseudo flight using ground transportation would arrive at the airport. The horizontal differences between scheduled arrival times and pseudo flight arrival times reflect the time differences between flying and using ground transportation, assuming that departure times from origin airports are the same. Triangles falling between the scheduled arrival curve and a given completed arrival curve represent flights that will save time from being substituted, assuming all other flights are flown. We will refer to these as Ground-substitutable flights, or GSFs for short. Table 3-1 shows that with the increase of capacity shortfall severity, the numbers of GSFs increases.

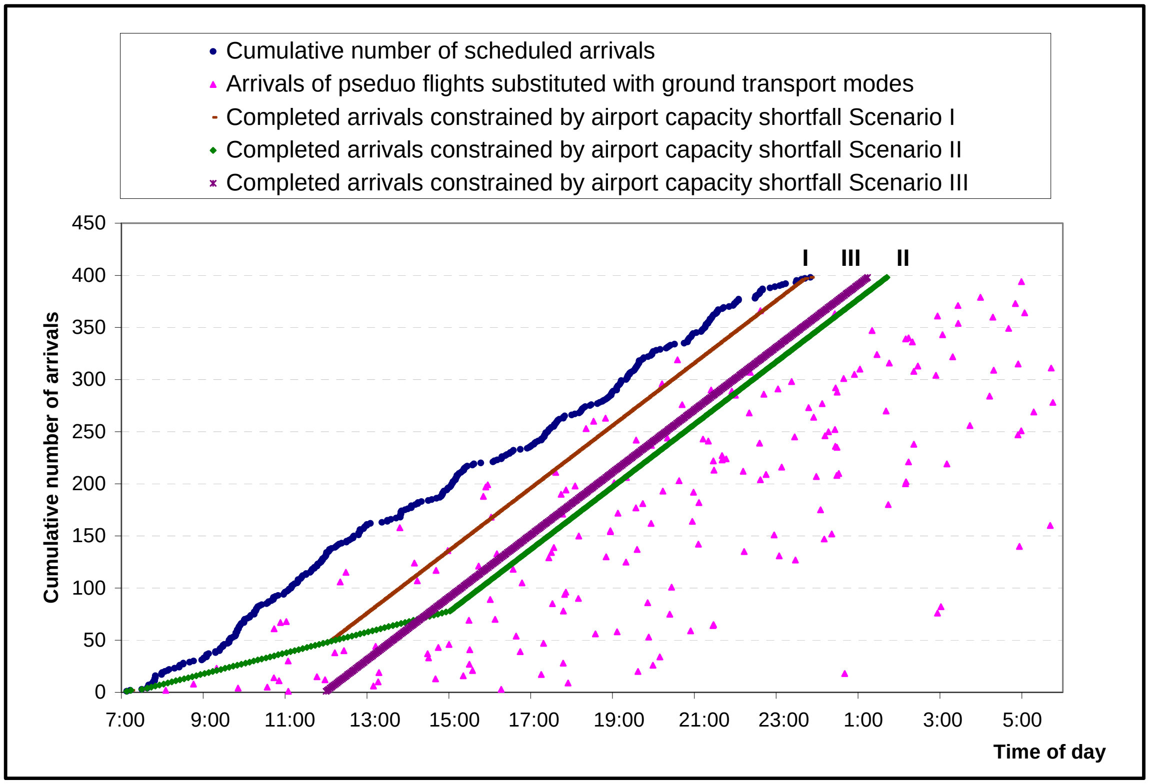

Figure 3- 4 Time Saving of Using Ground Transport Modes during Disruptions Caused by Capacity Shortfalls – Shortfalls with Various Severities

Table 3- 1 Comparisons of Flight Delays with and without Substitutions during Disruptions Caused by Capacity Shortfalls – Shortfalls with Various Severities

	Scenario I	Scenario II	Scenario III
Number of Flights Saving Time from being Substituted under Disruptions	17	38	46
Longest Flight Delay without Substitutions (hrs)	2.92	4.92	4.87
Longest Flight Delay with Substitutions (hrs)	2.70	4.70	4.50
Increase/ (Decrease)	(0.22)	(0.22)	(0.37)
Percentage of Increase or Decrease	-7.53%	-4.47%	-7.60%
Total Flight Delay without Substitutions (flight.hrs)	683.52	1354.35	1369.25
Total Flight Delay with Substitutions (flight.hrs)	533.72	977.58	821.68
Increase/ (Decrease)	(149.80)	(376.77)	(547.57)
Percentage of Increase or Decrease	-21.92%	-27.82%	-39.99%

Next, attention is turned to flight delays and the effect of substituting GSFs with ground transportation on the delays of other scheduled flights. Flight delay is calculated as the completed arrival time of a flight minus its scheduled arrival time. As shown in Table 3-1, the longest flight delays without substituting are 2.92 hours, 4.92 hours, and 4.87 hours for Scenarios I, II, and III, respectively. After GSFs are substituted with ground transportation, the remaining flights are moved up in the queue, reducing the longest flight delays to 2.70 hours, 4.70 hours, and 4.50 hours for Scenarios I, II, and III, respectively, which are 7.53%, 4.47%, and 7.60% less than the longest flight delays without substitutions in each scenario. In terms of total flight delay, the effect of inter-modal substitution is even more significant. If the GSFs are flown, the total flight delays are 684 flight-hours, 1,354 flight-hours, and 1,369 flight-hours under the three scenarios. Once they are substituted with ground transportation, the total flight delays go-down to 534 flight-hours, 978 flight-hours, and 822 flight-hours, respectively. Apparently, scenarios with more severe capacity shortfalls result in greater total delay reductions from inter-modal substitutions. The reductions can be as much as 40% of the original total flight delays. Thus inter-modal substitution is a “win-win”: it enables passengers on both the GSFs and the non-GSFs to arrive at the hub airport sooner.

3.2.2. Starting Time in the Day of Airport Capacity Shortfall

The starting time of the capacity shortfall will affect the number of GSFs and the operational impacts of inter-modal substitution. To illustrate the effect, we compare three scenarios —IV, V, and VI—in which the capacity shortfall starts at 7:00 am, 10:00 am, and 1:00 pm respectively and the shortfall lasts 5 hours before the capacity recovers to its normal level.

IV. Airport capacity drops to 10 arrivals per hour from 7:00 am to 12:00 pm and resumes 30 arrivals per hour afterwards.

V. Airport capacity drops to 10 arrivals per hour from 10:00 am to 3:00 pm and resumes 30 arrivals per hour afterwards.

VI. Airport capacity drops to 10 arrivals per hour from 1:00 pm to 6:00 pm and resumes 30 arrivals per hour afterwards.

Figure 3-5 shows the cumulative number of scheduled arrivals, arrivals of pseudo flights substituted with ground transportation, and completed arrivals under Scenarios IV, V, and

VI. Table 3-2 demonstrates that the latest starting time for the capacity shortfall leads to the smallest number of GSFs. For Scenarios IV and V, the numbers of GSFs are the same. Nevertheless, the longest flight delay and the largest total delay occur in Scenario V when the shortfall starts at 10:00 am. If GSFs are removed from the schedules, the largest total delay saving occurs in Scenario IV when the shortfall starts at the beginning of the day.

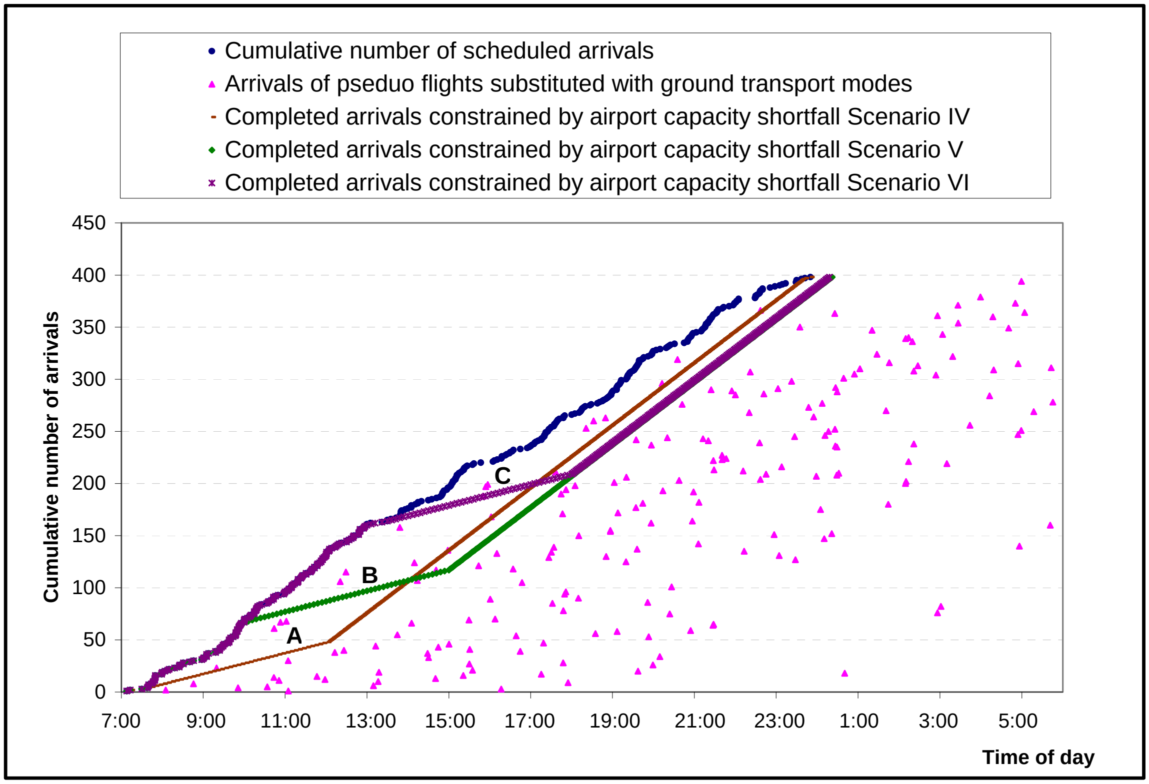

Figure 3- 5 Time Saving of Using Ground Transport Modes during Disruptions Caused by Capacity Shortfalls – Shortfalls with Different Starting Times of the Day

Table 3- 2 Comparisons of with and without Substitutions during Disruptions Caused by Airport Capacity Shortfalls – Shortfalls with Different Starting Time of the Day

	Scenario IV	Scenario V	Scenario VI
Number of Flights Saving Time from being Substituted under Disruptions	17	17	10
Longest Flight Delay without Substitutions (hrs)	2.92	3.58	2.82
Longest Flight Delay with Substitutions (hrs)	2.70	3.42	2.68
Increase/ (Decrease)	(0.22)	(0.16)	(0.14)
Percentage of Increase or Decrease	-7.53%	-4.47%	-4.96%
Total Flight Delay without Substitutions (flight.hrs)	683.52	715.82	386.60
Total Flight Dealy with Substitutions (flight.hrs)	533.72	565.05	318.73
Increase/ (Decrease)	(149.80)	(150.77)	(67.87)
Percentage of Increase or Decrease	-21.92%	-21.06%	-17.56%

3.2.3. Ground Transportation Speed

Previous sections have assumed that the average speed of ground transport modes used to substitute GSFs is 50 miles per hour. In this section, this parameter is varied to investigate how it will affect the numbers of GSFs and operational impact of substituting them. Three average speed values are assumed for ground transport modes: (VII) 40 miles per hour, (VIII) 50 miles per hour as in previous analyses, and (IX) 100 miles per hour (this assumes that high-speed rail is the ground transport substitute). Arrivals of pseudo flights substituted with ground transport modes with these three different speeds are shown in Figure 3-6. It is obvious that higher speed leads to a higher number of GSFs.

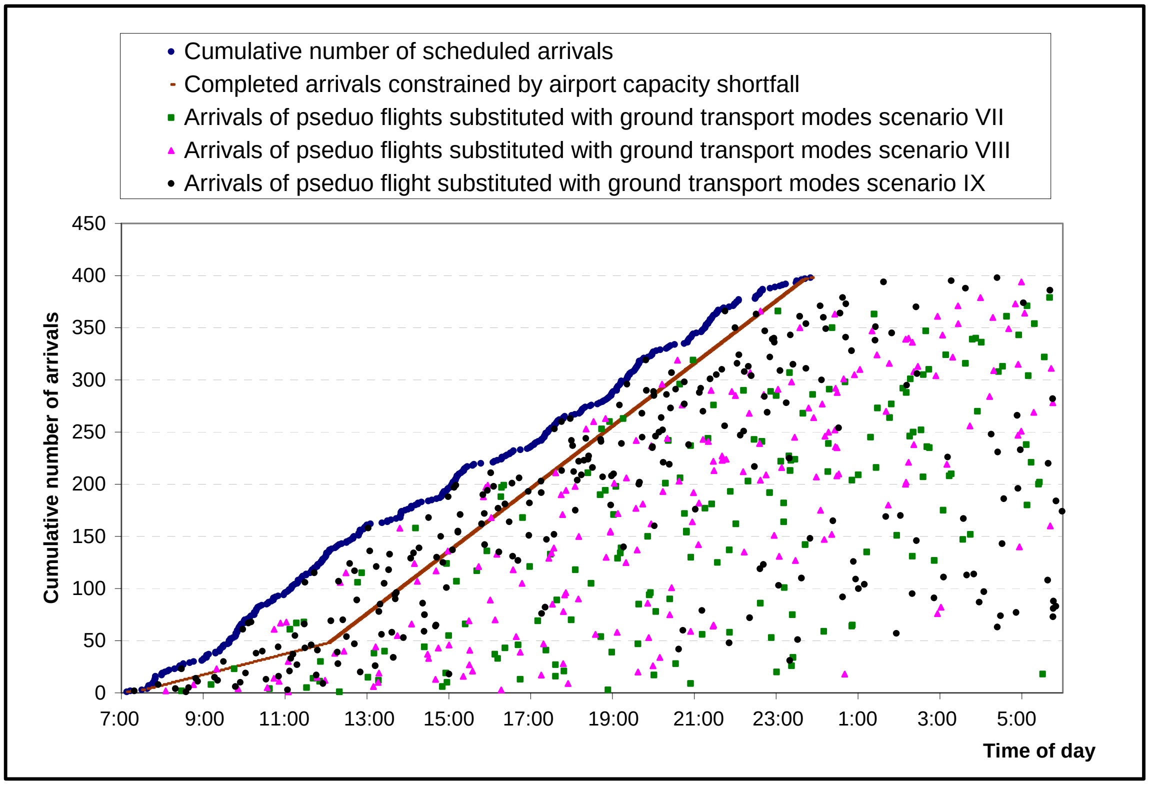

Figure 3- 6 Time Saving of Using Ground Transport Modes during Disruptions Caused by Capacity Shortfalls – Ground Transportation with Various Speeds

For Scenarios VII, VIII, and IX, without substitutions, the longest flight delay is 2.92 hours, as shown in Table 3-3. If GSFs are substituted with ground transportation and removed from the sequence of scheduled flights, the longest flight delays are 2.75, 2.70, and 2.35 hours for Scenarios VII, VIII, and IX, respectively. Although the differences of longest flight delays are insignificant, the total delay saving is substantial, especially when high-speed trains are options for ground transportation substitutions. The saving can be as high as 62% of the total flight delay without substitutions.

Table 3- 3 Comparisons of with and without Substitutions during Disruptions Caused by Capacity Shortfalls – Ground Transportation with Various Speeds

	Scenario VII	Scenario VIII	Scenario IX
Number of Flights Saving Time from being Substituted under Disruptions	12	17	64
Longest Flight Delay without Substitutions (hrs)	2.92	2.92	2.92
Longest Flight Delay with Substitutions (hrs)	2.75	2.70	2.35
Increase/ (Decrease)	(0.17)	(0.22)	(0.57)
Percentage of Increase or Decrease	-5.71%	-7.43%	-19.43%
Total Flight Delay without Substitutions (flight.hrs)	683.52	683.52	683.52
Total Flight Delay with Substitutions (flight.hrs)	569.32	533.72	257.1
Increase/ (Decrease)	(114.20)	(149.80)	(426.42)
Percentage of Increase or Decrease	-16.71%	-21.92%	-62.39%

This section investigated how the severity and starting time of an airport capacity shortfall affect the numbers of GSFs—flights that could reach their destination airport sooner if they were shifted to ground transport modes. In addition, the effect of using substitute modes with different ground transportation speeds was also investigated. In summary, inter-modal substitution is more effective in reducing delays when the capacity shortfall is more severe, when the shortfall starts earlier in the day, and when the ground transport mode is faster. Also, inter-modal substitution saves time for the passengers on flights that are substituted, as well as for passengers on un-substituted flights that can be moved up in the arrival sequence.

The above analysis is conducted only for an arrival sequence at SFO. When airlines make decisions to delay, cancel, and substitute the flights, they have to consider the connections between arrivals and departures, passenger reassignment, aircraft usage, network effect, and other constraints. Their decisions cannot be made only on the basis of a deterministic queuing analysis. Hence, the next section will propose a mathematical programming model to help a single airline to make its decisions during disruptions.

3.3. Mathematical Programming Model for Implementing RTIMS

By introducing ground transportation modes, an airline's hub-and-spoke network becomes a network with two sets of links. One set corresponds to flights between spoke airports and the hub airport, with short travel times but subject to delays caused by the insufficient capacity at the hub airport. The other set of links represent surface transport—usually by road—between spoke airports and the hub; these have longer transportation time but no airside delays because they do not require scarce airside capacity. As illustrated in Figure 3-7, the straight lines indicate links with short travel times but possible delays, while curvy lines indicate links with longer travel times but not subject to airside delays. The enlarged part on the right highlights spoke airports within a certain distance from the hub airport for which air service could be substituted with motor coach (or other road vehicle) service because of the relatively small differences between flying times and ground transportation times. As mentioned in the introduction, questions that airlines need to answer are as follows:

1. Which flight legs should be delayed, cancelled, or substituted with other transportation modes?

2. If a flight is not cancelled or substituted, how long it should be delayed?
3. How should passengers on cancelled flights be reassigned to other flights or ground transport services?
4. How should the ground transport services be scheduled?

These decisions are subject to constraints such as aircraft capacities, passenger flow conservations, aircraft availabilities, and airport capacities. To effectively make these decisions, it is necessary for managers to have knowledge of the predicted airport capacity profile (or assigned slots from the FAA GDP), and passenger itinerary information. Ideally decisions about recovering operations in the aftermath of a disruption should be made within a comprehensive framework. Currently, decisions related to scheduling, aircraft, crews, and passengers are made sequentially. Due to the complex interactions among these decisions and the limitations of operations research technology for solving large-scale nonlinear integer programming models, there is no model that successfully integrates all components. Encouragingly, although some assumptions still have to be made to simplify the problem, the research community is moving toward integrating two or more components in order to attain solutions that approach global optima. This study will focus on schedule and passenger recovery in disruptions within the proposed inter-modal framework, considering both arrival and departure capacity constraints. In this section, a mathematical programming model is constructed to help airlines answer the above questions.

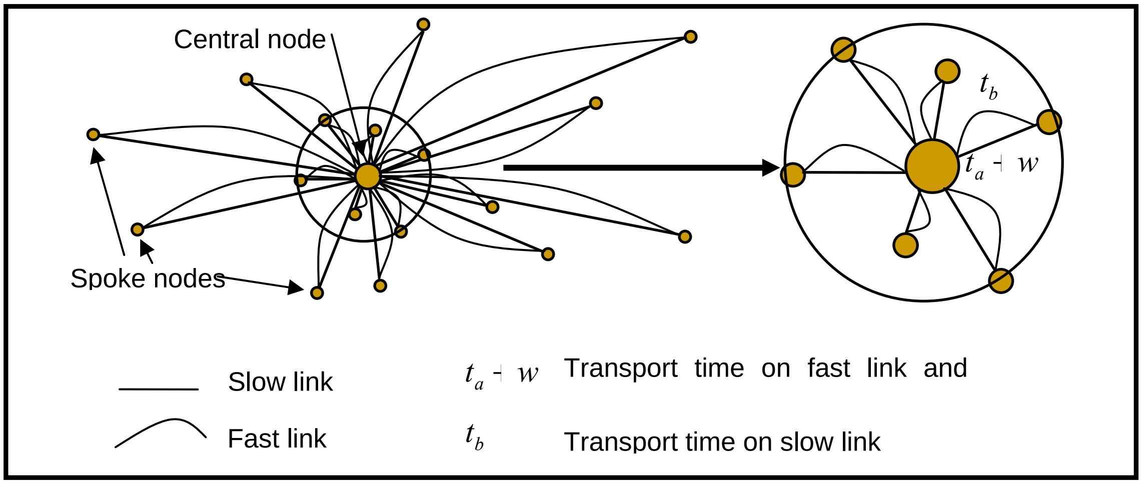

Figure 3- 7 Network with A Node Capacity Constraint

3.3.1. Notation

Table 3-4 lists all notations used in the mathematical programming model sorted in alphabetic sequence. Notations are categorized into four types: Input and Parameter as given, Decision Variables, and Intermediate Variables calculated based on decision variables and given information for the purpose of simplifying formulations.

Table 3- 4 Notation for RTIMS Model

Notation	**Category**	**Description**
$a \in A$	Index	Arrival flights
AAT_a	Input	Gate to gate time of arrival flight *a*
ADT_f	Input	Gate to gate times of departure flight *f*
AD_a	Input	Delay of arrival flight *a*
$AHPax_a$	Input	Number of local (i.e. not connecting) passengers

		originally booked on arrival flight a
$AHCap_t$	Input	Arrival capacity at the hub airport in time period t
$APax_a$	Input	Total number of passengers booked on arrival flight a
AT_a	Input	Scheduled arrival time of arrival flight a
BAT_a	Input	Transportation time if arrival flight a is substituted by ground transportation mode
BDT_f	Input	Transportation time if departure flight a is substituted by ground transportation mode
BO	Intermediate Variable	Total operating cost for ground transport mode
$CostB$	Parameter	Fixed cost of utilizing ground transportation per flight substitution
$CostBP$	Parameter	Variable cost of utilizing ground transportation per passenger unit time
$CostD$	Parameter	Estimated cost per disrupted passenger
$CostF$	Parameter	Airline operating cost of delaying a flight for one time unit
$CostP$	Parameter	Passenger delay cost per one time unit
DD_f	Intermediate Variable	Delay of departure flight f
$DHCap_t$	Input	Departure capacity at the hub airport in time period t
$DHPax_f$	Input	Number of local (i.e. not connecting) passengers

		originally booked on departure flight f
DP	Intermediate Variable	Number of disrupted passengers
$DPax_f$	Input	Total number of passengers booked on departure flight f
DT_f	Input	Scheduled departure time of flight f
$f \in \Phi$	Index	Departure flights
$IAircraft_k^{\tau}$	Intermediate Variable	Cumulative number of type k aircraft at the hub through time period τ
$IPax_s^{\tau}$	Intermediate Variable	Cumulative number of inbound transfer passengers with destination s arriving at the hub through time period τ
$k \in \mathrm{K}$	Index	Aircraft type
$maircraft$	Parameter	Minimum turnaround time of flights
$mpax$	Parameter	Minimum connecting time of transfer passengers at the hub airport
$OAircraft_k^{\tau}$	Intermediate Variable	Cumulative number of type k aircraft that have departed the hub through time period τ
$OPax_s^{\tau}$	Intermediate Variable	Cumulative number of outbound transfer passengers with destination s departed the hub through time period τ
$Pax_{a,f}$	Input	Number of passengers booked to transfer from arrival

		flight *a* to departure flight *f*
P_f	Decision Variable	The number of passengers on departure flight *f*
$s \in$ Є	Index	Spoke airports
$SpokeA_{as}$	Input	Indicators of origins of arrival flights (equals 1 if arrival flight *a* originated from spoke *s*, 0 otherwise).
$SpokeF_{fs}$	Input	Indicators of destinations of departure flights (equals 1 if departure flight *f* goes to spoke *s*, 0 otherwise).
$t \in \Gamma$	Input	A set of discrete time periods
$TypeA_{ak}$	Input	Indicators of aircraft type of arrival flights (equals 1 if arrival flight *a* uses aircraft type *k*, 0 otherwise).
$TypeF_{fk}$	Input	Indicators of aircraft type of departure flights (equals 1 if departure flight *f* uses aircraft type *k*, 0 otherwise).
$TypeO_k$	Input	Number of type *k* aircraft available at the hub airport at the beginning of the schedule perturbation
$TypeT_k$	Input	Number of type *k* aircraft required to be at the hub airport at the end of schedule perturbation
xf_a^t	Decision variable	Equals 1 if an arrival flight *a* is planned to arrive at the hub airport during time period *t*, 0 otherwise.
xs_a^t	Decision variable	Equals 1 if arrival flight *a* is substituted with ground transportation and is planned to arrive at the hub airport during time period *t*, 0 otherwise.

yf_f^t	Decision variable	Equals 1 if departure flight *f* is planned to take-off from the hub airport during time period *t*, 0 otherwise.
ys_f^t	Decision variable	Equals 1 if departure flight *f* is substituted with ground transportation and is planned to leave from the hub airport during time period *t*, 0 otherwise.

The time of day is divided into a finite set of time periods of equal duration, which are denoted by $t \in \Gamma = \{1..T\}$. For instance, Γ might be a set of 60 time periods of 10 minutes each, spanning to a planning horizon of 10 hours. The numbers of arrival and departure slots assigned to an airline in each time period *t* are denoted as $AHCap_t$ and $DHCap_t$ for $t \in \Gamma$, respectively. Airlines have arrival flights $a \in \mathrm{A} = \{1..A\}$ with the scheduled arrival time denoted by AT_a, and departure flights $f \in \Phi = \{1..F\}$ with the scheduled departure time denoted by DT_f. Gate-to-gate times of these flights are assumed to be fixed and recorded in two vectors AAT_a and ADT_f, respectively. Spoke airports in this hub-and-spoke network are denoted by a set $s \in Є = \{1..S\}$. $SpokeA$ and $SpokeF$ indicate arrival flights' origins and departure flights' destinations, respectively. If $SpokeA_{as} = 1$, arrival flight a comes from spoke airport s, 0 otherwise. Aircraft type is denoted by $k \in \mathrm{K} = \{1..K\}$. If $TypeA_{ak} = 1$, arrival flight a uses a type k aircraft, 0 otherwise. If $TypeF_{f,k} = 1$, departure flight f uses a type k aircraft, 0 otherwise. The number of type k aircraft available at the beginning of schedule perturbation period at the hub airport is denoted as $TypeO_k$. The number of type k aircraft required to be available at the end of perturbation period are indicated as $TypeT_k$.

Passenger itinerary information for this model is listed in Table 3-5. In Table 3-5, the first column is arrival flights and the first row is departure flights. $AHPax_a$ in the second column is the number of passengers on arrival flights a whose destination is the hub airport. $DHPax_f$ in the second row is the number of passengers on a departure flight f originating from the hub airport. Pax_{af} denotes the number of passengers whose itinerary involves a transfer from an arrival flight a to a departure flight f. In the last column and last row, the total passengers on arrival flights and departure flights are denoted by $APax_a$ and $DPax_f$, respectively.

Table 3- 5 Input Data: Passenger Itinerary Information

		F_1	F_2	F_3	…	
		$DHPax_1$	$DHPax_2$	$DHPax_3$	…	
A_1	$AHPax_1$	$Pax_{1,1}$	$Pax_{1,2}$	$Pax_{1,3}$	…	$APax_1$
A_2	$AHPax_2$	$Pax_{2,1}$	$Pax_{2,2}$	$Pax_{2,3}$	…	$APax_2$
A_3	$AHPax_3$	$Pax_{3,1}$	$Pax_{3,2}$	$Pax_{3,3}$	…	$APax_3$
…	…	…	…	…	…	…
		$DPax_1$	$DPax_2$	$DPax_3$	…	

It is assumed that the travel time for each ground transport link connecting a spoke airport with the hub is estimated from historical data and real-time traffic conditions. The time is then recorded into two vectors BAT_a and BDT_f, for $a \in \mathrm{A}$ and $f \in \Phi$. The other two parameters needed for this model are $maircraft$, the minimum turnaround time of flights, and $mpax$, the minimum connecting time of transfer passengers at the hub airport. We assume these values to be 45 minutes and 30 minutes, respectively. These two parameters could be varied according to flight or flight pair if the necessary information is available. The values of the cost parameters for this model are listed in Appendix A.

3.3.2. Decision Variables

The decision variables in the model include a set of binary decision variables for scheduled arrival flights, a set of binary decision variables for departure flights, and a set of integer decision variables for numbers of passengers on departure flights. For an arrival flight *a*, if it lands during time unit *t*, then xf_a^t equals 1, otherwise 0. If it is substituted with ground transport service that will arrive at the hub airport during time period t, xs_a^t equals to 1, otherwise 0. If all the xf_a^t and xs_a^t are set to 0 over the entire planning horizon, the arrival flight is cancelled without substitution. A similar set of decision variables is defined for departure flights. Additionally, a set of passenger decision variables determine the numbers of passengers on departure flights. These decision variables are defined as follows:

$$xf_a^t = \begin{cases} 1 & \text{if arrival flight } a \text{ is planned to arrive at the} \\ & \text{hub airport during time period } t; \\ 0 & \text{otherwise.} \end{cases} \qquad a \in \mathrm{A}$$

$$xs_a^t = \begin{cases} 1 \quad \text{if arrival flight } a \text{ is substituted with ground transportation} \\ \text{and arrives at the hub airport during time period } t \\ \\ 0 \quad \text{otherwise.} \end{cases} \qquad a \in \mathrm{A}$$

$$yf_f^t = \begin{cases} 1 \quad \text{if departure flight } f \text{ is planned to leave the hub airport} \\ \text{during time period } t \\ \\ 0 \quad \text{otherwise.} \end{cases} \qquad f \in \Phi$$

$$ys_f^t = \begin{cases} 1 \quad \text{if departure flight } f \text{ is substituted with ground transporation} \\ \text{and planned to leave the hub airport during time period } t \\ \\ 0 \quad \text{otherwise.} \end{cases} \qquad f \in \Phi$$

$P_f \geq 0$ Integer, the number of passengers on departure flight f $\qquad f \in \Phi$

3.3.3. *Objective Function*

The objective function to be minimized in this model is total disruption costs, including flight delay, cancellation cost, and passenger costs due to flight delay, cancellation, or substitution. In addition, the cost of ground transport, assumed here to be motor coach service, is taken into consideration. The total airline' disruption cost can be expressed by Equation 3.3.1.

$$\begin{aligned} Min \quad & \left(\sum_a AD_a \cdot AHPax_a + \sum_s \sum_t \left(IPax_s^t - OPax_s^t \right) + \sum_f DD_f \cdot P_f \right) \cdot CostP \\ & + DP \cdot CostD + \left(\sum_a AD_a + \sum_f DD_f \right) \cdot CostF + BO \end{aligned} \tag{3.3.1}$$

The first component of the objective function is the passenger delay cost including the delays of arriving passengers whose destination is the hub airport, passengers' waiting time at the hub, and delays of departing passengers. Passengers with the same itinerary are considered identical in the sense that the costs of delay or reassignment are the same, so that costs depend on aggregate numbers of passengers that are delayed or reassigned[15].The second component is the cost of the disrupted passengers, defined as those who missed their original connections and cannot be reassigned to a later flight before the end of the planning horizon. The third component is the flight delay cost, and the last is the ground transportation operating cost. The arrival and departure delays in the objective function are calculated as follows:

$$AD_a = \sum_t (t - AT_a) \cdot (xf_a^t + xs_a^t) + \left(1 - \sum_t (xf_a^t + xs_a^t)\right) \cdot ACanT_a \qquad (3.3.2)$$

$$DD_f = \sum_t (t - DT_f) \cdot (yf_f^t + ys_f^t) + \left(1 - \sum_t (yf_f^t + ys_f^t)\right) \cdot DCanT_f \qquad (3.3.3)$$

where $ACanT_a$ and $DCanT_f$ are estimated propagated delay hours if flights are cancelled. This is an approximate way to incorporate flight cancellation cost into airline disruption cost.

Disrupted passengers are passengers who missed their connections and could not get reassigned successfully during the disruption recovery horizon. Passengers who miss their original itinerary but are successfully reassigned to other flights or substitute ground

---

[15] Nevertheless, when airlines implement this strategy, they may consider passengers loyalty, value as customers, or other criteria in their reassignment decisions.

transportation, however, are not disrupted passengers (see Figure 3-8). The number of disrupted passengers is:

$$DP = \sum_a \left(1 - \sum_t \left(xf_a^t + xs_a^t \right) \right) \cdot APax_a + \sum_s \left(IPax_s^T - OPax_s^T \right) \tag{3.3.4}$$

$IPax_s^T$ and $OPax_s^T$ are total inbound and outbound transfer passengers to a destination s. Calculation of these numbers is elaborated in Section 3.3.4. The operating cost of a bus includes fixed cost per motor coach and a variable cost that is proportional to passenger times for those who are reassigned to ground transportation. It is calculated as:

$$\begin{aligned} BO = & \left(\sum_a APax_a \cdot ABT_a \cdot \sum_t xs_a^t + \sum_f P_f \cdot DBT_f \cdot \sum_t ys_f^t \right) \cdot CostBP \\ & + \left(\sum_a \sum_t xs_a^t + \sum_f \sum_t ys_f^t \right) \cdot CostB \end{aligned} \tag{3.3.5}$$

where $CostBP$ is the cost per passenger-hour spent on the motor coach; $CostB$ is the cost per motor coach utilized.

3.3.4. Constraints

The first set of constraints includes arrival and departure capacity constraints at the hub airport.

$$\sum_a xf_a^t \le AHCap_t \qquad \forall t \in \Gamma \tag{3.3.6}$$

$$\sum_f yf_f^t \le DHCap_t \qquad \forall t \in \Gamma \tag{3.3.7}$$

Besides being delayed or substituted, flights can also be cancelled. Therefore the summation of flying and substitution decision variables through the entire planning horizon should be less than 1. This is presented by the following set of constraints.

$$\sum_{t} \left(xf_{a}^{t} + xs_{a}^{t} \right) <= 1 \qquad \forall a \in \mathrm{A} \tag{3.3.8}$$

$$\sum_{t} \left(yf_{f}^{t} + ys_{f}^{t} \right) <= 1 \qquad \forall f \in \Phi \tag{3.3.9}$$

The next set of constraints conserves passenger flow. The outbound transfer passengers to a destination s leaving at time t should be no more than the inbound transfer passengers to the same destination arriving at the hub airport up to $t - mpax$, where $mpax$ is the minimum passenger connection time (as shown in Figure 3-8), so that:

$$IPax_{s}^{\tau} \geq OPax_{s}^{\tau + mpax} \qquad \forall s \in \Theta \quad \forall \tau \in \{1..(T - mpax)\} \tag{3.3.10}$$

where:

$$IPax_{s}^{\tau} = \sum_{t=1}^{\tau} \sum_{a} \sum_{f} \left(xf_{a}^{t} + xs_{a}^{t} \right) \cdot Pax_{af} \cdot SpokeF_{fs}$$

$$OPax_{s}^{\tau} = \sum_{t=1}^{\tau} \sum_{f} \left(yf_{f}^{t} + ys_{f}^{t} \right) \cdot \left(P_{f} - DHPax_{f} \right) \cdot SpokeF_{fs}$$

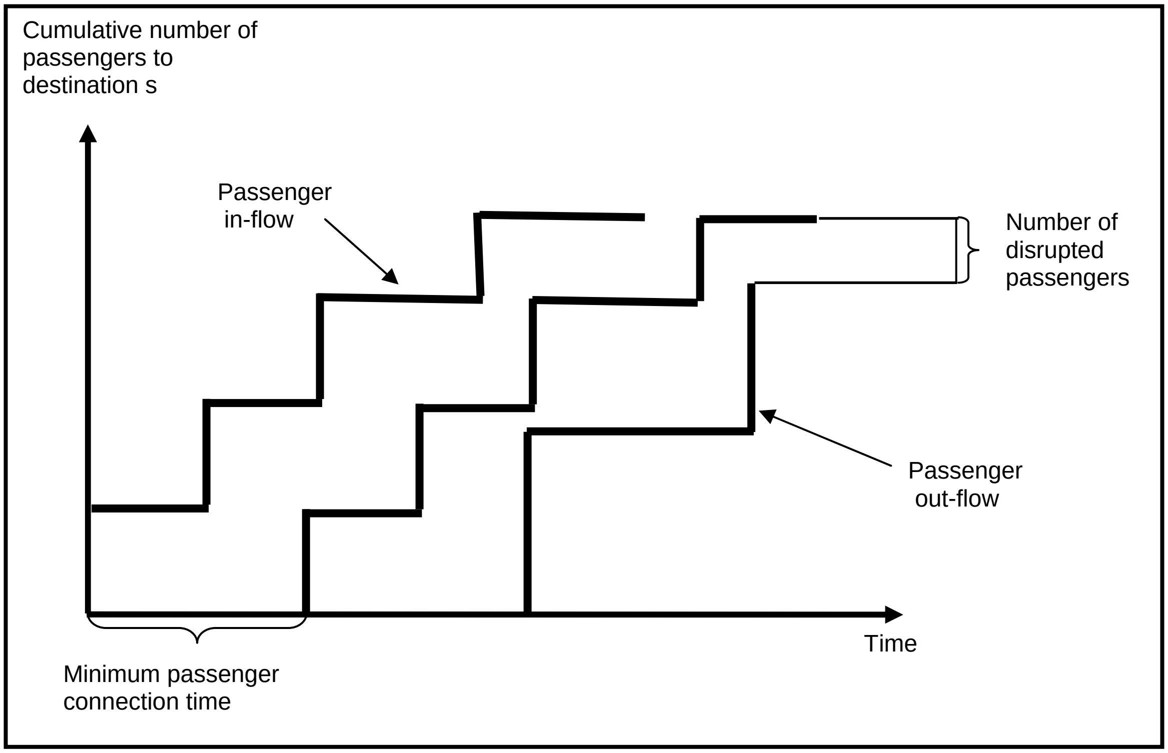

Figure 3- 8 Passenger Conservation and Disrupted Passengers

Another set of critical constraints are the aircraft flow conservations. Aircraft swapping is allowed among flights that utilize the same type of aircraft.

$$TypeO_k + IAircraft_k^{\tau} \geq OAircraft_k^{\tau+maircraft} \qquad \forall k \in \mathrm{K} \;\; \forall \tau \in \{1..(T - maircraft)\} \tag{3.3.11}$$

$$TypeO_k + IAircraft_k^{T} \geq OAircraft_k^{T} + TypeT_k \qquad \forall k \in \mathrm{K} \tag{3.3.12}$$

where:

$$IAircraft_k^{\tau} = \sum_{t=1}^{\tau} \sum_{a} xf_a^{t} \cdot TypeA_{ak}$$

$$OAircraft_k^{\tau} = \sum_{t=1}^{\tau} \sum_{f} yf_f^{t} \cdot TypeD_{fk}$$

In addition, motor coaches serving as pseudo flights are not allowed to be released from their origins earlier than the scheduled departure times of the original flights. For an arrival flight from spoke airports, there are two constraints, depending on whether that flight is substituted or not. The constraints are as follows:

$$\sum_{t} t \cdot xf_{a}^{t} \geq AT_{a} \cdot \sum_{t} xf_{a}^{t} \qquad \forall a \in \mathrm{A} \tag{3.3.13}$$

$$\sum_{t} t \cdot xs_{a}^{t} \geq \left(AT_{a} + \left(BAT_{a} - AAT_{a}\right)\right) \cdot \sum_{t} xs_{a}^{t} \qquad \forall a \in \mathrm{A} \tag{3.3.14}$$

In comparison, there is only one constraint for departure flights at the hub airport.

$$\sum_{t} t \cdot \left(yf_{f}^{t} + ys_{f}^{t}\right) \geq DT_{f} \cdot \sum_{t} \left(yf_{f}^{t} + ys_{f}^{t}\right) \qquad \forall f \in \Phi \tag{3.3.15}$$

A departure flight at the hub airport can be held to wait for transfer passengers. The cost of doing this is to sacrifice the time of passengers and crews who have been ready for departures. The number of passengers on a departure flight f, P_f, has to be less than the seating capacity of an aircraft used by flight f, $DCap_f$.

$$P_{f} \leq DCap_{f} \qquad \forall f \in \Phi \tag{3.3.16}$$

Airlines' operation is complicated due to the interaction and feasibility constraints of different resources. There are four main constraints that affect the feasibility of airline planning and airline disruption management, including aircraft maintenance checks, pilot work rules, fleet assignment, and passenger accommodation. Nevertheless, there is no model so far can solve all constraints at once. Our proposed model does not take aircraft maintenance checks and pilot work rules into consideration. This is because these constraints are less critical for disruption management than for planning. For instance, aircraft maintenance check is required once a week, once a month, and once a year or even longer for different aircraft type. It is important to consider this schedule for long-term airline planning, while it is somewhat less critical for short-term disruption management, especially when aircraft swapping is allowed in implementing real-time inter-modalism. FAA regulation Part 135 defines pilot work rules, listing maximum hours that they can serve within different time ranges. Disruptions may cause pilots to be ineligible to continue their scheduled tasks. In the proposed strategy, pilots, like passengers, can be transported via ground modes as reassigned passengers so that they can be centrally reassigned at the hub airport with consideration of the work rules, seniorities, and abilities.

3.4. Complexity and Solution Algorithm

The model presented in Sections 3.3 is a non-linear integer programming model with linear constraints but a non-linear objective function. The number of decision variables is $2\cdot\left(A+F\right)\cdot T+F$, and the number of constraints is $4\cdot(A+F)+T\cdot(S+3K+2)+(S+K+2)$, where A and F are the number of scheduled arrival and departure flights, respectively; S is the number of spoke airports; K is the number of aircraft types; and T is the number of time periods. Suppose the unit time is 5 minutes, an 18-hour daily schedule falls into 216 time units; a hub-and-spoke network with 20 spoke airports, 200 of daily arrivals and 200 daily departures, and there are 6 types of aircraft, then there are 173,000 decision variables and 10,268 constraints. An automatic pre-solve process reduces a model to its essential core, which can dramatically reduces the optimization time. Based on constraints (3.3.13), (3.3.14), and (3.3.15), a reasonable number of rows and columns of decision variables are reduced, but the problem will still take extensive computation time to solve.

Steps to further simplify the problem must therefore be taken. The first technique is to scrutinize the airlines schedule structure and partition the time horizon into serveal time windows. Fortunately, airlines usually apply banked scheduling strategies in hub airports to reduce the transfer time (time between connections), so it is possible to partition a daily schedule into several time windows within which a set of arrival and departure flights finish their operations. Previous research [Karow and Clarke 2002] set a time window of two hours. However, airlines have smoothed their schedules to reduce operating cost, as a response to the difficult financial environment in the aftermath of 9/11. Although the banking structure still exists, time windows should be elongated [Zhang et al. 2004].

Integer programming problems with a large number of variables are difficult to solve in general. Most successful approaches to solving nonlinear integer problems have involved linear approximation and relaxation techniques. In this study, the traditional relaxation techniques are used. An approximation algorithm for this programming is described as follows.

- First, the integrality of decision variables is relaxed and a non-linear programming problem is solved.

- Second, with solutions from the first step, the flight decision variables xf_a^t, xs_a^t, yf_f^t, ys_f^t are adjusted to 1 if the solution value is larger than 0.5; otherwise, they are set to 0.
- Third, the constraints (3.3.6) and (3.3.7), hub airport arrival and departure capacity constraints, are checked. If the constraints were violated, the arrival or departures are postponed according to "first scheduled first served".
- Finally, the model is rerun with the integer requirement of departure passenger decision variables retained and decision variables obtained from above steps.

The MINLP solver on NEOS Server 4.0[16] is used to perform the above tasks. The design, implementation, and more details of the NEOS Server were discussed by Cayayk et al. [1998], Gropp et al. [1997], and Dolan [2001].

3.5. Experimental Example

Numerical experiments are conducted to explore the performance of the above mathematical model. It is assumed that airline A is operating a hub-and-spoke network. There are four short-haul spoke airports, each with the same distance to the hub airport (one hour gate to gate by flying or three hour driving with ground transportation), and eight long-haul spoke airports that are also equidistant from the hub (five hours gate to gate) (see Figure 3-9).

---

[16] http://neos.mcs.anl.gov/neos/, accessed on July 31, 2007.

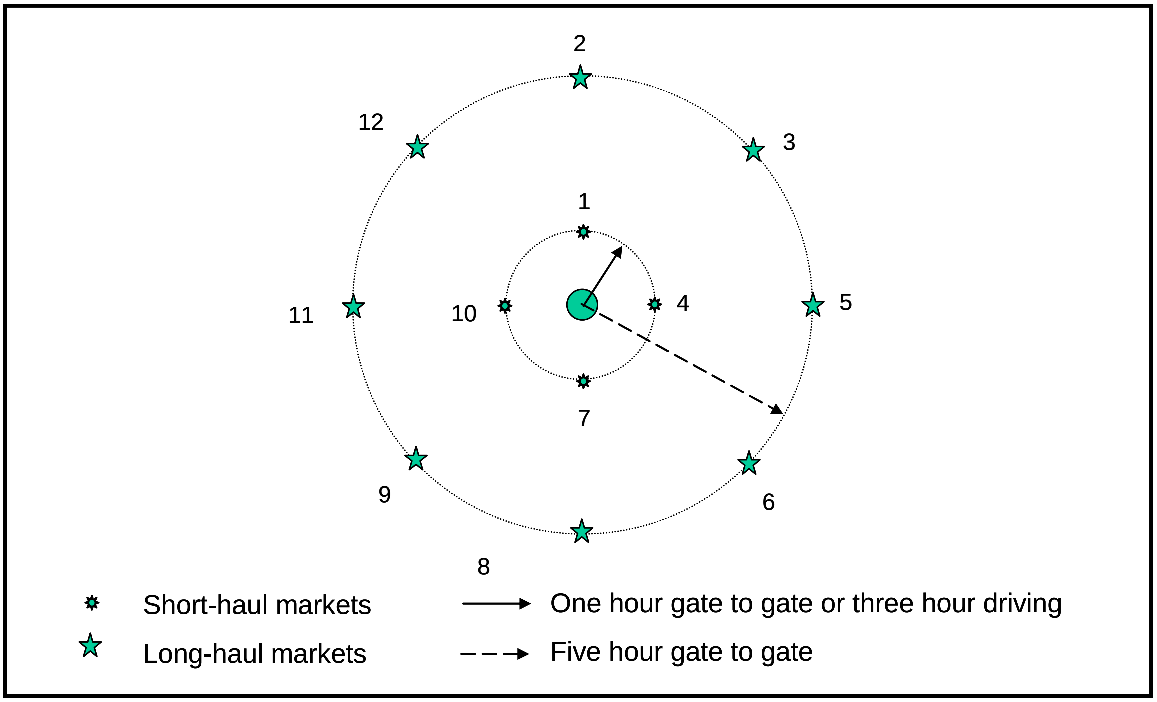

Figure 3- 9 Depiction of Network for Experimental Example

Twenty-four arrivals and twenty-four departures are scheduled and evenly distributed in a four-hour time window following the sequence of one short-haul, two long-hauls, then one short-haul, two long-hauls, and so forth. Given the numbers of spoke airports in Figure 3-9, the sequence of arrivals from spoke airports and departures to the spoke airports is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 as a first bank, and then 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 as a second bank, as shown in Figure 3-10. Thus one third of the flights are short-haul flights. Every spoke airport has two arrivals and two departures.

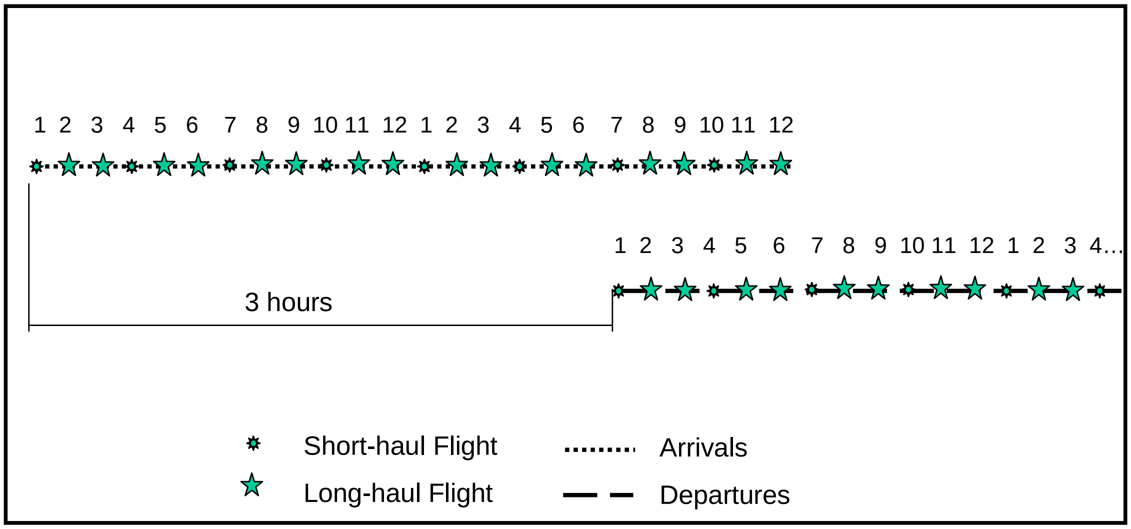

Figure 3- 10 Schedules of Arrival and Departure Flights

It is assumed that there are no passengers transferring from an arrival flight from one spoke airport to a departure flight to the same spoke airport, i.e. the diagonal elements of the transfer passenger matrix are all zeros. Other elements in the matrix are simplified by assuming that (1) all passenger transfers are between arriving and departing flights within the same bank; (2) the number of passengers transferring within a bank is determined by the category (short-haul or long-haul) of the arriving and departing flight. It is assumed that there are only two types of aircraft: large aircraft used for long-haul markets and small aircraft used for short-haul markets. The capacity of the hub airport drops to one third of the normal level for five hours. Consequently, the number of landing and take-off slots allocated to airline A is two each for the first five hours and six each afterwards.

3.5.1. Experimental Results

Table 3-6 shows the results of the numerical example. The “w/o Substitution” column lists the results for the scenario in which there is no ground transportation substitution for cancelled flights. The results are obtained by solving the RTMIS model with the approximation algorithm but with the substitution decision variables are forced to zero. The columns under the heading “with Substitution” are three types of results for the scenario in which there is ground transportation substitution: results from the approximation algorithm, results from exact solution, and a lower bound without any binary or integer constraints. The exact solution requires much more computation time than the approximation algorithm. It can be seen that for the scenario without substitution, there are three cancellations for inbound flights and five out of twenty-four departure flights are cancelled. About 15 percent of passengers are left behind as disrupted. The longest flight delay under this scenario is 6.5 hours. On the other hand, if there is ground substitution for cancelled flights, the number of flight cancellations increases and most of the cancellations are substituted with ground transport modes. The reduction of the objective function value in the “with substitution” scenario is mainly due to the decrease of flight delay and number of disrupted passengers.

Table 3- 6 Optimization Results for Numerical Experiment

	without Substitution	with Substitution		
		Approximation	Exact solution	Lower Bound
Total Cost ($)	885,720	514,930	495,125	463,523
Total Arrivals and Departures	48	48	48	
Inbound Cancellations	3	8	8	
Substitutions	-	8	8	
Outbound Cancellations	5	7	7	
Substitutions	-	5	5	
Longest Flight Delay (hours)	6.5	4.7	4.5	
Total Passengers	3,808	3,808	3,808	
Disrupted Passengers	560	339	339	
Percentage of Disrupted Pax	14.7%	8.9%	8.9%	
Computation Time (seconds)	10^4	10^2	10^4	10^2

Assuming there is no inter-modal substitution option for the airlines, and with the same cost coefficients and formulation, airlines' total cost is to $886 thousand. Optimization with relaxation of the integrality of decision variables provides a lower-bound for the problem, which is $464 thousand. With the solution algorithm described in Section 3.4, fractional decision variables are adjusted to binary variables for flight decision variables. Given adjusted flight decision variables as inputs and by retaining the integrality of the passenger decision variable, a set of new passenger numbers on departure flights is obtained, with the objective function value of $515 thousand, about ten percent higher than the lower-bound. The exact result from the non-linear integer programming falls between the lower-bound and the result of approximation algorithm. With inter-modal substitution, the number of disrupted passengers decreases from 560, about 15 percent of total passengers, to 339, about 9 percent.

The delay impacts of inter-modal substitution and cancellations on arrival flights are shown in Figure 3-11. The horizontal axis in the figure is the time of day, and the vertical axis is the cumulated number of arrival flights. The far left curve depicts the cumulated number of flights arriving at the airport following their original schedules. The curve for hypothetical cumulative arrivals constrained by airport capacity—assuming all flights are flown—is on the right. The dashed line on the left is the cumulative number of arrivals after removing cancelled flights, while the dotted line on the right is what happens during the disruption when flights are cancelled, reordered, or substituted by ground transport modes. It shows that eight arrival flights have been canceled. A shaded area in the figure indicates flight delay saving with RTIMS in comparison to the scenario in which no flights are cancelled. A similar plot was created for departure flights, as shown in Figure 3-12. In comparison, delay saving of arrival flights is much more than that of departure flights. This is because there is three hours lag between an arrival bank and a departure bank, so that for departure flights the capacity shortfall occurs later and is of shorter duration.

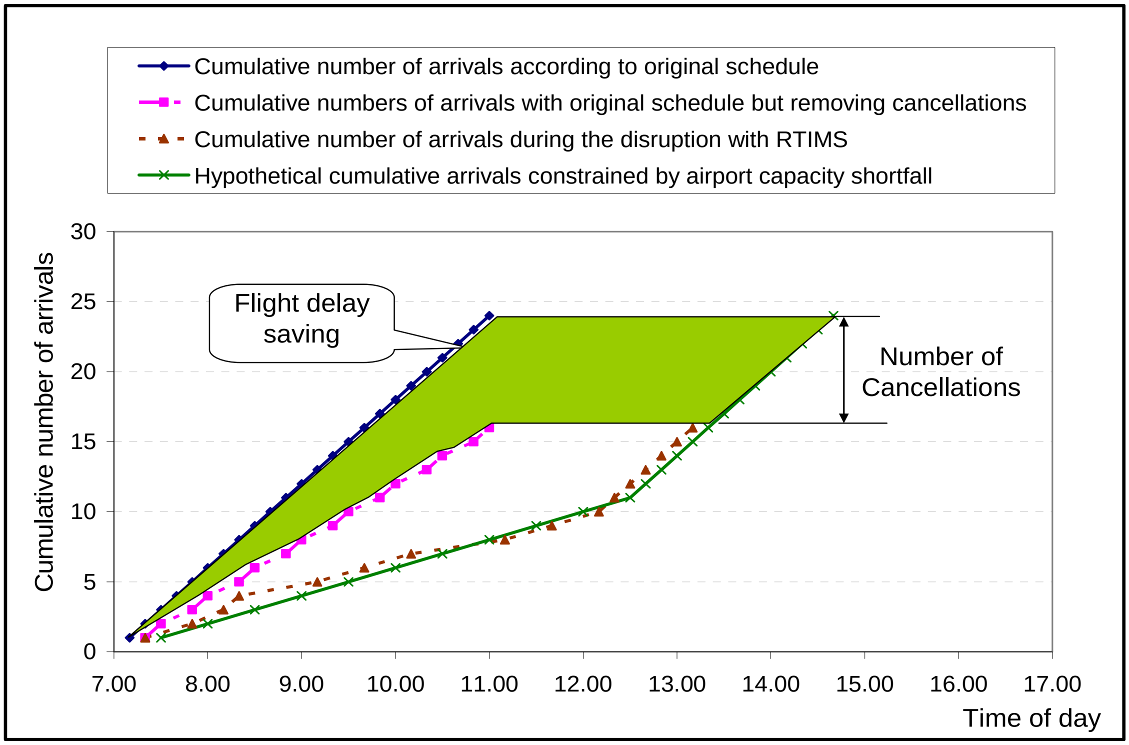

Figure 3- 11 Flight Delay Saving of Arrival Flights

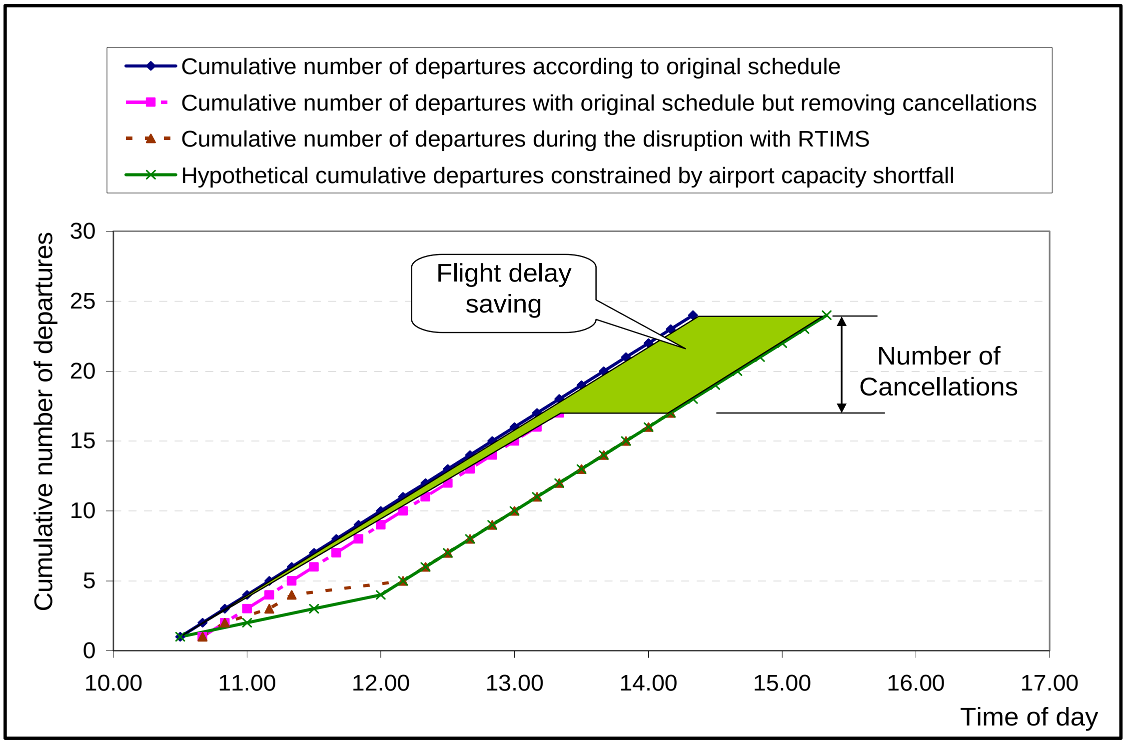

Figure 3- 12 Flight Delay Saving of Departing Flights

3.6. Sensitivity Analysis

We now investigate the sensitivity of our results to key features of the problem. We focus our attention on severity of capacity shortfalls, passenger value of time, distances of short-haul spoke airports, load factor of the flights, schedule peaking, and connecting patterns of transfer passengers.

3.6.1. Severity of Capacity Shortfall

In the base case, the hub airport capacity drops to one third for the first five hours and resumes to a normal level afterwards. Two more capacity shortfall cases are studied for the sensitivity analysis. In the low severity case, the capacity shortfall—which similarly to the baseline is one third the normal capacity—only lasts for three hours, while in the high severity case, the airport is completely shut down for five hours before it is reopened with a normal capacity.

The results for the three capacity shortfall cases are shown in Table 3-10. As expected, disruption cost increases with the severity of capacity shortfall. Airport closure generates a disruption cost three times that of the base case. In contrast, the less severe capacity shortfall avoids a large number of cancellations and only one ground substitution is needed. Percentage difference from total cost without substitution, as shown in the last row in Table 3-7, increases from 13 percent for less severe case to 44 percent for more severe case with airport closure. This analysis verifies that the proposed strategy, RTIMS, is more suitable when the capacity shortfall is estimated to be more severe, in terms of length and magnitude. However, the number of inter-modal substitutions changes very little between the Base and More Severe cases. This is because virtually all short-haul flights have been substituted in the Base case.

Table 3- 7 Sensitivity Analysis – Severity of Capacity Shortfall

	with Substitution		
	Less Severe	Base	More Severe
Total Cost ($)	409,603	514,930	1,696,750
Total Arrivals and Departures	48	48	48
Inbound Cancellations	-	8	18
Substitutions	-	8	8
Outbound Cancellations	2	7	16
Substitutions	1	5	6
Longest Flight Delay (hours)	4.7	4.7	7
Total Passengers	3,808	3,808	3,808
Disrupted Passengers	250	339	2,272
Percentage of Disrupted Pax	6.6%	8.9%	59.7%
% Diff. from w/o Substitution	13.1%	41.9%	44.2%

3.6.2. Passenger Value of Time

Keeping other coefficients the same, passenger value of time is changed to 1.5 times and 2 times of its value in the base case. Results from optimization show that the total disruption cost increases (see Table 3-8). The number of disrupted passengers increases from 339 to 940 and 1,130, respectively. This test demonstrates the trade-off between reducing the delay cost of undisrupted passengers and increasing the number of disrupted ones. With higher value of time, disrupted passengers, a relatively small percentage of total passengers whose unit cost retains its baseline value, are weighted less in airlines' disruption recovery consideration. Nevertheless, the percentage of cost saving from

RTIMS in comparison to without substitution varies only slightly across the scenarios with different passenger value of time.

Table 3- 8 Sensitivity Analysis – Passengers' Value of Time

	with Substitution		
	Base 1.0 CostP	Higher VOT 1.5 CostP	Highest VOT 2.0 CostP
Total Cost ($)	514,930	1,198,240	1,300,390
Total Arrivals and Departures	48	48	48
Inbound Cancellations	8	7	6
Substitutions	8	5	4
Outbound Cancellations	7	10	10
Substitutions	5	2	2
Longest Flight Delay (hours)	4.7	4.7	3.8
Total Passengers	3,808	3,808	3,808
Disrupted Passengers	339	940	1,130
Percentage of Disrupted Pax	8.9%	24.7%	29.7%
% Diff. from w/o Substitution	41.9%	37.9%	37.1%

3.6.3. Distance of Short-haul Spoke Airport

One critical condition for implementing the RTIMS is the presence of short-haul spoke airports where ground transportation can be used to substitute the flights. In the base case, it was assumed that the spoke airports are three-hour driving distance from the hub airport. As shown in Table 3-9, when the distance is shortened to two-hour driving distance (equivalent to if there is high-speed transport mode that shortens the driving time from three hours to two hours), although the disruption cost is lower, the number of cancellations, substitutions, and disrupted passengers remain the same as those of the base case. When the distance increases to four-hour driving, however, not only does the disruption cost increase dramatically, but far fewer flights are cancelled, and of those only one is substituted. In addition, the longest flight delay is seven hours, about 50 percent more than that in the baseline case. As shown in the last row of the Table 3-9, the cost saving resulting from inter-modal substitution drops dramatically from over 40 percent to about 5 percent when the spoke airports are a four-hour diving distance from the hub airport. This demonstrates that the availability of short-haul flights within reasonable driving distance is critical for implementing RTIMS.

Table 3- 9 Sensitivity Analysis – Distances of Short-haul Spoke Airports

	with Substitution		
Driving Time	Shorter Distance 2 hours	Base 3 hours	Longer Distance 4 hours
Total Cost ($)	508,640	514,930	1,441,160
Total Arrivals and Departures	48	48	48
Inbound Cancellations	8	8	3
Substitutions	8	8	-
Outbound Cancellations	7	7	5
Substitutions	5	5	1
Longest Flight Delay (hours)	4.7	4.7	7.0
Total Passengers	3,808	3,808	3,808
Disrupted Passengers	339	339	560
Percentage of Disrupted Pax	8.9%	8.9%	14.7%
% Diff. from w/o Substitution	43.5%	41.9%	5.4%

3.6.4. Load Factor

Higher load factors of flights leave less space for passenger reassignment and lead to more disrupted passengers and longer flight and passenger delays. The sensitivity analysis results, as shown in Table 3-10, indicate that the increase of load factors leads to less cancellations but more disrupted passengers, and more outbound flights are cancelled and substituted by ground transportation. When the load factor is higher, the percentage of cost saving in comparison to the case without substitution is lower.

Table 3- 10 Sensitivity Analysis – Load Factor

	with Substitution		
	Low LF ~65%	Base ~80%	High LF ~90%
Total Cost ($)	504,567	514,930	1,009,523
Total Arrivals and Departures	48	48	48
Inbound Cancellations	8	8	6
Substitutions	8	8	6
Outbound Cancellations	8	7	7
Substitutions	4	5	6
Longest Flight Delay (hours)	4.7	4.7	5.8
Total Passengers	3,808	3,808	3,808
Disrupted Passengers	108	339	486
Percentage of Disrupted Pax	2.8%	8.9%	12.8%
% Diff. from w/o Substitution	45.3%	41.9%	9.1%

3.6.5. Schedule Peaking

It is common to have peaks in a daily flight schedule. As an example, Figure 3-13 shows the arrival schedule of SFO on one Thursday in November 2005. As can be seen, the arrival rate is relatively low from 8:00 am to 9:30 am. It peaks from 9:30 am to 10:30 am, and then goes down from 10:30 am to 12:00 pm. In the early afternoon, the arrival rate reduces to the same level as that at the beginning of the day. For the experiment in Section 3.5, 24 arrival and departure flights are evenly distributed in four hours. In this section, peaks are created by: (1) compressing first 12 arrivals or departures into only one hour and advancing the second 12 flights right after the last flight in compressed bank; (2) compressing 24 arrivals or departures into only two hours. The hourly arrival and departure rates in the peaks are twice the level of normal capacity. As shown in Table 3-11, the change of schedule leads to higher disruption cost and more cancellations. Nevertheless, the percentage of cost saving from RTIMS in comparison to without substitution remains fairly constant across the scenarios with and without schedule peaking.

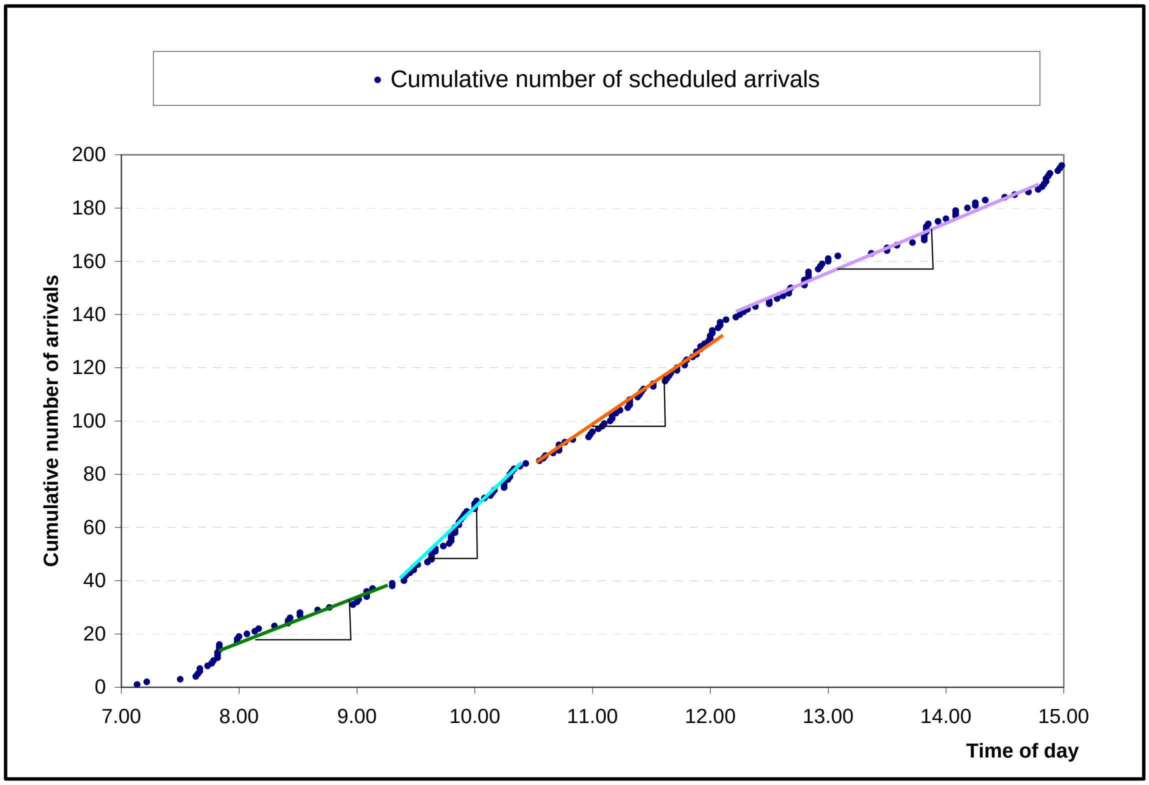

Figure 3- 13 Morning Peak in an Arrival Schedule at a Hub Airport

Table 3- 11 Sensitivity Analysis – Peaking Schedule

	with Substitution		
	Base	Peaking1	Peaking2
Total Cost ($)	514,930	624,729	831,092
Total Arrivals and Departures	48	48	48
Inbound Cancellations	8	8	8
Substitutions	8	8	8
Outbound Cancellations	7	11	11
Substitutions	5	8	8
Longest Flight Delay (hours)	4.7	4.3	4.9
Total Passengers	3,808	3,808	3,808
Disrupted Passengers	339	340	522
Percentage of Disrupted Pax	8.9%	8.9%	13.7%
% Diff. from w/o Substitution	41.9%	40.2%	39.6%

3.6.6. Transfer Passengers' Directional Characteristics

In reality, highly circuitous connection itineraries, for example, between spoke airports 1 and 2 in Figure 3-14 are uncommon. To reflect the directionality of a more typical connecting pattern, the transfer passenger matrix is adjusted by reducing the numbers of transfer passengers with circuitous itineraries while increasing those with low-circuity connections. Nevertheless, the adjustment keeps the total number of passengers on an arrival or a departure flight the same as in the Base case.

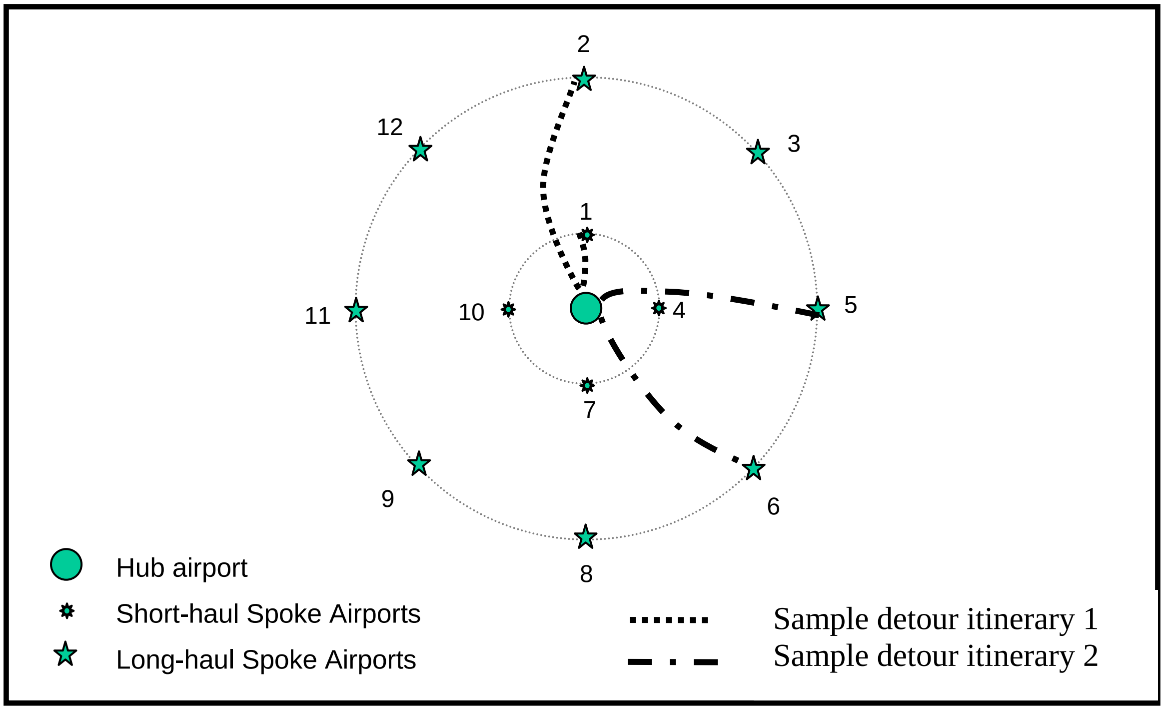

Figure 3- 14 Transfer Passengers' Circuitous Itineraries

The comparison of results is shown in Table 3-12. With the adjusted passenger matrix, the total cost is about 20% higher than in the Base case. There are slightly fewer inbound cancellations but all cancellations are substituted with ground transport service, just as in the baseline case. The number of disrupted passengers increases from about 9 percent of the total to 13 percent; however, the percentage cost savings from inter-modal substitution remain about the same.

Table 3- 12 Sensitivity Analysis – Connecting Pattern Circuity

	With Substitution	
	Baseline Connecting Pattern	Low Curcuity Connecting Pattern
Total Cost ($)	514,930	620,830
Total Arrivals and Departures	48	48
Inbound Cancellations	8	7
Substitutions	8	7
Outbound Cancellations	7	7
Substitutions	5	4
Longest Flight Delay (hours)	4.7	5.7
Total Passengers	3,808	3,808
Disrupted Passengers	339	500
Percentage of Disrupted Pax	8.9%	13.1%
% Diff. from w/o Substitution	41.9%	38.4%

3.7. Summary

In this chapter, we proposed and developed a model for optimizing an inter-modal strategy for airlines recovering from a temporary capacity shortfall at a hub caused by adverse weather or other temporary events. Capacity reduction at a hub airport caused by these temporary events results in implementation of a GDP in which flights are assigned CTAs. We suggest that, when delays resulting from GDPs are severe, airlines with hub-and-spoke networks use ground transportation as a substitute for short-haul flights. This proposed strategy, RTIMS, is different from strategic measures, such as air-rail integration, which are intended to permanently reduce short-haul flight frequency in daily operations. RTIMS is triggered by severe demand supply imbalance at major hub airports and it involves tactical integration of airside and ground transportation. A mathematical programming model is presented to assist airlines to make decisions on whether and how to delay, cancel, or substitute flights with ground transport, which we assume would take the form of motor coach service. The proposed model incorporates passenger reassignment at the hub airport, allowing aircraft to be swapped among flights that are assigned to utilize the same type of aircraft. The model calculates the number of disrupted passengers and accounts for this in the total cost. An approximation algorithm is proposed to get a solution while greatly reducing the substantial computation time required to solve large-scale non-linear integer programming problems.

We experiment with this model under a set of scenarios designed to reveal the sensitivity results to a wide range of factors. The factors considered are the severity of capacity shortfall, passenger value of time, distances of short-haul spoke airports, load factor of the flights, schedule peaking, and passenger itinerary circuity. The following observations are obtained:

- RTIMS yields the greatest reduction in cost compared to a more conventional strategy when the capacity shortfall at a hub airport is severe in terms of magnitude or duration.
- The value of the objective function varies with different passenger value of time.
- The distance between short-haul spoke airports and the hub airport is critical for implementing RTIMS, with a three hour driving time the approximate threshold for inter-modal substitution to be a useful strategy.
- Benefits of implementing RTIMS diminish as load factors increase.
- While flight schedule characteristics and transfer passengers’ connection patterns affect the outcomes of the model, the percentage of cost saving from RTIMS in comparison to a more conventional strategy is not very sensitive to these factors.

In the next chapter, real-time inter-modalism will be extended to a regional airport system, which includes all airports within 50 miles of a reference airport, usually a major hub airport. In addition to ground transport services between spokes and the hub, ground services connecting the hub airport and other airports in the region will be introduced in order to facilitate diversion of flights to other airports in the regional airport system.

CHAPTER 4: REAL-TIME INTER-MODAL DIVERSIONS (RTIMD) IN REGIONAL AIRPORT SYSTEMS

4.1. Problem Description and Motivation

4.1.1. Real-time Inter-modal Diversion for Regional Airport Systems

Bonnefoy and Hansman define a regional airport system as all commercial airports within 50 miles of a reference airport, usually a large hub airport [Bonnefoy and Hansman, 2005]. Figure 4-1 shows the 26 regional airport systems that they have identified. In our study, we propose that when the large hub airport in a regional airport system encounters a capacity shortfall, other airport(s) in the same region could be selected as alternative hub(s). In contrast to a reliever airport, which is assigned to attract general aviation operations from a large and congested airport as a long-term congestion mitigation strategy, an alternative hub would be used tactically, in response to airside capacity shortfalls or closures at major hub airports. The criteria of selecting alternative hubs will be elaborated in Section 4.2.

We propose to incorporate the use of alternative hubs into the real-time inter-modal substitution concept discussed in the previous chapter, so that the capacity available at these airports can be used to maximum advantage when the primary hub has a severe capacity shortage. Similar to the RTIMS strategy analyzed in Chapter3, this new strategy is intended to ameliorate temporary capacity reductions at a major hub through utilization of surface transport. Because it incorporates on the diversion of flights to alternative hubs, we will refer to it as "Real-time Inter-modal Diversion" (RTIMD), with the understanding that the strategy also encompasses RTIMS as described in Chapter 3. in RTIMD, in contrast to present-day operations, aircraft that are diverted to alternative hubs in the same region are not flown back to the major hub airport before they continue on to their original assigned or reassigned destinations. Airlines can reassign transfer passengers so that they can make their connections at the alternative hubs. The alternative hub can also replace the primary one in the itineraries of originating and terminating passengers. Ground transport modes are provided to connect alternative hubs and the major hub airport, complementing existing ground services that connect these airports with other parts of the region.

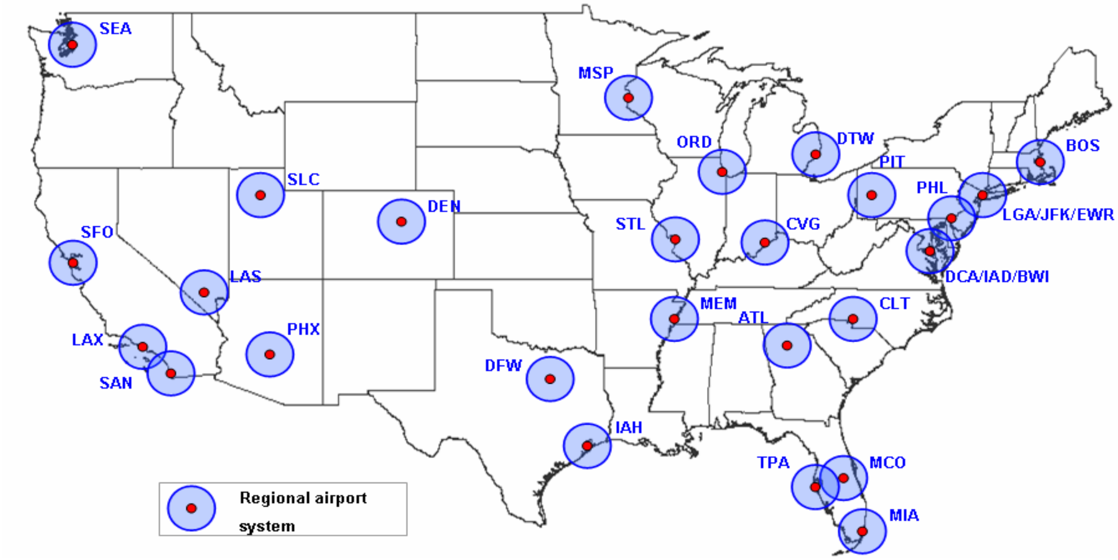

Figure 4- 1 Regional Airport Systems in the Continental US[17]

---

[17] Source: Bonnefoy, P. A., and Hansman, R.J. , *Emergence of Secondary Airports and Dynamics of Regional Airport Systems in the United States*, Report No. ICAT-2005-02, MIT International Center for Air Transportation, May 2005.

A network supporting this new strategy is depicted in Figure 4-2. Solid lines represent air links while dotted lines stand for ground transportation links between hubs, and grey dashed lines denote ground transport links between spokes and the major hub.

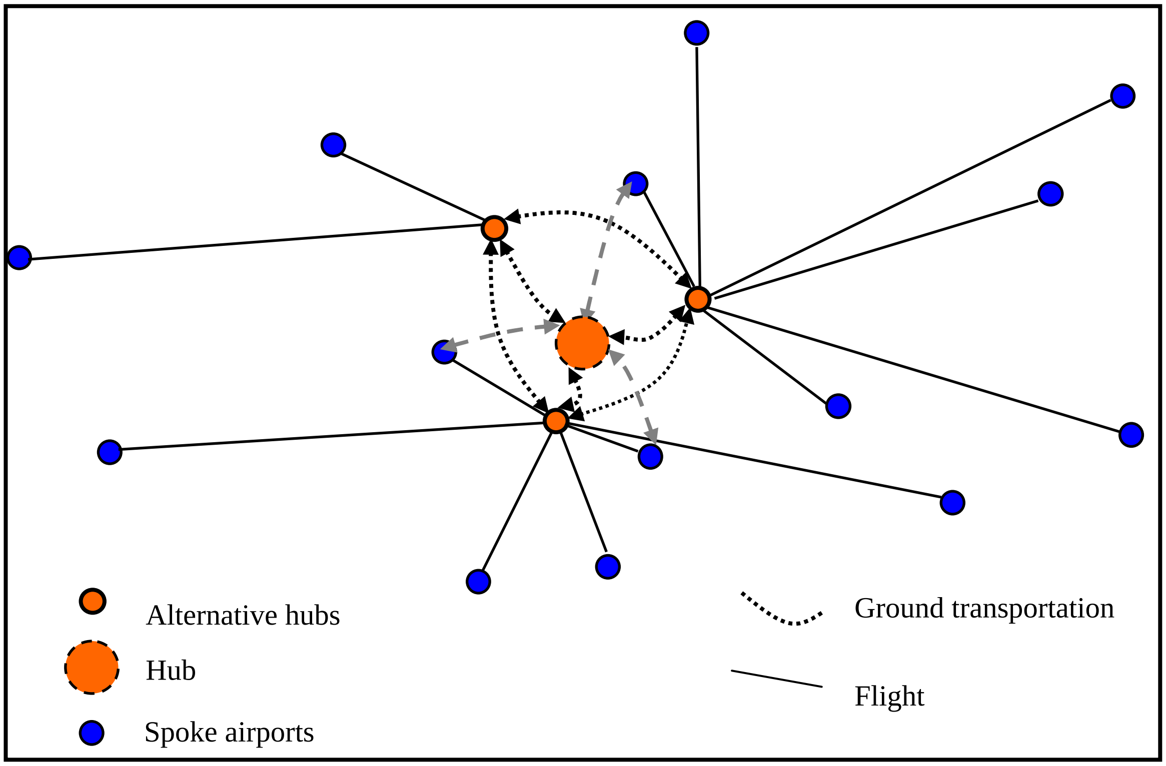

Figure 4- 2 Network for Real-time Inter-modal Diversion (RTIMD)[18]

Our proposed RTIMD strategy, in summary, includes the following components:

- Substitution of short-haul flights with ground transport modes—i.e. Real-time Inter-modal Substitution—for passenger bound for the major hub.
- Diversion of flights to alternative hub(s)

---

[18] This figure shows the network if the original hub airport is complete shut-down. Otherwise, there should be links from spoke airports to the original hub airport, which are omitted here for clarity.

- Transport of passengers between the major hub airport and alternative hubs so that they may access the diverted flights or transfer between flights using different airports as a result of the diversion.

4.1.2. Current Flight Diversion and Time Factors

In the current system, under some circumstances, such as severe airport capacity shortfalls or closures due to weather, equipment outage, or emergencies, flights have to be diverted to alternate airports instead of landing at their original destinations. Once a flight is diverted, it remains at the alternate until clearance is received from Air Traffic Control ensuring that its original destination has enough capacity to allow it to proceed there. Michael Irrgang conducted delay estimations of flight diversions [Irrgang 1995]. He estimated that the total delay caused by diversions ranges from 85 minutes to 2 hours plus the destination airport closure time. This total includes:

- 10 - 20 minutes: Extra flying time to reach the alternate
- Original destination closure time, (30-75 minutes refueling time at the alternate is also included in this period)
- 30-60 minutes: Wait for new departure clearance to original destination
- 45 minutes: Fly to original destination

In the current system, the total demand at the destination airport is not reduced as the result of diversions because diverted flights must ultimately fly to it. Demand from diverted flights creates additional delays for flights that are not diverted, further compounding the problem. The situation worsens if the capacity of the destination airport is affected for a longer duration or the capacity shortfall or airport closure occurs in the early hours in the day.

Our proposed RTIMD strategy, in contrast, avoids flying aircraft back to the original hub airport and its associated cost and delay. By utilizing nearby airports in the same region and the ground transport connection between them, and guided by the mathematical programming model that we will present in Section 4.4, an airline can integrate their operations at the original hub and alternative hub airport(s).

The remainder of this chapter is organized as follows. Section 4.2 elaborates on the evolution of passenger itineraries under RTIMD. Section 4.3 discusses the criteria for selecting alternative hubs. A mathematical programming model is proposed in Section 4.4 for a single airline to solve the RTIMD. Section 4.5 conducts a case study by using the model. A so-called Regional GDP, which supports RTIMD by ensuring that diversions do not exceed the capacities of alternate hub airports, is presented in Section 4.6, followed by a summary in Section 4.7.

4.2. Evolution of Passengers' Itineraries under RTIMD

Figure 4-3 depicts how passenger itineraries may be adjusted in response to a disruption under RTIMD. Our example is based on the San Francisco Bay Area regional airport system, consisting of the SFO, Oakland International Airport (OAK), and Mineta San Jose International Airport (SJC) airports. Suppose a flight from SEA to DFW via SFO is diverted to OAK. Original itineraries of passengers using this flight are shown on the left of the figure. A solid arrow represents flying, a dotted arrow dedicated inter-airport motor coach service, and a dashed arrow regular local transport services.

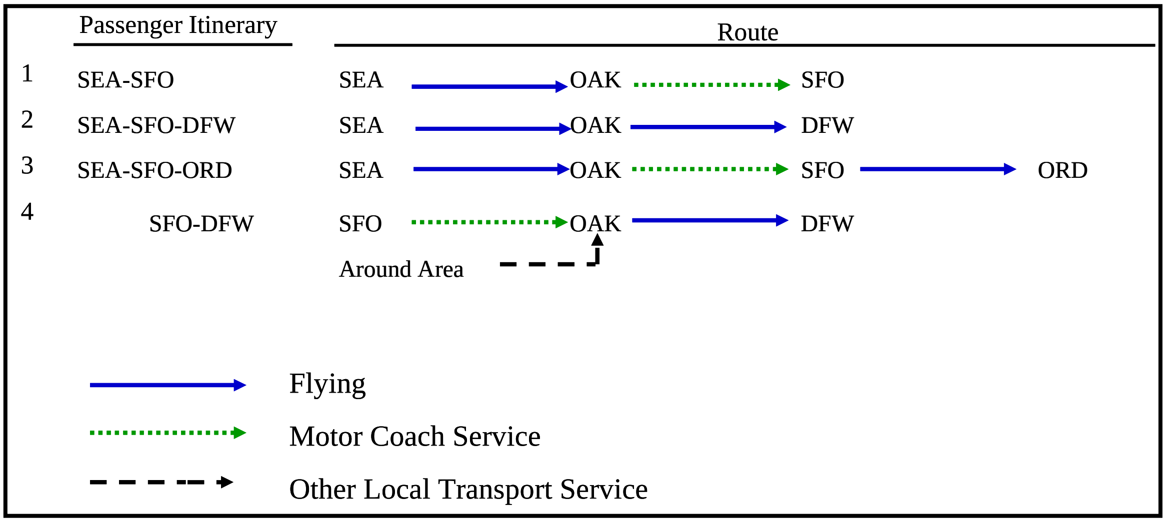

Figure 4- 3 Passenger Itineraries under RTIMD

The first passenger has a final destination of SFO. With the diversion, after getting off the plane at OAK, she can take motor coach service to SFO or reach her local destination using the local surface transport system. The second passenger has a final destination of SEA. With her itinerary of SEA-SFO-DFW, she can continue her adjusted itinerary SEA-OAK-DFW, without a significant adjustment. The destination of the third passenger is Chicago O'Hare International airport (ORD). If the originally scheduled ORD outbound flight has not also been diverted to OAK, then this passenger needs to be rebooked on a regularly scheduled flight from OAK to ORD, or be transported to SFO to board a flight bound for ORD. Meanwhile, the fourth passenger, whose itinerary is from SFO to DFW, needs to be reassigned to another flight from SFO or go to OAK to catch her flight. If a passenger with such an itinerary is already at the airport, she can use the motor coach service provided by airlines to get to OAK. Otherwise, if she is still on her way to SFO or yet to begin her trip, with appropriate communications, she can get reassignment information in advance and go to OAK with local surface transport. The four examples depicted in Figure 4-3 illustrate all possible cases of the evolution of passengers' itineraries with the flight diversion under RTIMD.

4.3. Selection Criteria for an Alternative Hub

Alternative hubs should be selected with input from many stakeholders, including airlines, the air traffic service provider, airport ownership and management, airport neighbors, the regional transport authority, and others. Many factors will influence the decisions, but certain criteria are clearly of paramount importance from a technical point of view. First, the alternative hubs should have excess airfield capacity during periods when these airports likely to be needed. This generally implies that capacity at the alternative hubs should not be strongly affected by the factors that cause capacity shortfalls at the major hub airport. Second, alternate hubs should have runways with sufficient lengths for large transport aircraft. Third, these facilities should be within reasonable driving distance from the major hub airport so that ground travel between the hubs is economical in terms of time and money. Finally, airlines will prefer to divert flights to airports where they also have operations, making it easier for them to consolidate passengers and swap aircraft and crew.

4.3.1. Runway Length

To identify potential alternative hubs within a region, we have to first look into the maximum runway length at the airports. For instance, required take-off and landing runway lengths for typical aircraft types at SFO are presented in Figure 4-4. These lengths were determined based on the assumption that aircraft operate at maximum take-off or landing operating weight with air conditioning off, runways are at 15 feet above mean sea level (AMSL) with dry surface, and it is under international standard atmosphere (ISA) +15 C^o for taking-off and ISA for landing with zero wind. Given daily flight schedules at the major hub airport, and fleet mix of the flights, it is easy to determine the number of flights a potential alternative hub can accept on top of its own scheduled arrivals and departures.

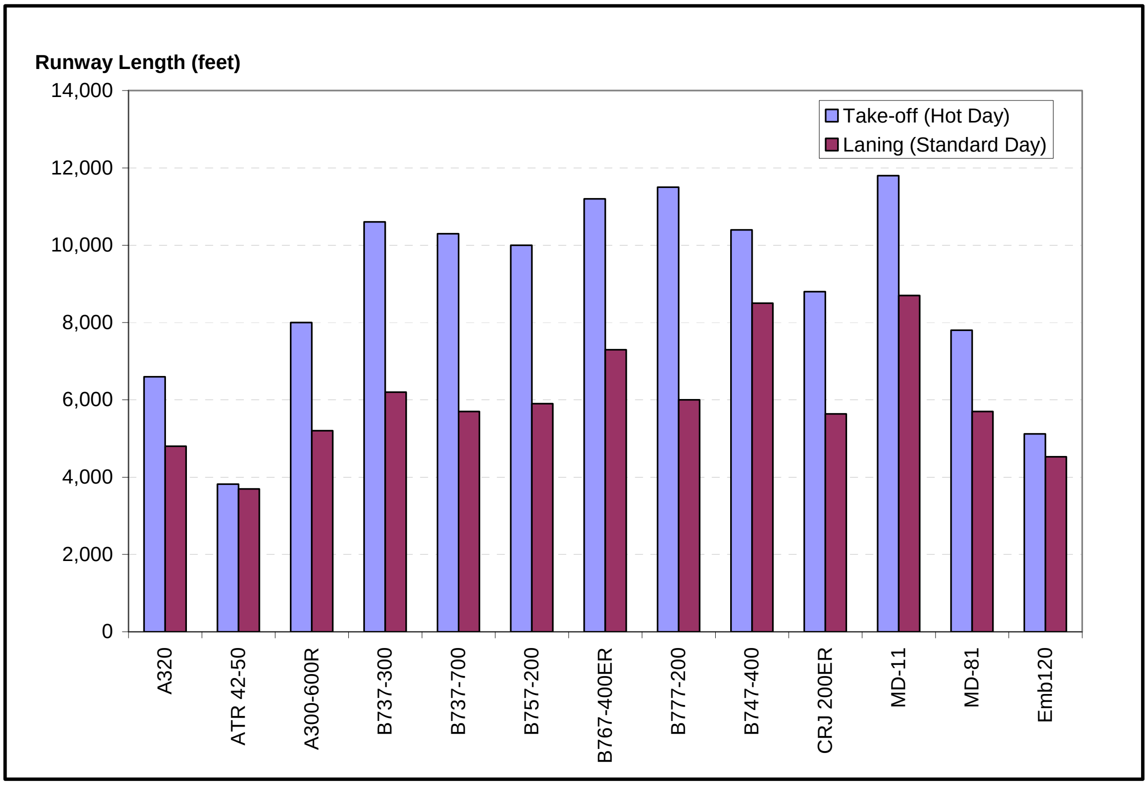

Figure 4- 4 Take-off and Landing Runway Length Requirements by Aircraft Type

4.3.2. Driving Distance between Airports in the System

Driving time between an alternative and a major hub airport is another important selection metric. This depends on the freeway network in the area and accessibility of airports on both ends. As a real-time reaction to disruptions, the driving time should ideally be evaluated with instantaneous information such as traffic conditions, road construction activities, weather conditions, and so on. We assume that these information can be obtained from local traffic management center. In the case that real-time road information is not available, conservative estimates will be used in the proposed mathematical programming model.

4.3.3. Weather Correlation

Weather correlation has to be considered in selecting alternative hubs if the capacity reduction or temporary closure at the major hub airport is caused by adverse weather. Historical meteorological indicators or capacity data could be used to investigate the correlation. The airports with a lower correlation should receive stronger consideration for selection as alternative hubs.

4.3.4. Correlation of Demand Profiles

Another metric that needs to be taken into account is the correlation of demand profiles at a potential alternative hub and the major hub airport. Lower correlation means that the demands at two airports are more complementary, thus it is easier to divert major-hub-airport-bound flights to that alternative hub and fit that traffic into light traffic periods at the alternative hub.

4.4. Mathematical Programming Model of Real-time Inter-modal Diversion (RTIMD)

With short-haul ground substitution, alternative hubs, and ground transport links between alternative and original hub airports, airlines need to decide how to delay, cancel, substitute, and divert flights. The following mathematical programming finds the mix of flight delays, cancellations, substitutions by surface modes, and diversions to alternate hubs that minimizes the cost of disruption resulting from a temporary capacity shortfall at the major hub. It assumes that an alternative hub have already been selected based on the criteria in Section 4.3.

4.4.1. Notation

Most of the notation is the same as that used in the RTIMS model, presented in Chapter 3. For reader's convenience, we listed the notation for the RTIMD model in its entirety in Table 4-1, in alphabetical order.

Table 4- 1 Notation for RTIMD Model

Notation	**Category**	**Description**
$a \in$ A	Index	Arrival flights
AAT_a	Input	Gate to gate time of arrival flight *a*

$AAltCap_t$	Input	Arrival capacity at the alternative hub airport in time period t
ADT_f	Input	Gate to gate time of departure flight f
AD_a	Input	Delay of arrival flight a
$AHPax_a$	Input	Number of local (i.e. not connecting) passengers originally booked on arrival flight a
$AHCap_t$	Input	Arrival capacity at the hub airport in time period t
$APax_a$	Input	Total number of passengers booked on arrival flight a
AT_a	Input	Scheduled arrival time of arrival flight a
BAT_a	Input	Transportation time if arrival flight a is substituted by ground transportation mode
BDT_f	Input	Transportation time if departure flight f is substituted by ground transportation mode
BO	Intermediate Variable	Total operating cost for ground transport mode
$CostA$	Parameter	Cost of utilizing the alternative hub per diverted landing or taking-off
$CostB$	Parameter	Fixed cost of utilizing ground transportation per substitution
$CostBP$	Parameter	Variable cost of utilizing ground transportation per passenger-unit time
$CostD$	Parameter	Estimated cost per disrupted passenger

$CostF$	Parameter	Airlines' operating cost of delaying a flight for one time unit
$CostP$	Parameter	Passenger delay cost per one time unit
$CostTA$	Parameter	Cost of ferrying remaining aircraft back from the alternative hub to the original hub
$CostTF$	Parameter	Cost of transporting passengers between the original and alternative hubs, which is based on passenger value of time and the driving time between the two airports
$DAltCap_t$	Input	Departure capacity at the alternative hub airport in time period t
DD_f	Intermediate Variable	Delay of departure flight f
$DHCap_t$	Input	Departure capacity at the hub airport in time period t
$DHPax_f$	Input	Number of local (i.e. not connecting) passengers originally booked on departure flight f
DP	Intermediate Variable	Number of disrupted passengers
$DPax_f$	Input	Total number of passengers on departure flight f
DT_f	Input	Scheduled departure time of departure flight f
$f \in \Phi$	Index	Departure flights
$IAircraft_k^{\tau}$	Intermediate Variable	Cumulative number of type k aircraft at the hub through time period τ

$IAltAircraft_k^\tau$	Intermediate Variable	Cumulative number of type *k* aircraft at the alternative hub through time period τ
$IAltPax_s^\tau$	Intermediate Variable	Cumulative number of inbound transfer passengers with destination *s* arriving at the alternative hub trough time period τ
$IPax_s^\tau$	Intermediate Variable	Cumulative number of inbound transfer passengers with destination *s* arriving at the hub trough time period τ
$k \in K$	Index	Aircraft type
maircraft	Parameter	Minimum turnaround time of flights
mpax	Parameter	Minimum connecting time of transfer passengers at the hub and alternative hub airport
$OAircraft_k^\tau$	Intermediate Variable	Cumulative number of type *k* aircraft that have departed the hub through time period τ
$OAltAircraft_k^\tau$	Intermediate Variable	Cumulative number of type *k* aircraft that have departed the alternative hub through time period τ
$OAltPax_s^\tau$	Intermediate Variable	Cumulative number of outbound transfer passengers with destination *s* departed the alternative hub trough time period τ
$OPax_s^\tau$	Intermediate Variable	Cumulative number of outbound transfer passengers with destination *s* departed the hub trough time period τ
$Pax_{a,f}$	Input	Number of passengers booked to transfer from arrival flight *a* to departure flight *f*

P_f	Decision Variable	The number of passengers on a departure flight f
$s \in$ Є	Index	Spoke airports
$SpokeA_{as}$	Input	Indicators of origins of arrival flights (equals 1 if arrival flight a originated from spoke s, 0 otherwise)
$SpokeF_{fs}$	Input	Indicators of destinations of departure flights (equals 1 if departure flight f goes to spoke s, 0 otherwise)
$t \in \Gamma$	Index	A set of discrete time periods
$TypeA_{ak}$	Input	Indicators of aircraft type of arrival flights (equals 1 if arrival flight a uses aircraft type k, 0 otherwise)
$TypeF_{fk}$	Input	Indicators of aircraft type of departure flights (equals 1 if departure flight f uses aircraft type k, 0 otherwise)
$TypeO_k$	Input	Number of type k aircraft available at the beginning of schedule perturbation
$TypeT_k$	Input	Number of type k aircraft required to be at the hub airport at the end of schedule perturbation
xd_a^t	Decision variable	Equals 1 if an arrival flight a is planned to be diverted to an alternative hub and arrive there during time period t, 0 otherwise.
xf_a^t	Decision variable	Equals 1 if an arrival flight a is planned to arrive at the hub airport during time period t, 0 otherwise.
xs_a^t	Decision variable	Equals 1 if arrival flight a is substituted with ground transportation and is planned to arrive at the hub airport

		during time period t, 0 otherwise.
yd_f^t	Decision variable	Equals 1 if departure flight f is planned to take-off at an alternative hub during time period t, 0 otherwise.
yf_f^t	Decision variable	Equals 1 if departure flight f is planned to take-off from the hub airport during time period t, 0 otherwise.
ys_f^t	Decision variable	Equals 1 if departure flight f is substituted with ground transportation and is planned to leave from the hub airport during time period t, 0 otherwise.

The time of day is divided into a finite set of time periods of equal duration, and is denoted by $t \in \Gamma = \{1..T\}$. For instance, Γ might be a set of 60 time periods of 10 minutes each, summing to a planning horizon of 10 hours. Arrival and departure slots assigned to an airline at the primary hub airport are denoted as $AHCap_t$ and $DHCap_t$, while slots at the alternative hub are labeled $AAltCap_t$ and $DAltCap_t$ for $t \in \Gamma$. The airline has arrival flights $a \in \mathrm{A} = \{1..A\}$ that are scheduled to fly to the hub airport with scheduled arrival time denoted by AT_a and departure flights $f \in \Phi = \{1..F\}$ with scheduled departure time denoted by DT_f . En-route flight times of these flights are assumed to be constant and recorded in two vectors AAT_a and ADT_a, respectively. Spoke airports in this hub-and-spoke network are represented by a set $s \in Є = \{1..S\}$. $SpokeA$ and $SpokeF$ record arrival flights' origins and departure flights' destinations. If $SpokeA_{as} = 1$, arrival flight a originates from spoke airport s, 0 otherwise; and likewise for $SpokeF$. Aircraft type is denoted by $k \in \mathrm{K} = \{1..K\}$. If $TypeA_{ak} = 1$, arrival flight a uses a type k aircraft, 0 otherwise. If $TypeF_{f,k} = 1$, departure flight f uses a type k aircraft, 0 otherwise. The number of type k aircraft available at the beginning of the schedule perturbation period at the hub airport is denoted as $TypeO_k$. The number of type k aircraft required to be available at the end of perturbation period is denoted as $TypeT_k$.

Passenger itinerary information for the model is shown in Table 4-2. In Table 4-2, the first column is arrival flights and the first row departure flights. $AHPax_a$ in the second column is the number of passengers on arrival flights a with destinations at the hub airport. $DHPax_f$ in the second row is the number of passengers on a departure flight f originating from the hub airport. $Pax_{a,f}$ denotes the number of passengers whose itinerary involves a transfer from an arrival flight a to a departure flight f. In the last column and last row, the total passengers on arrival flights and departure flights are presented by $APax_a$ and $DPax_f$, respectively.

Table 4- 2 Input Data: Passenger Itinerary Information

		F1	F2	F3	…	
		DHPa,x_1	DHPa,x_2	DHPa,x_3	…	
A_1	AHPa,x_1	$Pax_{1,1}$	$Pax_{1,2}$	$Pax_{1,3}$	…	$APax_1$
A_2	$AHPax_2$	$Pax_{2,1}$	$Pax_{2,2}$	$Pax_{2,3}$	…	$APax_2$
A_3	$AHPax_3$	$Pax_{3,1}$	$Pax_{3,2}$	$Pax_{3,3}$	…	$APax_3$
…	…	…	…	…	…	…
		$DPax_1$	$DPax_2$	$DPax_3$	…	

We assume ground transportation time for substituting each hub-airport-bound flight (arrival or departure) is estimated based on historical data and real-time traffic conditions. The time is then recorded in two vectors BAT_a and BDT_f, for $a \in \mathrm{A}$ and $f \in \Phi$. Other parameters needed for this model include minimum turnaround times of flights and minimum connecting time of transfer passengers at the hub airport, denoted as $maircraft$ and $mpax$ respectively. As same as what we have done in Chapter 3, we assume constant values 45 minutes for $maircraft$ and 30 minutes for $mpax$. The values of the cost parameters for this model are listed in Appendix A.

4.4.2. Decision Variables

The decision variables in the model include a set of binary decision variables for scheduled arrival flights, a set of binary decision variables for departing flights, and a set of integer decision variables for numbers of passengers on departure flights. If an arrival flight a lands at the hub airport during time unit t, xf_a^t equals 1, otherwise, 0. If it is substituted for by ground transport mode that will arrive at the hub airport during time unit t, xs_a^t equals 1, otherwise 0. If flight a is diverted to the alternative hub where it arrives during time period t, then xd_a^t equals 1, otherwise 0. If none of previous actions occure during the entire planning horizon, then the arrival flight is cancelled. The process is similar for departure flights. The set of passenger decision variables determine the numbers of passengers on departure flights. A departure flight can be held to wait for connecting passengers. The cost of doing so is the sacrifice of the time of the passengers and crews who have been ready for departure. The decision variables are listed as follows:

$$xf_a^t = \begin{cases} 1 \text{ if arrival flight a is planned to arrive at the hub airport} \\ \text{during time period t;} \\ \\ 0 \text{ Otherwise} \end{cases} \qquad a \in A$$

$$xs_a^t = \begin{cases} 1 \text{ if arrival flight a is planned to be substituted with ground} \\ \text{transportation and arrive at the hub airport during time period t;} \\ \\ 0 \text{ Otherwise} \end{cases} \qquad a \in A$$

$$xd_a^t = \begin{cases} 1 \text{ if arrival flight a is planned to be diverted to the alternative hub} \\ \text{and arrive during time period t;} \\ \\ 0 \text{ Otherwise} \end{cases} \qquad a \in A$$

$$yf_f^t = \begin{cases} 1 \text{ if departure flight } f \text{ is planned to leave the hub airport} \\ \text{during time period } t \\ \\ 0 \text{ otherwise.} \end{cases} \qquad f \in \Phi$$

$$ys_f^t = \begin{cases} 1 \text{ if departure flight } f \text{ is substituted with ground transporation} \\ \text{and planned to leave the hub airport during time period } t \\ \\ 0 \text{ otherwise.} \end{cases} \qquad f \in \Phi$$

$$yd_a^t = \begin{cases} 1 \text{ if departure flight f is planned to take-off from the alternative} \\ \text{hub airport during time period t;} \\ \\ 0 \text{ Otherwise} \end{cases} \qquad f \in \Phi$$

$P_f \geq 0$ Integer, the number of passengers on departure flight f $\qquad f \in \Phi$

4.4.3. Objective Function

Given original flight schedules, airport capacity profiles, passenger matrix, and parameters, an airline implementing RTIMD needs to determine the decision variables so as to minimize the disruption cost which is the summation of the following components.

<u>Passenger Delay Cost</u>

$$\begin{pmatrix} \sum_a AD_a \cdot AHPax_a + \sum_s \sum_t \left(IPax_s^t - OPax_s^t \right) \\ + \sum_s \sum_t \left(IAltPax_s^t - OAltPax_s^t \right) + \sum_f DD_f \cdot P_f \end{pmatrix} \cdot CostP \tag{4.4.1}$$

Passenger delay cost includes the cost related to (1) the delay of passengers on arrival flights with destinations at the hub airport; (2) passengers' waiting time at the hub and alternative hub airports; (3) the delay of passengers on departure flights. Passengers with the same itinerary are considered identical in the sense that they suffer the identical cost for any particular form of delay during disruptions[19]. The arrival and departure delays in the objective function are calculated as follows.

$$AD_a = \sum_t \left(t - AT_a \right) \cdot \left(xf_a^t + xd_a^t \right) + \left(1 - \sum_t \left(xf_a^t + xd_a^t \right) \right) \cdot ACanT_a \tag{4.4.2}$$

$$DD_f = \sum_t \left(t - DT_f \right) \cdot \left(yf_f^t + yd_f^t \right) + \left(1 - \sum_t \left(yf_f^t + yd_f^t \right) \right) \cdot DCanT_f \tag{4.4.3}$$

---

[19] Nevertheless, airlines may give priorities to different passengers when they implement this strategy, based on passengers' loyalty, value to airlines, or other criteria. For instance, if a flight cancellation causes some passengers to be rebooked immediately and others to be held at the hub for a longer period, these criteria may be used to decide who gets the more desirable option.

Where $ACanT_a$ and $DCanT_f$ are estimated delay hours if flights are cancelled. According to previous studies, the value is in the range of 6 to 7 hours [Bratu and Barnhart 2006]. Using these parameters in this formulation is an approximate way to count flight cancellation cost into airline disruption cost.

The passenger inflow and outflow at the major hub airport with respect to a specific spoke airport is calculated as follows:

$$IPax_s^\tau = \sum_{t=1}^{\tau} \sum_a \sum_f \left(xf_a^t + xs_a^t \right) \cdot Pax_{a,f} \cdot SpokeF_{f,s} \tag{4.4.4}$$

$$OPax_s^\tau = \sum_{t=1}^{\tau} \sum_f \left(yf_f^t + ys_f^t \right) \cdot \left(P_f - DHPax_f \right) \cdot SpokeF_{f,s} \tag{4.4.5}$$

Likewise, the passenger inflow and outflow at the alternative hub airport is as follows:

$$IAltPax_s^\tau = \sum_{t=1}^{\tau} \sum_a \sum_f xd_a^t \cdot Pax_{a,f} \cdot SpokeF_{f,s} \tag{4.4.6}$$

$$OPax_s^\tau = \sum_{t=1}^{\tau} \sum_f yd_f^t \cdot \left(P_f - DHPax_f \right) \cdot SpokeF_{f,s} \tag{4.4.7}$$

As discussed in Chapter 3, and shown in Figure 3-8, passenger waiting time at one airport can be captured by calculating the area between the inflow and outflow passenger curves. *CostP* in Expression (4.4.1) is a parameter reflecting passenger value of time.

Inter-hub Connection Cost

$$\begin{pmatrix} \sum_a \sum_f \sum_t Pax_{a,f} \cdot yd_f^t + \sum_s \sum_t \left(IAltPax_s^T - OAltPax_s^T \right) \\ + \sum_a \sum_t AHPax_a \cdot xd_a^t + \sum_f \sum_t DHPax_f \cdot yd_f^t \end{pmatrix} \cdot CostTF \tag{4.4.8}$$

Recalling the evolution of passenger itineraries discussed in Section 4.2, for a transfer passenger, if her inbound flight has been diverted to an alternative hub while the original scheduled outbound flight has not, this passenger needs to be rebooked on a regularly scheduled flight from the alternative hub to her destination, or be transported to the major hub airport to catch flights there going to her destination. A similar situation occurs in the case of transfer passengers whose outbound flight has been diverted to the alternative hub after they have arrived at the major hub airport. Multiplication of binary flight decision variables is required if we introduce the exact calculation of these numbers of passengers. This would greatly increase the complexity of the model. To avoid this complication, we use the first component in Expression (4.4.8) to estimate the total number of transfer passengers that need to be transported between the primary and alternative hubs. Additionally, we will make the simplifying assumption that all passengers on diverted arrival flights with destinations at the hub airport need to take the motor coach service to get to the major hub airport, although in many cases they could proceed directly to their local destination. Similarly, local passengers booked to take departure flights will have to go to the alternative hub if their outbound flights have been diverted there. *CostTF* is a cost coefficient for calculating inter-hub connection cost and it is determined by passenger value of time and the driving time between the original and alternative hubs.

Cost of Utilizing the Alternative Hub

$$\left(\sum_a \sum_t xd_a^t + \sum_f \sum_t yd_f^t\right) \cdot CostA \tag{4.4.9}$$

Expression (4.4.9) is the cost of utilizing ground facilities at the alternative hub, which depends on the number of flight operations arriving at and departing from the alternative hub, excluding operations of ferrying surplus aircraft. *CostA* is the cost coefficient for calculating this cost.

Flight Delay Cost

$$\left(\sum_a AD_a + \sum_f DD_f\right) \cdot CostF \tag{4.4.10}$$

Flight delay cost is calculated as the flight delay multiplied by a cost coefficient *CostF*.

Cost of Ferrying Aircraft Back to the Original Hub Airport

$$\sum_k \left(IAltAircraft_k^T - OAltAircraft_k^T\right) \cdot CostTA \tag{4.4.11}$$

The model allows the total number of outbound diverted flights be less than the total inbound diverted flights. In another word, it allows aircraft to be left over at the alternative hub and then be ferried back to the original hub airport at the end of planning horizon. *CostTA* is the cost coefficient of ferrying one such aircraft. The aircraft inflow and outflow of type k at the end of planning horizon is calculated as follows:

$$IAltAircraft_k^T = \sum_{t=1}^{T} \sum_a xd_a^t \cdot TypeA_{a,k} \tag{4.4.12}$$

$$OAltAircraft_k^T = \sum_{t=1}^{T} \sum_f yd_f^t \cdot TypeD_{f,k} \tag{4.4.13}$$

Disrupted Passenger Cost

$$DP \cdot CostD \tag{4.4.14}$$

The number of disrupted passengers at the hub airport and the alternative hub airport are the following:

$$\begin{aligned} DP = \sum_a \left(1 - \sum_t \left(xf_a^t + xs_a^t + xd_a^t \right) \right) \cdot \alpha \cdot APax_a + \sum_s \left(IPax_s^T - OPax_s^T \right) \\ + \sum_s \left(IAltPax_s^T - OAltPax_s^T \right) \end{aligned} \tag{4.4.15}$$

The inflow and outflow at a spoke airport is not tracked in this study, so the first component is the estimated number of disrupted passengers if the inbound flight is cancelled. The second and third components indicate the number of transfer passengers left over at the major or alternative hub at the end of the planning horizon. *CostD* is the unit cost parameter for passenger disruption. It is based on passenger value of time and estimated average waiting time of disrupted passengers before they can be assigned to another itinerary beyond the planning horizon.

Short-haul Flight Substitution Cost

$$\begin{aligned} BO = \left(\sum_a APax_a \cdot ABT_a \cdot \sum_t xs_a^t + \sum_f P_f \cdot DBT_f \cdot \sum_t ys_f^t \right) \cdot CostBP \\ + \left(\sum_a \sum_t xs_a^t + \sum_f \sum_t ys_f^t \right) \cdot CostB \end{aligned} \tag{4.4.16}$$

Expression (4.4.16) is the operating cost of motor coaches used to substitute cancelled short-haul flights. The cost includes a fixed cost and a variable cost that is proportional to passenger times for those who are reassigned to ground transportation. *CostBP* and *CostB* are cost coefficients for fixed and valuable costs, respectively.

Hence, the entire objective function is presented as the following:

$$
\begin{aligned}
Min \quad & \left(\begin{array}{l} \sum_a AD_a \cdot AHPax_a + \sum_s \sum_t \left(IPax_s^t - OPax_s^t\right) \\ \sum_s \sum_t \left(IAltPax_s^t - OAltPax_s^t\right) + \sum_f DD_f \cdot P_f \end{array}\right) \cdot CostP \\
& + \left(\begin{array}{l} \sum_a \sum_f \sum_t Pax_{a,f} \cdot yd_f^t + \sum_s \sum_t \left(IAltPax_s^T - OAltPax_s^T\right) \\ + \sum_a \sum_t xd_a^t \cdot AHPax_a + \sum_f \sum_t yd_f^t \cdot DHPax_f \end{array}\right) \cdot CostTF \\
& + \left(\sum_a \sum_t xd_a^t + \sum_f \sum_t yd_f^t\right) \cdot CostA + \left(\sum_a AD_a + \sum_f DD_f\right) \cdot CostF \\
& + \sum_k \left(IAltAircraft_k^T - OAltAircraft_k^T\right) \cdot CostTA + + DP \cdot \cos tD + BO
\end{aligned}
\tag{4.4.17}
$$

4.3.4. Constraints

The first set of constraints is arrival and departure capacity constraints at the hub airport.

$$\sum_a xf_a^t \le AHCap_t \qquad \forall t \in \Gamma \tag{4.4.18}$$

$$\sum_f yf_f^t \le DHCap_t \qquad \forall t \in \Gamma \tag{4.4.19}$$

$$\sum_a xd_a^t \le AAltCap_t \qquad \forall t \in \Gamma \tag{4.4.20}$$

$$\sum_f yd_f^t \le DAltCap_t \qquad \forall t \in \Gamma \tag{4.4.21}$$

Besides being delayed or substituted for, flights can also be cancelled. This is presented by the following set of constraints.

$$\sum_t \left(xf_a^t + xs_a^t + xd_a^t\right) <= 1 \qquad \forall a \in \mathrm{A} \tag{4.4.22}$$

$$\sum_{t}\left(yf_f^t + ys_f^t + yd_f^t\right) <= 1 \qquad \forall f \in \Phi \tag{4.4.23}$$

The next two constraints reflect passenger conservation, i.e. the cumulative number of departed transfer passengers from the major hub airport to a destination s through time t should be no more than the cumulative number of transfer passengers arrived at the airports through t - $mpax$, where $mpax$ is the minimum passenger connecting time. A similar constraint is applicable for the passengers at the alternative hub.

$$IPax_s^{\tau} \geq OPax_s^{(\tau+mpax)} \quad \forall s \in \Theta \quad \forall \tau \in \{1..(T-mpax)\} \tag{4.4.24}$$

$$IAltPax_s^{\tau} \geq OAltPax_s^{(\tau+mpax)} \qquad \forall s \in \Theta \quad \forall \tau \in \{1..(T-mpax)\} \tag{4.4.25}$$

Another critical constraint is aircraft flow conservation. We allow aircraft swapping among flights utilizing the same type of aircraft. At the hub airport, there are inventories of different aircraft types at the beginning of the disruption and also aircraft requirements at the end. Thus, the constraints are as follows.

$$TypeO_k + IAircraft_k^{\tau} \geq OAircraft_k^{(\tau+maircraft)}$$

$$\forall k \in \mathrm{K} \quad \forall \tau \in \{1..(T-maircraft)\} \tag{4.4.26}$$

$$TypeO_k + IAircraft_k^{T} \geq OAircraft_k^{T} + TypeT_k \qquad \forall k \in \mathrm{K} \tag{4.4.27}$$

where:

$$IAircraft_k^{\tau} = \sum_{t=1}^{\tau}\sum_{a} xf_a^t \cdot TypeA_{a,k} \tag{4.4.28}$$

$$OAircraft_k^{\tau} = \sum_{t=1}^{\tau}\sum_{f} yf_f^t \cdot TypeD_{f,k} \tag{4.4.29}$$

At the alternative hub airport, we suppose there are no aircraft inventories and no requirements. The constraints are as follows.

$$IAltAircraft_k^{\tau} \geq OAltAircraft_k^{(\tau + maircraft)}$$

$$\forall k \in \mathrm{K} \quad \forall \tau \in \{1..(T - maircraft)\} \tag{4.4.30}$$

where:

$$IAltAircraft_k^{\tau} = \sum_{t=1}^{\tau} \sum_{a} xd_a^t \cdot TypeA_{a,k} \tag{4.4.31}$$

$$OAltAircraft_k^{\tau} = \sum_{t=1}^{\tau} \sum_{f} yd_f^t \cdot TypeD_{f,k} \tag{4.4.32}$$

In addition, relative flights are not allowed to be released from their origins earlier than their originally scheduled times. For an arrival flight from a spoke airport, there are two constraints depending on if that flight is substituted or not. The constraints are as follows:

$$\sum_t t \cdot \left(xf_a^t + xd_a^t\right) \geq AT_a \cdot \sum_t \left(xf_a^t + xd_a^t\right) \qquad \forall a \in \mathrm{A} \tag{4.4.33}$$

$$\sum_t t \cdot \sum_h xs_a^{t,h} \geq \left(AT_a + \left(BAT_a - AAT_a\right)\right) \cdot \sum_t xs_a^t \qquad \forall a \in \mathrm{A} \tag{4.4.34}$$

In comparison, there is only one constraint for departure flights at the hub airport.

$$\sum_t t \cdot \left(yf_f^t + ys_f^t + yd_f^t\right) \geq DT_f \cdot \sum_t \left(yf_f^t + ys_f^t + yd_f^t\right) \quad \forall f \in \Phi \tag{4.4.35}$$

The number of passengers on departure flights P_f has to be less than the seating capacity of aircraft used by flight f, $DCap_f$.

$$P_f \le DCap_f \qquad \forall f \in \Phi \tag{4.4.36}$$

4.4. A Case Study : Comparison of RTIMS and RTIMD

4.4.1. Case Study Airport

A west coast hub airport, San Francisco International Airport (SFO), is chosen as our case study airport. SFO is a large hub airport located in a regional airport system together with Oakland International Airport (OAK), another large hub airport and Mineta San Jose International Airport (SJC), a medium hub airport.[20] As shown in Figure 4-5, there are several spoke airports located within a three-hour driving radius around SFO.

4.4.2. Data Preparation

Official Airline Guide (OAG)

We obtain flight information from the OAG[21] database. The database contains origins of arrival flights and destinations of departure flights. It also includes their scheduled arrival times and departure times at the hub airport, operating airlines, equipment information, and aircraft seating capacities.

---

[20] FAA defined large, medium, small, and non-hub airports according to annual revenue passengers. Large hub airports are those which process at least one percent of revenue passenger boardings annually, medium hub airports are those which process between 0.25 percent and one percent of revenue passenger boardings annually, small hub airports are those which process between 0.05 percent and 0.25 percent of revenue passenger boardings annually, whether or not in scheduled service.

[21] OAG is best known for its airline schedules database. This holds future and historical flight details for 1,000 airlines and more than 3,500 airports.

Figure 4- 5 San Francisco International Airport and Nearby Airports

<u>DOT Data Bank 1A</u>

Passenger itinerary information is not available directly from airlines. The way of simulating the information is described as follows.

- Obtain OD traffic table from the hub traffic at itinerary level processed by Data Base Products Inc. based on the DOT data bank which contains DOT 10% percent coupon survey data.
- Obtain flights’ arrival and departure times from OAG data.

- With information obtained from the above database, use the mathematical programming model described in Appendix B to synthesize the passenger itinerary information.

Capacity Profiles at the Original and Alternative Hubs

Liu and Mark in their book, "*Managing Uncertainty in the Single Airport Ground Holding Problem*", propose to use statistical cluster analysis to classify arrival capacity data into patterns of arrival capacity profiles (refer to the book for details on how the technique is applied to determine capacity profiles) [Liu and Hansen 2008]. We borrowed one chart from their book, which demonstrates capacity profiles at SFO. The data used to construct this chart are quarter-hourly Arrival Acceptance Rates (AARs) from Aviation System Performance Metrics (ASPM) database for year 2003.

For the case study, profile "SFO6" in Figure 4-6 is picked as the base capacity profile, i.e. the capacity drops to the half of the normal level for 5 hours, from 7:00 am to 12:00 pm, then returns to the normal level. Slots are first given to international flights and long-haul flights beyond a certain distance that are exempted from GDP. The rest of the slots are allocated to airlines according to their original schedules. Thus the airline in this case study gets four slots per hour for first five hours, from 7:00 am to 12:00 pm, and eight slots per hour for the rest of the day. We assume that OAK has been selected as the alternative hub with excess capacity of two arrivals and two departures per hour.

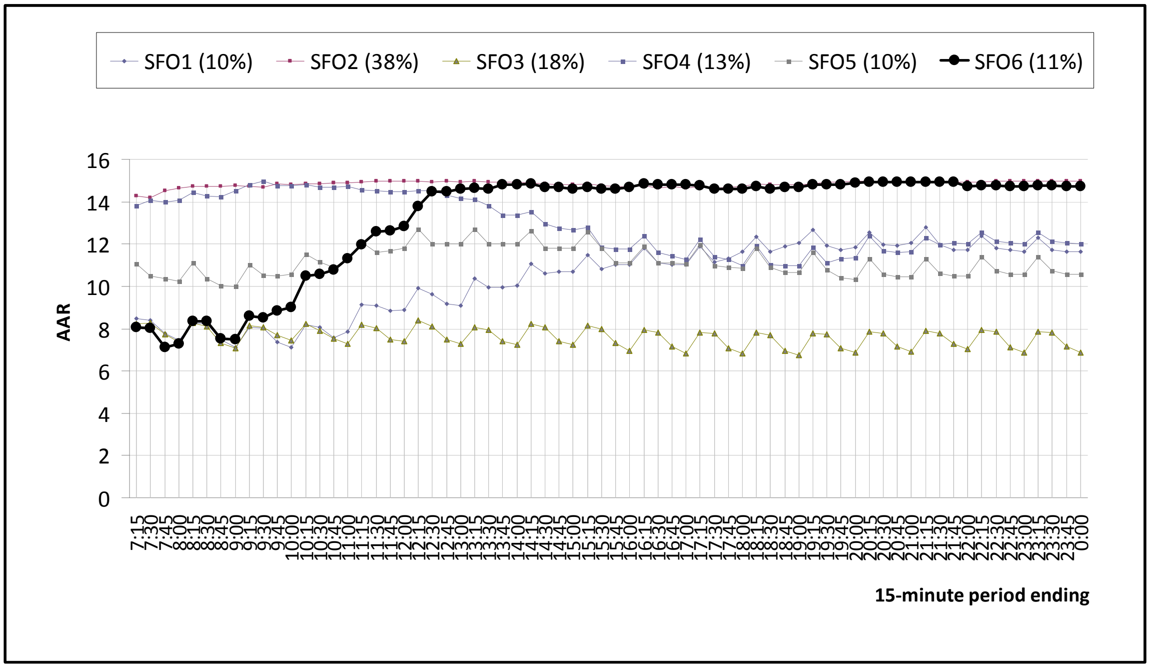

Figure 4- 6 Quarter-hourly Capacity Scenarios at SFO [Liu and Hansen 2008; Used by Permission]

4.4.3. Optimization Results

Table 4-3 summarizes optimization results. As shown in the first column, among 40 flights scheduled to arrive at the major hub airport within two and half hours, 12 are cancelled and all of these are substituted by ground transport, eight are diverted to the alternative hub, while the remaining 20 flights continue flying to the major hub airport. There are more cancellations for outbound flights, and among the 20 cancellations, 15 are substituted with ground transportation. Six outbound flights are reassigned to depart from the alternative hub. The substitution rate is somewhat high because we have assumed a constant passenger value of time. The solution consequently includes some rather long motor coach trips, up to about six-hour driving. These might not be in the optimal solution if the travel time cost function were convex, as it may well be. A useful extension to our model, which we did not attempt, would be to incorporate a piecewise linear travel time cost function so that the effects of convexity could be captured.

Table 4-3 also shows the results from RTIMS, and the strategy of diversion-only without short-haul substitution. It shows that diversion-only strategy yields the highest cost and number of disrupted passengers. RTIMS has only slightly higher cost and number of disrupted passengers than RTIMD. The last row of the Table 4-3 shows the percentage decrease in disruption cost in comparison to a strategy without any inter-modal components. RTIMD reduces disruption cost about 28 percent, RTIMS, 24 percent, and the “Diversion Only” strategy only 12 percent. The benefit of RTIMD over RTIMS is marginal because the capacity constraints at the alternative hub imposed in this case study is relatively tight. It can be seen from the Figure 4-6 that the excess capacity constraints at the alternative hub are binding in the solution.

Table 4- 3 Optimization Results with Different Strategies

	RTIMD	RTIMS	Diversion Only
Total Cost ($)	804,646	846,734	988,481
Inbound Cancellation	12	15	2
Substitution	12	13	-
Inbound Diversion	8		10
Outbound Cancellation	20	22	7
Substitution	15	17	-
Outbound Diversion	6		8
Disrupted Passengers	421	462	671
% Diff. from without Inter-modalism	28.1%	24.3%	11.6%

Figure 4-7 shows the evolution of arrival flights at both hub airports under the RTIMD solution. The curve on the left of the chart is the cumulative number of arrivals according to their original schedules. The piece-wise linear curve on the right of the chart corresponds the output curve if no arrival flights were delayed, cancelled, substituted, or diverted. Triangles, squares, and diamonds on the chart represent outcomes from RTIMD. Each triangle indicates an inbound flight arriving at the major hub airport, with the ordinate indicating its arrival time and the abscissa its plane in the original arrival sequence. Each square depicts, using the same plotting convention, the arrival at the major hub airport of a motor coach that is replacing an inbound flight. Each diamond likewise symbolizes a diverted inbound flight arriving at the alternative hub. With RTIMD, most flights arrive at either the original or the alternative hub earlier than when they would arrive at the original hub, given the capacity shortfall and the original arrival curve. Nevertheless, there are several ground transportation trips that arrive later than the flights they are replacing would under the original “no action” scenario. This has implications for equity and passenger acceptance that will be discussed in Chapter 5.

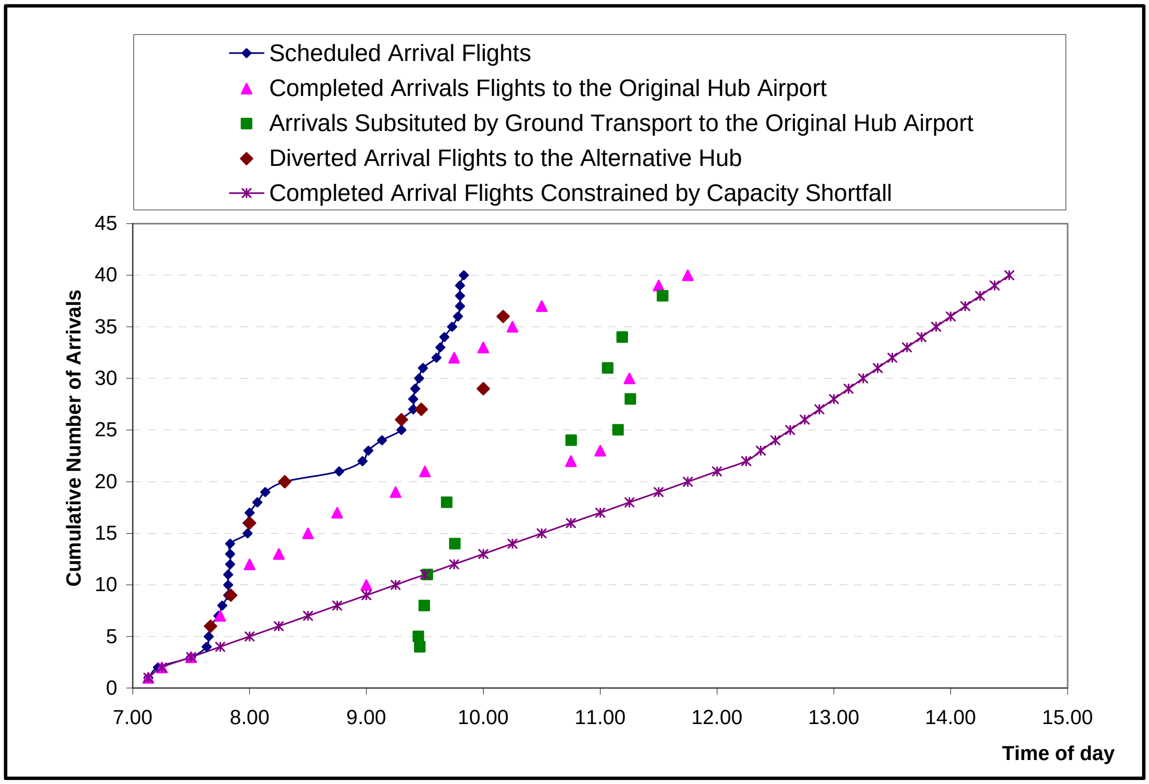

Figure 4- 7 The Evolution of Arrival Flights with RTIMD Strategy

4.4.4. Comparison of RTIMS and RTIMD for Various Capacity Shortfall Scenarios

To investigate how performance of the model varies with the severity of the capacity shortfall, we ran it under three scenarios with different levels of severity. Table 4-4 summarizes the scenarios.

Table 4- 4 Hub Airport Capacity Scenarios for the Case Study

Scenario 1	The capacity at the major hub airport is reduced to half of normal level for three hours, and then returns to normal
Scenario 2	The capacity at the major hub airport is reduced to half of normal level for five hours, and then returns to normal
Scenario 3	The major hub airport is closed for five hours and then the capacity returns to normal

The severity of capacity shortfall increases with the order of the scenarios. Table 4-5 summarized the results of this experiment for both the RTIMS and RTIMD strategies. Since RTIMD also includes inter-modal substitution, the differences in the results reflect the value of being able to divert flights to alternates as well as substitute flights with motor coach service.

Table 4- 5 Comparison of RTIMS and RTIMD

		Scenario1	Scenario2	Scenario3
RTIMS	Total Cost	537,843	846,734	1,121,040
	Inbound Cancellation	8	15	22
	Substitution	8	13	21
	Outbound Cancellation	12	22	26
	Substitution	12	17	21
	Disrupted Passengers	265	462	553
RTIMD	Total Cost	512,019	804,646	1,000,571
	Inbound Cancellation	7	12	16
	Substitution	7	12	16
	Inbound Diversion	4	8	12
	Outbound Cancellation	10	20	22
	Substitution	4	15	18
	Outbound Diversion	4	6	8
	Disrupted Passengers	260	421	502

As expected, with a particular strategy, the total disruption cost increases when the severity of the capacity shortfall increases. The flight cancellations, substitutions, and disrupted passengers have similar trends. Comparing the two strategies, RTIMD leads to lower costs, as well as fewer fight cancellations and substitutions in all scenarios. In comparison to RTIMS, the delay saved by diverting flights with RTIMD increases when the severity of capacity shortfall increases. This is depicted in Figure 4-8.

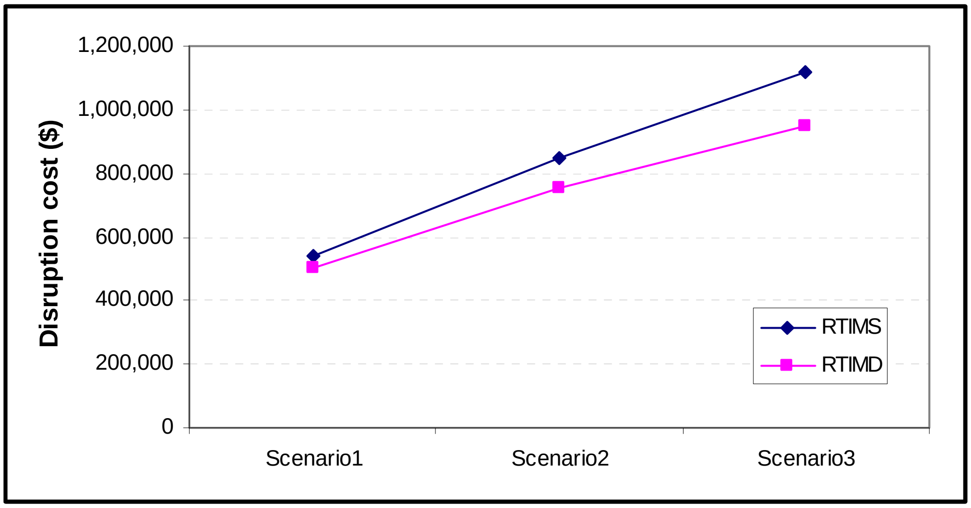

Figure 4- 8 Comparisons of RTIMS and RTIMD

4.4.5. The Impact of Rapid Ground Connection between Hubs

Heretofore is has been assumed that ground driving time between SFO and OAK is one hour. However, the driving time could be much shorter if Bus Rapid Transit (BRT) was available or there were a tunnel connecting the two airports. Hence, we reduce the driving time between the airports to half an hour and investigate the impact of this change on the optimum solution. The comparison of base case and the case with reduced ground travel time is shown in Table 4-6. Outcomes of cancellation, substitution, and diversions are not affected, a consequence of the fact that the alternative hub capacity constraints are binding under both scenarios. However, the number of disrupted passengers decreases and the value of objective function is slightly lower when the ground travel time is reduced. Conversely, if we increase the ground travel time between hubs from one hour to 1.5 hours, the number of diversions decreases markedly while the number of disrupted passengers increases. The minimum disruption cost also increases substantially under this scenario.

Table 4- 6 Impact of Ground Connection Time between Original and Alternative Hubs

RTIMD	Ground Connection Between Hubs		
	1 1/2 hour	1 hour	1/2 hour
Total Cost ($)	883,820	804,646	780,506
Inbound Cancellation	15	12	12
Substitution	13	12	12
Inbound Diversion	4	8	8
Outbound Cancellation	22	20	20
Substitution	17	15	15
Outbound Diversion	3	6	6
Disrupted Passengers	450	421	389

4.5. Regional GDP

RTIMD encourage airlines to redistribute flights from a hub where its capacity is temporarily inadequate to alternative hubs where excess capacities may be available. While benefits of this strategy are obvious, one consequence is that, in the absence of coordination, diverted traffic at the alternative hubs may exceed their capacities. This poses an ATFM problem that the current TFM system was not designed to handle. GDPs in the current system evaluate the imbalance between supply and demand at one airport and provide airlines with CTAs for their flights into that airport. To implement RTIMD in the most efficacious manner, the current GDP should be enhanced to determine CTAs at not only the major hub airport but other airports in a regional airport system, and do so in the manner that airlines consider equitable. Furthermore, the enhanced GDP should reveal the shadow price of slots at the alternative hubs, in order to determine appropriate limits on the quantity of flight traffic that should be diverted to any given alternate. While overly stringent limits would reduce the effectiveness of RTIMD in mitigating congestion and delay at the major hub, ones that are too lax could result in the alternates being overwhelmed.

4.5.1. Mathematical Programming Model for Regional GDP

A mathematical programming model below is developed to determine the CTAs and investigate the shadow price of slots at the alternative hubs. The model works from the point of view of the air traffic service provider. Given airlines' original schedule and capacity profiles at major and the alternative hub airports, it determines which arrival flight should be diverted to the alternative hub so as to minimize the total cost of recovering from a disruption. The results of this model would be used to set capacity constraints for individual airlines as inputs to the RTIMD model discussed in previous section.

Notation

Table 4-7 lists all notations used for Regional GDP model.

Table 4- 7 Notation for Regional GDP Model

Notation	**Category**	**Description**
$i \in I$	Index	Arrival flights.

BT_{ij}	Input	Ground transportation time from the alternative hub j to the original hub for flight i.
c_I	Input	Reduced capacity rate at the major hub airport
c_V	Input	Resumed capacity rate at the major hub airport
C_j	Parameter	Cost of utilizing airport j as an alternative hub
D_k	Intermediate Variable	Cumulative number of arrivals through time period k.
Δ_i^a	Input	Indicator of ownership of flight i (equals 1 if flight i belongs to airline a, 0 otherwise)
Λ_{ij}	Input	Indicator of runway qualification (equals 1 if the runway length requirement for flight i is satisfied, 0 otherwise)
$ECap_{nj}$	Input	Hourly excess capacity at alternative hub j.
ϕ_n^a	Input	Percentage of airline a's flights to the total flights in hour n.
HSA_i	Input	Indicator for which hourly time period flight i's original schedule time falls into
$j \in J$	Index	Airports other than the major hub in the system
M	Parameter	A very large number.
Mis_i	Parameter	Penalty for passengers on flight i assume there is certain possibility that they may miss their connection.
Pax_i	Input	Number of passengers booked on flight i.

P_k	Intermediate Variables	Total number of passengers for time period k .
T_I	Input	Length of capacity reduction
$TPax_i$	Input	Number of transfer passengers on flight i .
w_k	Intermediate Variables	Delay occurred in time period k
x_{ij}	Decision Variable	Equals to 1 if flight i is diverted to alternative hub j
y_j	Decision Variable	Equals to 1 if airport j is used as an alternative hub, 0 otherwise.

Objective Function

The objective function for the regional GDP model, like that of the previous models, is the total cost of recovery from the disruption caused by the loss of capacity at the major hub. In this case, however, the cost is for the whole system rather than an individual airline. The function is:

$$Min \qquad \sum_k w_k \cdot P_k + \sum_i \sum_j x_{ij} BT_{ij} Pax_i + \sum_i \sum_j x_{ij} TPax_i Mis_i + \sum_j C_j y_j \qquad (4.5.1)$$

with:

$$\begin{aligned} w_k &= \min\left(\frac{D_k}{c_I} - t_k, \frac{D_k - c_I T_I}{c_V} - (t_k - T_I) \right) \qquad 0 < t_k \le T_I \\ &= \max\left(0, \frac{D_k - c_I T_I}{c_V} - (t_k - T_I) \right) \qquad t_k > T_I \end{aligned} \qquad (4.5.2)$$

$$D_k = \sum_{i \in \{i | HSA_i < t_k\}} \left(1 - \sum_j x_{ij} \right) \qquad \forall k \in \mathrm{K} \qquad (4.5.3)$$

$$P_k = \sum_{i \in \{i | t_{k-1} \le HSA_i < t_k\}} \left(1 - \sum_j x_{ij} \right) \cdot Pax_i \qquad \forall k \in \{1..K\} \qquad (4.5.4)$$

This function includes both passenger delay cost and cost of utilizing the alternative hub airport. The components of the objective function are described as follows.

$\sum_k w_k \cdot P_k$ represents the delay of passengers on flights that land at the original hub airport, where w_k is the average delay of flights landed at the original hub airport during time period k, and P_k is the total number of passengers on those flights, obtained from Equation (4.5.4). The calculation of w_k will be elaborated in next section.

The term $\sum_i \sum_j x_{ij} BT_{ij} Pax_i$ represents extra ground transportation time for transfer passengers whose flights are diverted to alternative hubs, where x_{ij} is a decision variable indicating whether Flight i is diverted to an alternative hub j, BT_{ij} is the ground transportation time for passengers on Flight i from the alternative hub j to the hub airport, and Pax_i is the number of passengers on Flight i.

The term $\sum_i \sum_j x_{ij} TPax_i Mis_i$ in (4.5.1) represents the estimated misconnection cost for transfer passengers. It depends on $TPax_i$, the number of transfer passengers on Flight i, and Mis_i, the estimated unit penalty of missing connections.

The term $\sum_j C_j y_j$ represents the cost of utilizing airports as alternative hubs where C_j is a cost coefficient for using airport j and y_j is a decision variable indicating that airport j is used as an alternative hub.

Constraints

The minimization of the objection function (4.5.1) is subject to the following constraints:

1. A flight can only be diverted to alternative hub j where the runway length at airport j satisfies the landing requirement

$$x_{ij} = 0 \qquad \forall \Lambda_{ij} = 0 \tag{4.5.5}$$

where Λ_{ij} indicates if the runway length requirement is satisfied, 0 otherwise.

2. A flight can be diverted to at most one alternative hub

$$\sum_j x_{ij} \leq 1 \qquad \forall i \in \mathrm{I} \tag{4.5.6}$$

3. Flights can only be diverted to alternative hub j if airport j is used as an alternative hub

$$\sum_i x_{ij} \leq M \cdot y_j \qquad \forall j \in \Gamma \tag{4.5.7}$$

4. The total number of diverted flights to alternative hub j cannot exceed the excess capacity at the alternative hub

$$\sum_{i\in\{\mathrm{I}|n-1\le HSA_i<n\}} x_{ij} \le ECap_{n\,j} \qquad \forall j\in \Gamma\, \forall n\in \mathrm{N} \tag{4.5.8}$$

5. The remaining slots at the major hub should be distributed to airlines proportional to their original schedules

$$\sum_{n-1\le HSA_i\le n}\left(1-\sum_{j} x_{ij}\right)\cdot \Delta_i^a \ge Floor\left(\sum_{n-1\le HSA_i\le n}\left(1-\sum_{j} x_{ij}\right)\cdot \phi_n^a\right) \qquad \forall a\in \mathrm{A}\, \forall n\in \mathrm{N} \tag{4.5.9}$$

In the first step of a regional GDP, Constraint 5 assumes that airlines will conform to the model solution. In reality; however, airlines may not use the slots assigned at the alternative hub. In that case, relative x_{ij} s which equal to 1 in the solution will be set as 0 and the model has to be rerun to get updated solution and objective function value.

4.5.2. Continuous Approximation Delay Estimation Method

Average delay of flights that have landed at the original hub airport during time period k , w_k (see Equation 4.5.2), is calculated using the continuous approximation method as shown in Figure 4-9. As can be seen, the cumulative number of scheduled arrival flights is approximated as a continuous curve on the left. The piecewise line on the right represents cumulative number of arrivals restricted by a capacity shortfall at a hub airport. To obtain a closed form for flight delay, the time of day is divided into a finite set of time periods of equal duration. For flights whose scheduled time is in time period k , the flight delay is either $\min\left(\frac{D_k}{c_I}-t_k, \frac{D_k-c_I T_I}{c_V}-(t_k-T_I)\right)$ or

$\max\left(0, \frac{D_k - c_I T_I}{c_V} - (t_k - T_I)\right)$ depending on whether the time period k is before or after the time period when the capacity recovers. In these expressions, D_k is the cumulative number of arrival flights up to time period k, c_I is the capacity level during disruption, c_V is the normal capacity level, and T_I is the length of capacity shortfall.

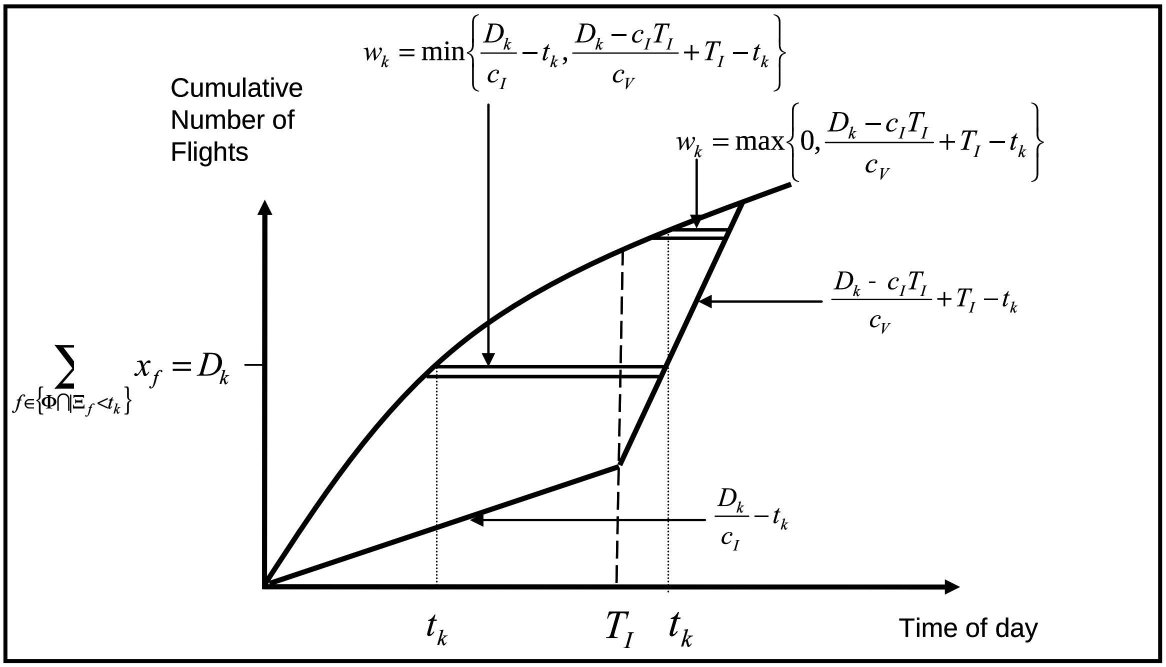

Figure 4- 9 Illustration of Delay Continuous Approximation

To test the performance of the continuous approximation method, mean absolute differences of average flight delay between an exact calculation method and continuous approximation are compared. The experiment is constructed as follows: (1) a series of daily arrival rates is assumed, from 50 to 300 arrivals per day; (2) the sequence of arrivals in one day is created randomly, and 20 repetitions are used for each daily arrival rate; (3) the time of day is divided into two finite sets of time periods with equal durations of 2 minutes and 5 minutes. The results demonstrate, as shown in Figure 4-10, the mean absolute difference decreases when the daily arrival rate increases. It becomes less than 10 seconds when the daily arrival rate reaches 150 arrivals per day. Daily arrival rates at all large hub airports in the U.S. are more than 150, hence, calculated delays from the continuous approximation method are close to those from the exact solution. This method provides a closed form that can be used in the Regional GDP model described above.

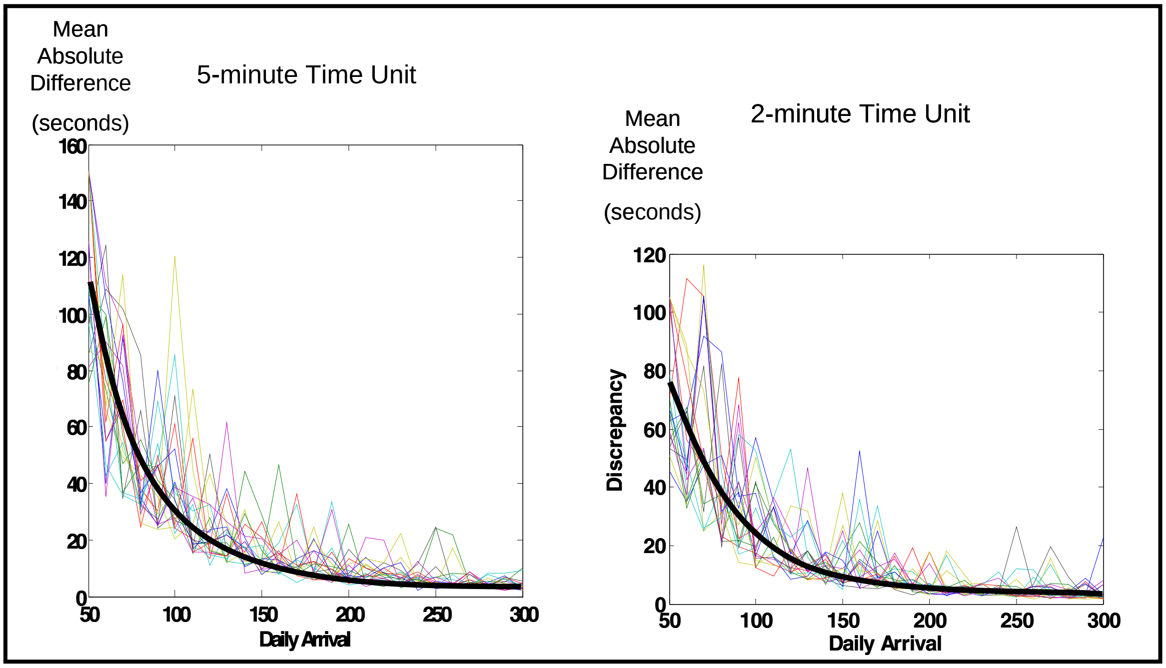

Figure 4- 10 Performance of Continuous Delay Approximation

A daily arrival sequence containing 343 scheduled flights are obtained. It is assumed that the capacity at a major hub drops to 10 flights per hour for 6 hours and recovers to 30 flights per hour. The value of the objective function for without diversion is about 57 thousand passenger-hours. The optimal objective function of Regional GDP model is about 30 thousand passenger-hours with 26 diversions. In the model, misconnection cost for transfer passengers is estimated based on $TPax_i$, the number of transfer passengers on flight i , and Mis_i , the estimated unit penalty of missing connections. This is a worst case estimation because in reality, not all transfer passengers would miss their connections. If we adjust the model assuming there is no misconnection, the optimal objective function would be about 27 thousand passenger-hours with 30 diversions. These two solutions represent upper and lower bounds for a solution with a more realistic value for the misconnection penalty, which could be obtained from experience implementing Regional GDPs. The shadow price of utilizing one more hourly slot at an alternative hub is 521 passenger-hours for the first slot, and 383 passenger-hours for the second slot, and with little benefit for subsequent slots afterwards. Once the schedule information at the alternative hub is given, the impact of transferring hourly slots from the regular schedule to diverted traffic can be estimated and compared with the above results to see if it is cost-effective to change the number of slots available to major-hub-bound flights at the alternative hub.

4.6. Summary

In this chapter, we extend the concept of real-time inter-modalism to regional airport systems. In addition to substituting short-haul flights with ground transport modes to a major hub airport during disruptions, we propose diverting major-hub-bound flights to nearby alternative hubs and reallocating passengers between the major hub and alternative hubs via ground transport modes. This strategy allows transfer passengers to connect at alternative hubs, as well as aircraft swapping at both the major hub airport and alternative hubs. It reduces aircraft reallocations after disruptions, reduces disruption costs, and helps airlines recover from disruption more promptly.

We reviewed the literature on flight diversions, reliever hubs, and other methodologies for airline disruption management. We have also discussed the criteria of selecting alternative hubs in a regional airport system. As a proof of concept, we demonstrated the computational feasibility of our model for solving a real-world disruption problem with the approximation solution algorithm proposed in Chapter 3. We used the schedule data and load factor data for a major hub airport from public accessible database maintained by federal agencies, and synthesized passenger itineraries based on that. Comparisons of case study with RTIMS and RTIMD show that RTIMD has better performance in terms of reducing the total disruption cost and leads to lower numbers of disrupted passengers.

To avoid alternative hubs being overwhelmed by diverted major-hub-bound flights, a Regional GDP is proposed to enhance the current system to provide CTAs at not only the major hub airport but also the regional airports. For providing a Regional GDP advisory, a mathematical programming model was proposed and demonstrated.

CHAPTER 5: IMPLEMENTAION ISSUES OF REAL-TIME INTER-MODALISM

To implement the real-time inter-modal strategies proposed in this study, several fundamental issues need to be considered. The goals of this chapter are to identify these issues, assess their importance and difficulty, suggest solutions based on preliminary investigation, and highlight needs for further research and policy decision making.

5.1. Motor coach Charter Service Provision

5.1.1. Entities Involved and Their Roles

There are different roles in implementing inter-modal operations and various entities can fulfill these roles. The most likely options are summarized in Table 5-1. Airlines will undoubtedly be the parties that determine the need for and initiate inter-modal operations. Then, in supplying motor coach service, there are two roles: owning vehicles and operating/maintaining vehicles. Airlines could purchase a fleet of vehicles and dedicate them for inter-modal operations so that they can furnish the vehicles with amenities such as more luggage space, complementary beverage service, and entertainment, appropriate for their particular clientele. These featured services can help airlines retain passengers' loyalty and reduce their customers' perceived costs of inter-modal diversion.

Airlines could also operate the fleet by themselves so that they can have the control of vehicles dispatch, ensure the integration of airside and landside operations, and coordinate ground stations at the hub and spoke airports. Despite these advantages, this option may not be economical because the strategies proposed in this study are for days with severe capacity shortfalls, which occur infrequently. A key to overcome this would be to control labor cost, a major component of motor coach operating cost. To do so, airlines could contract with part-time drivers that serve on-call while leaving the vehicle fleet idle—or perhaps offering charter service—on the days without inter-modal operations. It may even be possible to use airline personnel—with suitable training—as part-time drivers during disruptions. Airlines could also outsource the operating and maintenance to a third party with access to professional charter services and who agrees to give top priority to airline inter-modal service requests.

Table 5- 1 Alternatives for Inter-modal Service

	Alternative 1	Alternative 2	Alternative 3
Initiate inter-modal service	Airlines	Airlines	Airlines
Own vehicles	Airlines	Airlines	Third party
Operate and maintain vehicles	Airlines	Third party	Third party
Provide airport facilities	Airport	Airport	Airport

A third alternative would be for airlines to contract with existing local charter companies and order services when they are needed. In this scenario, the local companies would own, operate, and maintain vehicles for inter-modal operations. This would require much less effort and investment on the part of airlines, and would enable them to experiment with real-time inter-modal concepts without making a large commitment. A major concern for this approach is the ability of the charter operators to respond in a timely manner to airline service request. We consider this issue in the next section.

5.1.2. Supply Capacity and Lead Time of Motor coach Service

If airlines contract with local companies for motor coach service, the ability of such companies to respond in a timely manner to requests for service, which will be inherently urgent and unpredictable, will be critical. To assess this capability, we used internet sources to identify supply characteristics of local charter companies and conducted a telephone survey to determine response time of charter companies to the type of urgent request that would arise with real-time inter-modal strategies. In addition, a sample of GDP logs was analyzed so that charter response time could be compared with the lead times that would be available to airlines making the requests.

Table 5-2 lists the numbers of charter companies in six metropolitan regions offering different type of vehicles. The seating capacity of deluxe motor coaches ranges from 36 to 68 passengers. The motor coaches have shelf and belly space to accommodate luggage. Restrooms and air conditioning are standard equipment of the motor coach. It is also very common to have entertainment equipment such as TV/VCR. Table 5-1 also lists the numbers of companies who can provide executive coaches or limo buses. In comparison to deluxe motor coaches, executive coaches and limo buses are characterized by plush perimeter seating, tables, TV monitors, and on-board concierge to serve drinks from the on-board bar.

Table 5- 2 Numbers of Charter Companies

	Range of Seating	San Francisco	Los Angeles	New York	Chicago	Miami	Texas
All Type		100	188	260	153	88	114
Deluxe Motorcoach	36-68	63	113	189	83	59	80
Executive Coach	18-30	1	5	11	9	4	6
Limo Bus	18-30	10	9	34	15	10	13

Source: BusRates.com, accessed in May 2008.

As mentioned in Chapter 1, we interviewed a customer service manager of United Airlines at ORD about her experiences working with local charter companies. The manager stated that United could obtain motor coaches within one hour of making the request. To get a more general idea about how promptly charter companies could respond to service requests, we conducted a telephone survey for ten randomly picked charter companies in each of the six regions in Table 5-2. We constructed a scenario motivating an urgent request for motor coach service at an airport. We asked for a motor coach that can accommodate at least 30 passengers and their personal belongings and be available for at least 6 hours. The results of the survey are shown in Table 5-3. About 30 percent of the charter companies did not have vehicles or drivers available at the time of the request. For companies who can provide the service, half of them can get to the airport within one and half hours, while the other half need about three to four hours. The longer lead time is not due to the unavailability of vehicles but the time required obtaining operators. Many charter companies hire part-time drivers and schedule their work load according to reservations. For urgent requests, they need to check the availability of drivers and reorganize their shifts. Some charter companies asked if the surveyor had a business relationship with their company, implying that they might respond more quickly if this were the case.

Table 5- 3 Charter Company's Response to Urgent Service Request

	San Francisco	Los Angeles	New York	Chicago	Miami	Texas
	SFO	LAX	JFK	ORD	MIA	DFW
Not available	3	2	3	5	4	4
1-1.5 hours	2	3	4	2	3	2
3-4 hours	4	5	3	3	3	4
Total	9	10	10	10	10	10

We next investigated whether these response times would be adequate in light of the lead times that would be available to airlines implementing real-time inter-modal strategies. Logs for a sample of 27 GDPs at seven major airports in the years 2006 and 2007 were examined. GDP durations versus GDP lead time, i.e. the time difference when a GDP was issued and when it was activated, are plotted in Figure 5-1. For the 13 GDPs in the circle, the average lead time is about two hours (124 minutes). Curiously, there are 14 GDPs with negative or zero lead times. To further investigate these cases, we consulted with GDP experts at Metron Aviation, the major developer of CDM and the source of the GDP logs.

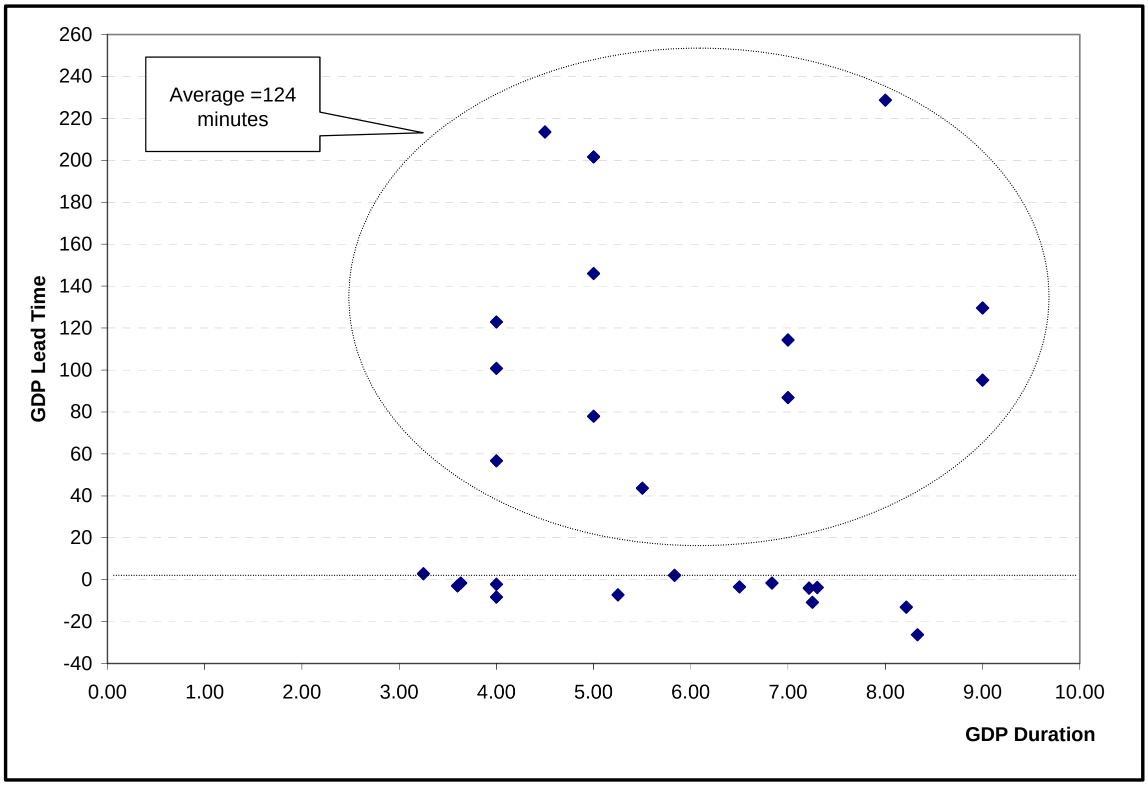

Figure 5- 1 Historical GDP Duration and Lead Time

According to the GDP experts at Metron, the GDPs with zero lead time occurred because airport weather conditions changed suddenly and the GDPs needed to be implemented right away. Negative GDP lead times resulted from the time lapse after a traffic specialist modeled the program but before he sent it out. Nevertheless, GDP specialists running the program have a "now_plus" parameter that they can set. By default, it is set to 45 minutes, which means that when a GDP is issued, flights within 45 minutes of departing are exempted. This policy recognizes that within 45 minutes of departure, passengers are probably already boarding. The GDP experts also noted that, while lead time is a good measure of how much notice the airlines received, airlines take actions during the course of GDP, even if it begins before they can have coach service in place. They also observed that GDPs—even those with little or no lead time according to the GDP logs—are not a total surprise to airlines because the likelihood of their occurrence at various airports is usually discussed on morning telephone conferences between the command center, traffic centers, and airlines. This is particular encouraging for the practicability of implementing real-time inter-modal strategies.

In sum, based on the survey of motor coach charter companies and investigation GDP lead time and durations, we conclude that airlines would have enough time to arrange motor coach service in most cases, even when they have to contract with local charter companies to do so.

5.2. Passenger Acceptance and Communication

Considerable resistance is expected from at least some of the passengers who will be reassigned to motor coaches or any other ground transport modes in a real-time inter-modal strategy. In comparison to flying, using ground transport modes is considered inferior in terms of speed and—in some cases—level of service. When confronted with such resistance, airlines could tell passengers that they are being given an extra option and retain the right to be reassigned to a later flight instead if they choose so. Passengers who opt for this, however, should be accommodated in a manner that does not greatly increase disruption costs over those obtained from the optimum solutions of the inter-modal models presented in Chapters 3 and 4. The seemingly contradictory goals of controlling costs and affording flexibility to passengers can be reconciled in various ways.

For example, if a passenger at a spoke airport chooses to drive to the hub airport, then assuming his/her driving time is similar to that of the motor coaches, the passenger inflow at the hub airport will about the same as if the passenger accepted the reassignment. Hence, the disruption cost will not be substantially affected in this scenario.

If the passenger insists on waiting at the spoke airport for a flight instead of being reassigned to the motor coach, the airline policy can be to rebook him/her so that the cost incurred is no more than that of reassigning him/her on the motor coach. This may well result in a very long wait so that the airline can first accommodate passengers who have been involuntarily disrupted. Indeed, the airline can use the prospect for such a lengthy delay as an inducement for the passenger to accept the motor coach substitute. This is possible because airlines have limited obligations to provide compensations to passengers if their disruption is caused by adverse weather. Many passengers are aware of this and would thus accept the substitute ground transport as the best option for completing their itinerary. This is illustrated in Figure 5-2, a modified version of Figure 4-9, in which each square represents an arrival flight substituted by ground transportation arriving at the original hub airport. In comparison to arrival times if nothing is done to deal with the capacity shortfall, most of those shifted to surface transport are better off. The few counterexamples are highlighted by a circle on the chart. Airlines may provide in-kind or monetary compensation for those passengers who are made worse-off in order to increase the system efficiency of the disruption recovery.

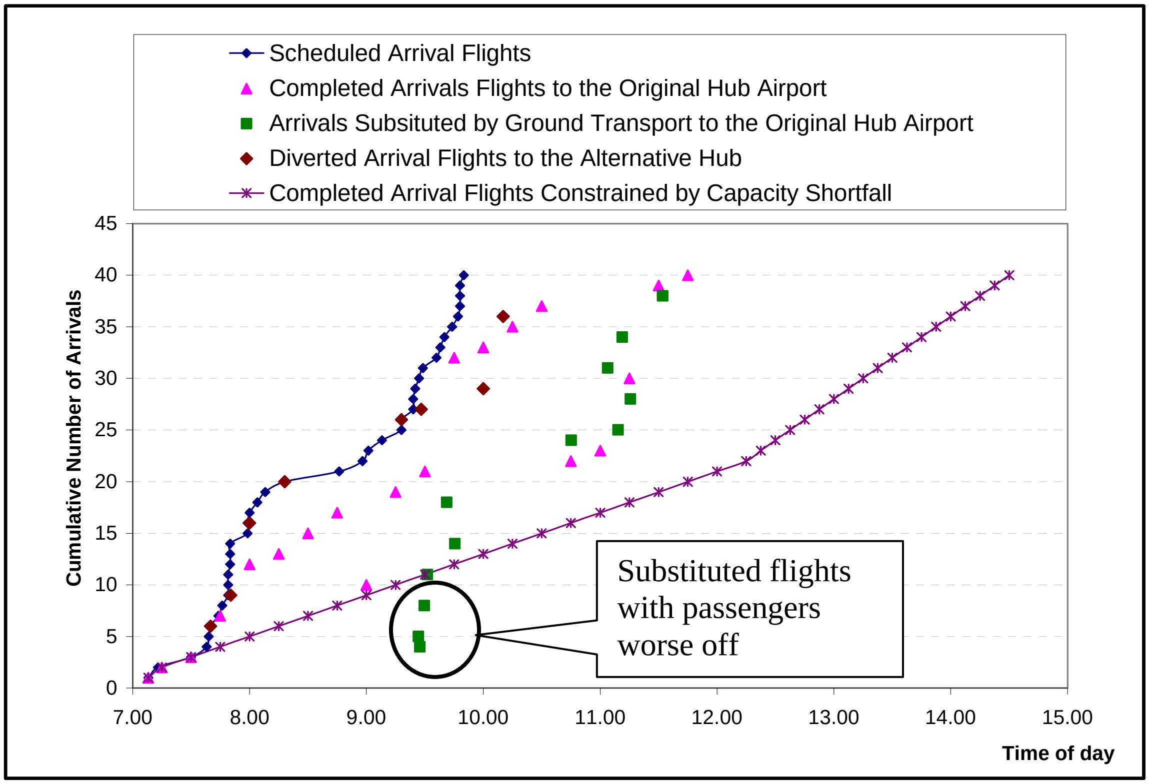

Figure 5- 2 Evolutions of Arrival Flights with RTIMD and Substituted Flights with Passengers Worse Off

A final important passenger-related issue is communication. Airlines now have initiated programs such as "EasyUpdate" used by UA to inform passengers with real-time updated flight information via email, text message, pager, or phone. Such programs, with regular promotion to encourage subscribers, will provide airlines a convenient channel to communicate with passengers during disruptions.

5.3. Security Issue

Real-time inter-modalism introduces a new class of passengers (and passenger baggage) into the airport system—those who have been reassigned from flights to ground transport services. This raises the question of how these passengers (and their bags) will be processed through security. Since many reassignments will occur at the last minute, it is reasonable to assume that reassigned passengers will have gone through security processing at their origin airport. The main issue is how they will be processed when they reach the hub, where, had they flown, they would arrive in the secure area and not require further screening. There are two options. Option 1 is to locate inter-modal passengers' loading and unloading inside the airport secure area. The secure area, also called sterile area, is the part of the airport in which only authorized airport and airline employees, and passengers who have gone through security screening, are allowed. Transfer passengers discharged from flights enter into this area so that they do not have to be re-screened prior to boarding their connecting flight. To implement Option 1, motor coaches carrying airline passengers must be screened before entering the secure area. The screening, which may require advanced sensing technology, will ensure that the motor coach has not been "breeched" since the time it was loaded at the spoke airport. Thus, with this option, passengers can go through security screening at the spoke just as they would in regular operations. Thus no extra screening is required for transfer passengers and their luggage.

Option 2 sets up loading and unloading areas outside the secure area. Transfer passengers will have to go through the security screening at the hub airport, in many cases after already having done so at the spoke airport. From the point of view of the Transportation Security Administration (TSA), Option 1 does not increase workload but may enhance the security risk, while Option 2 increases workload but eliminates any threat associated with motor coaches entering the secure area. Airlines, the airport, and TSA would all have important roles in deciding which of these options to pursue and how to implement substitution securely.

5.4. Airport Facility Issue

5.4.1. Airport Facility Requirement

Concerns arise from ground traffic management centers and airports about how the real-time inter-modalism will affect local traffic, especially local ground access to airports. The results of the numerical examples in Chapter 4 show that among 40 scheduled arrival flights over a three-hour period, 12 of them are substituted by ground transportation. For SFO, the average aircraft capacity is 46 seats per short-haul flight. A load factor of 85% will lead to about 40 passengers, which is about the amount of passengers that can be accommodated by a deluxe motor coach or two executive coaches or limo buses. In addition, there will be about 400 passengers that need to be transported between the original and alternative hubs on each direction. Because the distance between the hubs is shorter, motor coaches used to transport passengers on this route can turn around faster. Hence, 12 motor coaches are estimated to be needed for short-haul flight substitution and 5 for inter-hub connections. This volume of traffic flow will have trivial impact on local traffic and ground access to the airport.

Regardless of whether Option 1 or 2 in Section 5.3 is chosen, a parking area is needed to accommodate the inter-modal motor coach operations. Based on our estimates of the number of coach trips required and assuming an average motor coach size of 46 feet by

13 feet, a parking lot about 20,000 square feet is needed to accommodate the 17 motor coaches simultaneously. Moreover, a passenger walkway is needed to connect the parking lot and terminal entrance. Similar facilities but of much smaller sizes are needed at spoke airports. If airport authorities want to support real-time inter-modal operations, they need to invest capital to provide such facilities. In next section, we will discuss funding source of capital improvement program (CIP) for provide the required facilities.

5.4.2. Funding Source of CIP for Inter-modal Facilities

Airport CIP funding comes from different sources as illustrated in Figure 5-3. A major part of CIP, 58 percent, is funded by bond proceeds backed by general airport revenue or passenger facility charge (PFC) revenues or both. Other than that, 21 percent of CIP projects are funded by federal airport improvement program (AIP) funds and 11 percent by PFC revenues collected at the airport. The eligibility of projects for using AIP and PFC funds is summarized Figure 5-4. In comparison, AIP funds are more commonly used for airside projects, while PFC funds go toward landside and paying bond debt service with only 18 percent of these funds used for airside projects.

As stated by Secretary Mary E. Peters, transportation system congestion is one of the single largest threats to US economic prosperity and way of life. Back in 2003, Americans lost $9.4 billion as a result of airline delays. With demand going up and fuel prices soaring to record highs, the delay cost of air transportation congestion continues to grow. The National Strategy to Reduce Congestion on America's Transportation Network (National Strategy), announced in May 2003, encourages federal agencies and states to tap private sector resources and expertise to improve transportation systems. A new response to aviation congestion included in the National Strategy is to give priority treatment and agency resources to projects and technologies that enhance aviation system capacity. Real-time inter-modalism is a framework that can enhance aviation system capacity by (1) integrating ground transportation in the aviation system when there are airside capacity shortfalls; (2) utilizing excess capacities at secondary or third airports in regional airport systems. It therefore qualifies for the priority treatment included in the National Strategy. Thus, airports seeking funds for inter-modal facilities can make this argument and increase their prospects for obtaining Federal AIP funds, or approval for using PFC revenues, to pay the project cost of providing the required facilities.

.

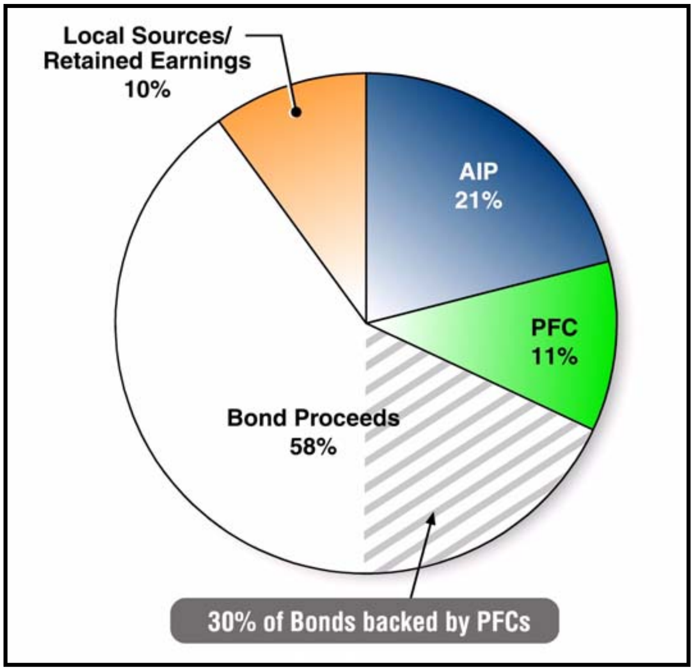

Figure 5- 3 Funding Sources of Airport CIP[22]

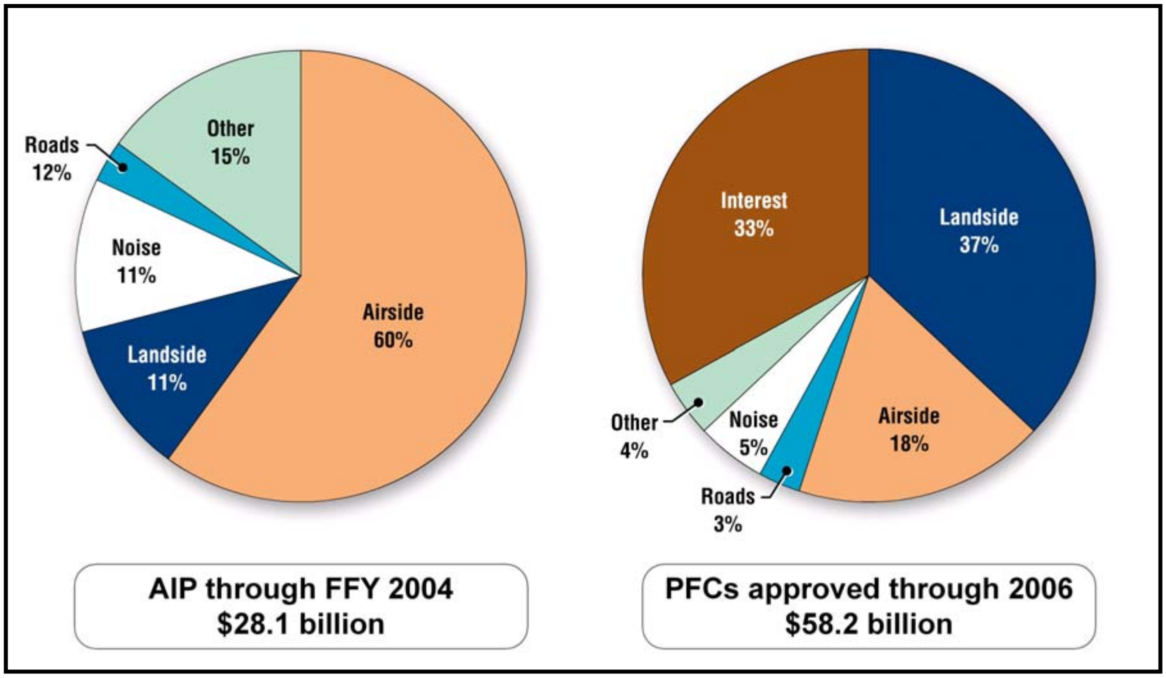

Figure 5- 4 AIP vs. PFC Funded Projects[22]

5.5. Summary

This chapter discussed some of the most critical issues that must be faced when implementing real-time inter-modalism for airlines disruption management. These issues cannot be solved on a purely technical level but also require policy making. Table 5-4 summarizes these issues and their degree of urgency, indicated by two levels: fairly urgent and urgent. Fairly urgent issues include the ownership of coach fleet, maintenance and operating the vehicles, and passenger compensation. Airlines need to figure out which alternative, as discussed in Section 5.1, would be the best. However, as the first step, they can obtain motor coach services from local charter companies. For passenger

---

[22] Source: Tomoson Financial, FAA, and ACI-NA

compensation, airlines need to obtain a better understanding of passengers' preferences through several real-time inter-modal operation experiments before they set up a compensation procedure. Airlines also need time to incorporate the procedure into their existing passenger reassignment program. The other issues listed in Table 5-4 are identified as urgent because they must be solved before inter-modal operations can be implemented.

Table 5- 4 Summary of Implementation Issues

	Issue	Degree of Urgency
Motorcoach		
Ownership	Airlines or a third party?	Fairly urgent
Operating and maintenance	Airlines or a third party?	Fairly urgent
Passenger		
Compensation	Who should be compensated and how?	Fairly urgent
Communication	What program should be set-up and how to increase the subscription?	Urgent
Security		
Loading and unloading	Inside secured area or outside?	Urgent
Facility		
Parking Lot and Walkway	What size?	Urgent
Funding Source	Where to get the funding?	Urgent

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

6.1. Conclusions

This study proposes a new strategy, featuring the use of surface transport both as a substitute mode and an enabler for the use of alternative hubs, for airline recovery from schedule perturbation caused by adverse weather or other temporary events. Chapter 1 investigated causes of airside capacity shortfalls, discussed the consequences, and introduced current airline operations and existing air traffic flow management systems. Based on this information, it suggested utilizing other transportation modes and argued the possibility of integrating them into the air transportation network. A thorough literature review on airline disruption recovery and airport congestion management was presented in Chapter 2.

Capacity reduction at a hub airport caused by temporary events is reflected as reduced slots (or later controlled time of arrivals [CTAs]) for airlines in Ground Delay Programs (GDPs). Chapter 3 suggests that, under these circumstances, airlines with hub-and-spoke networks should integrate ground transportation into their operations, replacing flights in order to reduce disruption cost, passenger delays, and passenger misconnections. This strategy is termed Real-time Inter-Modal Substitution (RTIMS). RTIMS is different from the air-rail cooperation practiced in Europe because it is only triggered by severe demand and supply imbalance at major hub airports and it consists of operational integration of airside and ground transportation. Chapters 3 analyzed inter-modal substitution for a simple hub-and-spoke network. We first investigated the impact of short-haul flight cancellation and substitution on an arrival sequence at SFO. We determined the number of flights for which arrival time at the destination could be hastened by substituting a slower surface mode. We also considered the impact of inter-modal substitution on total flight delay and the longer individual flight delays. We proposed a mathematical programming model to help a single airline to implement the RTIMS. The complexity of the model was discussed and an approximation algorithm was suggested to reduce computation time so that decisions can be made in real time. Numerical examples are constructed to explore the performance of the mathematical model. The model was then applied to a set of scenarios designed to reveal its sensitivity to various inputs. Key finding are as follows:

- In our baseline experiment, RTIMS reduced the cost of recovering from a disruption by about 40 percent compared to a system without inter-modal substitution.
- RTIMS yields the greatest reduction in cost compared to a more conventional strategy when the capacity shortfall at a hub airport is severe in terms of magnitude or duration.
- The value of the objective function varies with different passenger value of time.
- The distance between short-haul spoke airports and the hub airport is critical for implementing RTIMS, with a three hour driving time the apparent maximum for inter-modal substitution to be a useful strategy.
- Benefits of implementing RTIMS diminish when the load factor goes higher.

- While flight schedule characteristics and transfer passengers' connection patterns affect the outcomes of the model, the percentage of cost saving from RTIMS in comparison to a more conventional strategy is not very sensitive to these factors.

In Chapter 4, real-time inter-modalism was extended to regional airport systems. For a major airport located in a regional airport system, the inter-modal strategy includes diverting major-hub-airport-bound flights to nearby airports (alternative hubs) in addition to substituting short-haul flights. Decisions about delaying, flying, canceling, substituting, and diverting flights could be determined with the use of a mathematical programming model that takes a comprehensive disruption cost as its objective function. The model was run for a case study. Various capacity shortfall scenarios and ground connection assumptions were constructed to compare the performance of RTIMS and RTIMD. Key findings are as follows:

- In comparison to RTIMS, RTIMD results to lower disruption costs due to fewer flight cancellations, substitutions, and disrupted passengers. For the baseline case, however, the reduction was only about four percent.
- The reduction in disruption costs from utilizing RTIMD instead of RTIMS increases with the severity of the capacity shortfall.
- Restraining the travel time between hubs to one hour or less is essential in implementing RTIMD.

Considering the possibility of alternative hubs being overwhelmed by diverted flights, Chapter 4 also suggested an enhancement to the current GDP so that not only the CTAs at a major hub airport but also those at airports in a regional airport system could be determined. This enhanced GDP was termed Regional GDP.

Although the benefits of real-time inter-modalism are projected to be promising, several important issues need to be addressed in order to implement the strategies. These issues were discussed in Chapter 5. It was suggested that airlines can own, maintain, and operate a motor coach fleet, in which case they would have full control of the configuration and dispatching of vehicles and could respond to urgent needs of inter-modal operations promptly. It was also suggested that airlines can contract with local charter companies, an alternative that would entail less risk. Based on a survey of motor coach charter companies and investigation of GDP lead time and duration, we conclude that airlines responding to a temporary capacity shortfall would have enough time to arrange motor coach service from a local charter company in most cases. Passenger related issues, such as compensation and communication, were also discussed in Chapter 5, as well as security issues. Furthermore, airport facility needs for accommodating inter-modal operations were estimated. Funding sources for supporting these facilities were suggested from the point of view of airport finance. At the end of Chapter 5, implementation issues were summarized and their degrees of urgency were identified.

6.2. Recommendations for Further Research

This study proposed a new approach to alleviating delay and disruption in the air transportation network and provides the ground work for the decision support tools required to efficiently employ this strategy. There remain many significant challenges and research questions to address before implementing real-time inter-modal strategies. This section lists some of the problems for future research.

6.2.1. Airlines' Massive Cancellations

Concerned about possible faulty wiring that could cause a short-circuit or even cause a fire and explosion, the FAA suddenly grounded hundreds of MD-80[23] aircraft operated by airlines such as United Airlines, American Airlines, and Delta in March and April of 2008. This emergency inspection led airlines to cancel thousands of flights for several days, which affected more than one million passengers. For instance, on April 9th, 2008, American Airlines cancelled 1,094 flights, about half the size of their daily operation[24]. More than 100,000 travelers were estimated to be affected. Passengers were stranded at the terminals or in hotels in the vicinity of airports. It took days for airlines to reassign passengers and recover to normal schedules. While the circumstances surrounding this event were unique, other similar situations are quite conceivable, including widespread airport closures due to storms, or a temporary system-wide shutdown due to the perceived threat, or actual occurrence of a 9/11-style terrorist attack.

Real-time inter-modalism can be used to alleviate disruption costs caused by this kind of network-wide massive cancellation situation by replacing MD-80 with other type of

---

[23] The Boeing MD-80 is a quiet, fuel-efficient twinjet, certified by the Federal Aviation Administration in August 1980 and entered airline service in October 1980. A standard two-class configuration of MD-80 supplies 152 seats.

[24] Yahoo! Finance, http://biz.yahoo.com/ap/080409/american_airlines_cancellations.html, accessed in may 2008.

aircraft and substituting flights with ground transport services. A suitable replacement of the MD-80 is Boeing 737-400, which supplies 146 seats with a two-class configuration. Other aircraft with capacity from 100 to 150 could also be the replacement contingent on the number of passengers who have purchased the itineraries for the flight scheduled to use the Boeing MD-80. The model proposed in this study can be extended to a network with multiple hubs. Given the aircraft that have to be cancelled from a daily schedule, airlines would decide which flight to cancel, which to substitute with ground transport, how to swap aircraft, and how to reassign passengers on a massive scale.

6.2.2. Other Strategies to Manage Computational Complexity

Section 3.4 discussed the complexity of the mathematical programming model and proposed an algorithm to obtain solution that approximate optimality. The approximation algorithm reduces computation time from hours to minutes for solving numerical examples with flights scheduled in a 4-hour window. As the first step of the approximation algorithm, relaxed nonlinear programming is solved by using the solver MINLP on a remote server. Other, more robust or sophisticated strategies to manage computation complexity may yield better results with little additional computation time, or allow large problems with longer time windows to be solved.

6.2.3. Passenger Attitudes and Response

Passengers’ resistances to inter-modal substitution, and possible solutions to it, have been elaborated in Chapter 5. For real-time inter-modalism to be acceptable, passengers should have the right to accept or reject the motor coach service. Although their decisions could probably be accommodated so as not to affect the solutions of the models, their preferences for conventional or innovative disruption management are still critical for the framework to work successfully. Hence, it is helpful to conduct a carefully designed survey to understand passengers' preferences.

Most surveys ask questions and expect answers from passengers based on their own experience. In this case, however, passengers would have to state their preferences in situations that they may have not encountered before. We suggest combining survey questions with customer service feedback systems and distributing the forms to passengers when they are encountering disruptions. In comparison to general passengers who may have never gone through any disruptions, these passengers are presumably more interested in responding to the survey form and have incentives to answer carefully. The possible questions in a survey form should include, besides basic descriptive statistics, travelers' prior experience with irregular operations, such as delay duration and types of reassignment, and the perceived degree of inconvenience. Finally, hypothetical stated preference questions would be used to assess the wiliness of passengers to be reassigned to motor coach under different delay scenarios. Passengers would also be asked about features of the motor coach service that may affect their acceptance, such as voucher for a future airline reservation, complementary beverage and meals, entertainment, wireless internet access, comfortable seats, and so on.

6.2.4. Enhancement of CDM

Ground delay programs (GDPs) in the current system evaluate the imbalance between supply and demand at one airport and provide airline with CTAs at that airport for airlines scheduled flights. To implement RTIMS in the most efficacious manner, CTAs at not only the major hub airport but also airports in a regional airport system are needed. Section 4.5 proposed a mathematical programming model to obtain such CTAs assuming FAA acts as the central coordinator. Similar to GDP advisory in the current CDM procedure, the CTAs from the model can be considered as a Regional GDP advisory. There are two options to deal with the possible rejections on utilizing the CTAs at alternative hubs. Option 1, as mentioned in Section 4.5, would take airlines' response and rerun the model with given updated decision variables. Option 2, however, would create a set of airline specific cost parameters—on the consensus of airlines—and build them into the objective function. In this case, airlines would agree not to reject CTAs at the alternative hubs in the Regional GDP advisory while not using the CTAs are allowed.

Correspondingly, the current CDM system needs to be enhanced to realize this regional GDP by including regional transport agencies, regional airport authorities, airlines serving regional airports, and others. The flow chart presented in Chapter 3 should be adjusted as shown in Figure 6-1 with the enhancements highlighted in bold-bordered boxes. These enhancements could not be realized without collaboration between FAA, airlines, airports, and passengers, and consensus on the importance of integrating underutilized regional airports into disruption recovery strategies. It is a challenging task for every entity involved and wisdom is needed to ensure fairness, effectiveness, and efficiencies of the enhanced CDM.

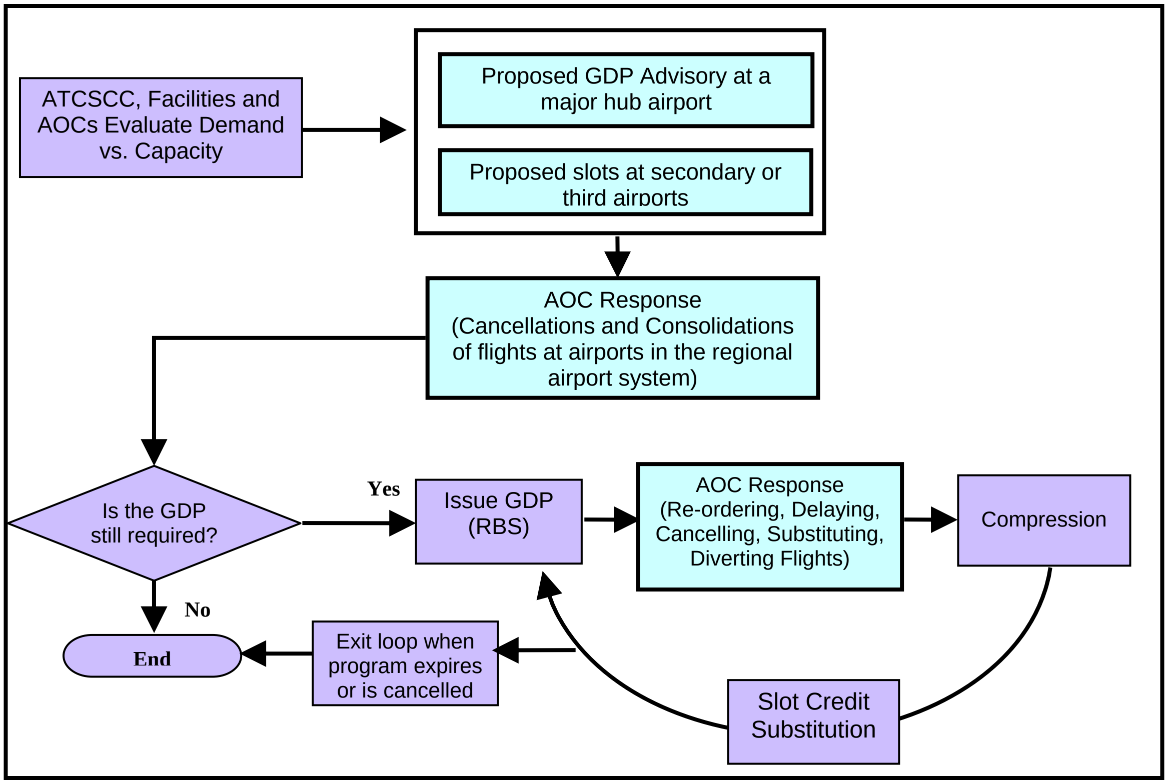

Figure 6- 1 Flow Chart of Enhanced Collaborative Decision Making

The inter-modal framework proposed in this thesis will substantially reduce the costs of recovering from major hub capacity shortfalls by providing alternatives to simple cancellation in airlines' schedule perturbation recovery, thus reducing the number of disrupted passengers and the delay propagated to later flights and other parts of the network. Although it has been introduced as a real-time, temporary-event-triggered operational reaction, the inter-modal framework could be extended and implemented on a daily base. Oil prices were skyrocketed to an all-time high and keep on rising, from about $40 per barrel in 2004 to about $130 per barrel in May 2008. Such high oil prices are harming many industries, especially airlines for which where fuel cost counts for more than one third of operating expenses. In the March and April 2008, five airlines, Aloha, ATA, Skybus, Frontier, and EOS airlines, filed Chapter 11 bankruptcy applications and four of these have discontinued their operations. Airlines are trying different ways to cut the cost and improve their profit margin, such as raising airfares, pursuing acquisition (e.g. Delta and Northwest) and cooperation (e.g. United and Continental), charging for checked bags (e.g. United and American), and charging for the privilege of getting window or aisle seats (e.g. US Airways). We suggest that, in addition to these approaches, airlines should now think about integrating inter-modal framework into their daily operations so that they can reduce frequencies of short-haul flights, which are the least fuel-efficient in terms of passenger-miles and travel time savings. The motor coach charter service proposed in this study could be a good choice for the ground transport mode because they do not require massive capital investment and could be implemented quickly once the issues discussed in Chapter 5 are resolved.

APPENDIX A: THE VALUES OF THE COST PARAMATERS USED IN SOLVING MATHEMATICAL PROGRAMMING MODELS OF RTIMS AND RTIMD

Notation	Description	Value
CostA	Cost of utilizing the alternative hub per diverted landing or taking-off	$1000
CostB	Fixed cost of utilizing ground transportation per substitution	$200
CostBP	Variable cost of utilizing ground transportation per passenger-unit time (unit=10 minutes)	$2.5
CostD	Estimated cost per disrupted passenger	$300
CostF	Airlines' operating cost of delaying a flight for one time unit (unit=10 minutes)	$500
CostP	Passenger delay cost per one time unit (unit=10 minutes)	$5
CostTA	Cost of ferrying remaining aircraft back from the alternative hub to the original hub	$2000
CostTF	Cost of transporting passengers between the original and alternative hubs, which is based on passenger value of time and the driving time between the two airports (driving time=1 hour)	$30

APPENDIX B: THE MATHEMATICAL PROGRAMMING MODEL FOR SYNTHESIZING PASSENER ITINERARIES

Table A.1 list the notation used in this mathematical programming model in alphabetical order.

Table A.1 List of Notations

Notation	**Category**	**Description**
$i \in I$	Input	A set of arrival flights
$ACap_i$	Input	Aircraft capacity of arrival flight i
$DCap_j$	Input	Aircraft capacity of departure flight j
D_j	Input	Destination of departure flight j
O_i	Input	Origin of arrival flight i
AP_i	Input	Percentage of seats on arrival flight i reserved for local passengers destined for the hub airport
DP_j	Input	Percentage of seats on departure flight j reserved for local passengers originated at the hub airport
$T_{i,j}$	Input	Time of transferring between arriving flight i and departing flight j.
$T_{\min}$	Input	Bus transportation time if arrival flight a is substituted by ground transportation mode
$X_{i,j}$	Decision Variable	The number of passengers transferring between arriving flight i and departure flight j.

The objective is to minimize passengers' total travel time, which is equivalent to minimizing passenger transfer time at the hub airport since the airborne time is assumed to be constant among the two airports. Thus, the formulation is as follows.

$$\text{Minimize } \sum_{i,j} X_{i,j} * T_{i,j} \tag{A.1}$$

Subject to;

$$\sum_{\substack{i \in \{O_i = o\} \\ j \in \{D_j = d\}}} X_{i,j} = S_{o,d} \qquad \forall (o,d) \tag{A.2}$$

$$\sum_{j} X_{i,j} \leq ACap_i * (1 - AP_i) \qquad \forall i \tag{A.3}$$

$$\sum_{i} X_{i,j} \leq DCap_j * (1 - DP_j) \qquad \forall j \tag{A.4}$$

$$X_{i,j} \text{ integer} \qquad \forall (i,j) \tag{A.5}$$

The constraint (A.2) is to make sure that all transfer passengers are connected via arriving and departing flights. Constraints (A.3) and (A.4) are aircraft capacity constraints of arriving and departing flight. Constraint (A.5) denotes that numbers of passengers should be integers.

BIBLIOGRAPHY:

Ball, M.O., Hall, W., Hoffman, R., and Rifkin, R. 1998. *Collaborative Decision Making in Air Traffic Management: a Preliminary Assessment*, National Center of Excellence in Aviation Operations Research (NEXTOR), University of Maryland, College Park, MD.

Ball, M.O., Chien, C-Y., Hoffman, R., and Vossen, T. 1999. *Collaborative Decision Making in Air Traffic Management: Current and Future Research Directions*, Proceedings of ATM 99, Workshop on Advanced Technologies and Their Impact on Air Traffic Management in the 21st Century.

Ball, M O. &, Hoffman, R. & Odoni, A. & Rifkin, R. 2003, *Static Integer Program with Dual Network Structure and Its Application to the Ground Holding Problem*, Operations Research, Vol. 51, No. 1, pp 167-171, January-February 2003, INFORMS.

Ball, M.O., Hoffman R., Knorr, D. & Wetherly, J. & Wambsganns, M. 2000. *Assessing the Benefits of Collaborative Decision Making in Air Traffic Management*, Third USA/Europe Air Traffic Management R&D Seminar. Naples, Italy. June 13-16. 2000

Bonnefoy, P. A., and Hansman, R.J. 2005, Emergence of Secondary Airports and Dynamics of Regional Airport Systems in the United States, *Report No. ICAT-2005-02, MIT International Center for Air Transportation, May 2005.*

Barnett, A I. & Paull, G. 2000, *Runway Collisions: Crying Wolf?* INFORMS, San Antonio, 2000.

Bratu S. & Barnhart, C. 2006, *Flight Operations Recovery: New Approaches Considering Passenger Recovery,* Journal of Scheduling, Springer Netherlands Vol. 9, No.3, pp279-298, June 2006.

Cao, J.-M. & Kanafani, A. 1997a. *Real-Time Decision Support for Integration of Airline Flight Cancellation and Delays, Part I: Mathematical Formulation,* Transportation Planning and Technology, Vol. 20, pp183-199, 1997.

Cao, J.-M. & Kanafani, A. 1997b. *Real-Time Decision Support for Integration of Airline Flight Cancellation and Delays, Part II: Algorithm and Computational Experiments*, Transportation Planning and Technology, Vol. 20, pp201-217, 1997.

Carlson, P. 2000. *Exploiting the Opportunities of Collaborative Decision Making, a Model and Efficient Solution Algorithm for Airline Use*, Transportation Science, Vol. 34, No. 4, pp381-393, Nov. 2000.

Carr, F., Clarker, J-P & Duong V. 2002, *Robust Dynamic Routing of Aircraft under Severe Flow Restrictions*, the Proceeding of the 2002 American Control Conference, Vol. 2, pp1314-1319. 2002.

Chinneck, J.W., *Practical Optimization: a Gentle Introduction,* http://www.sce.carleton.ca/faculty/chinneck/po.html.

Czyzyk, J., Mesnier M., and Moré, J. 1998, *The NEOS Server*, IEEE Journal on Computational Science and Engineering, pp68-75, Vol. 5, No. 3, 1998.

Dolan, E., 2001, *The NEOS Server 4.0 Administrative Guide*, Technical Memorandum ANL/MCS-TM-250, Mathematics and Computer Science Division, Argonne National Laboratory, May 2001.

FAA 2007, *Capacity Needs in the National Airspace System*, The MITRE Corporation, May 2007.

Filar, J.A. & Manyem, M. & White, K. 2001. *How Airlines and Airports Recover from Schedule Perturbation, A Survey,* Annals of Operations Research, 108, pp315-333, 2001.

Forman, B.E., Wolfson M.M., Hallowell, R.G., & Moore, M.P. 1999, *Aviation User Needs for Convective Weather Forecasts*, American Meteorological Society 79th Annual Conference, Wyndham Anatole Hotel, Dallas, TX, January 10-15, 1999.

Gerhardt, M. 1999, *Inter-modal Freight Transportation*, 4th Edition. Transportation Foundation and Inter-modal Association of North America.

Goetz, A. R., & Vowles, T. M. 2000, *Progress in Inter-modal Passenger Transportation: Private Sector Initiatives*, Research report of National Center for Inter-modal Transportation at University of Denver.

Gropp, W. & Moré, J. 1997, *Optimization Environments and the NEOS Server, Approximation Theory and Optimization*, M. D. Buhmann and A. Iserles, eds., pages 167-182, Cambridge University Press, 1997.

Hoffman, R. & Ball, M.O. 2000, *A Comparison of Formulations for the Single-airport Ground Holding Problem with Banking Constraints*. Operations Research, Vol. 48, No.4 pp. 578-590, July-August 2000, INFORMS.

Hansen, M. & Zhang, Y., 2005. *Operational Consequences of Alternative Airport Demand Management Policies: Case of LaGuardia Airport in New York,* Transportation Research Board 84th Annual Meeting Compendium of Paper CD-ROM, January 2005,, Washington D.C. USA.

Inniss, T.R. & Ball, M.O. 2002. *Estimating One-Parameter Airport Arrival Capacity Distributions for Air Traffic Flow Management*, NEXTOR report, August 2002.

Irrgang, M.E., *Irregular Airline Operations*, Handbook of Airline Economics (1st edition), McGrawHill Aviation Week Group: New York, pp/ 349-365, 1995.

Jarrah, A. et al. 1993. *A Decision Support Framework for Airline Flight Cancellation and Delays*, Transportation Science Vol. 27, No. 3, pp266-280, 1993.

Kasper, D.M., 2004. *Second Markets & Efficiency in the Allocation of Slots at US Airports*, Working paper.

Karow, M. & Clarke, J.-P. 2002, *Virtual Hubs, A Schedule Recovery Approach*, AGIFORS Airline Operation, Rome, Italy.

Kenyon, S. & Lyons, G. 2003. *The Value of Integrated Multimodal Traveler Information and its Potential Contribution to Modal Change*, Transportation Research Part F: Traffic Psychology and Behavior, Vol. 6, No. 1, March 2003, pp. 1-21.

Knorr, D. 2006, *NAS Operational Analysis and Challenges for Modeling the Future*, National Airspace System Performance Workshop, March 14-17, 2006, Asilomar Conference Center, Pacific Grove, CA.

Kohl, N. et al. 2004. *Airline Disruption Management – Perspectives, Experiences and Outlook*, Technical report, IMM-Technical Report-2004-16, Informatics and Mathematical Modeling, Technical University of Denmark, September 2004.

Liu, P.-C. B. & Hansen, M. 2008, *Managing Uncertainty in the Single Airport Ground Holding Problem*, IBSN-10: 3836484455, VDM Verlag Dr. Mueller, e.k., May 16, 2008.

Luo, S. & Yu, G. 1998a, *Airline Schedule Perturbation Problem: Landing and Takeoff with Nonsplitable Resource for the Ground Delay Program*, Operations Research in the Airline Industry, Kluwer Academic Publishers, Boston 1998.

Luo, S. & Yu, G. 1998b, *Airline Schedule Perturbation Problem: Ground Delay Program with Spittable Resources*, Operations Research in the Airline Industry, Kluwer Academic Publishers, Boston 1998.

Meyer, E. et al. 1997. *Evaluating the Feasibility of Reliever and Floating Hub Concepts when a Primary Airline Hub Experiences Excessive Delays*, Research Report SWUTC/96/467503, Southwest Region University Transportation Center, Center for Transportation Research, The University of Texas at Austin, August 1997.

Milner, J.M. 1995. *Dynamic Slot Allocation with Airline Participation*, Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA, June 1995.

Mukherjee, A. & Hansen, M. 2007, A *Dynamic Stochastic Model for the Single Airport Ground Holding Problem,* Transportation Science, Vol. 41, No. 4, November 2007, pp. 444-456.

NBAA Press Releases 2005. *NBAA Strongly Supports Funding of Minneapolis Reliever Airports*, http://web.nbaa.org/public/news/pr/2003/20030515-026.php, accessed at January 27, 2005.

de Neufville R., and Odoni, Amedeo, (2003), *Airport Systems: Planning, Design, and Management*, McGraw-Hill Company, Inc., 2003

Nilim, A., Ghaoui, L., & Duong, V. 2002, *Robust Dynamic Routing of Aircraft under Uncertainty*, IEE 21st Digital Avionics Systems Conference, 2002.

Nolan, M. S. 1999, *Fundamentals of Air Traffic Control*, Wadswordth, Inc., Belmont, California.

Northwest Airlines, 1991, *Cost per Delay Minute*, Report from Northwest airlines, 1991.

Rakas, J., and Schonfeld, P.M. 1998, *Airspace system performance after equipment outages,* Research Report for FAA Grant No. 96-G-006, Dept. of Civil Engineering, Univ. of Maryland, College Park, MA, 1998.

Rakas, J. and Schonfeld P.M. 2004, *Deterministic Models for Degraded Airside Capacity and Delays*, Journal of Transportation Engineering, ASCE, September/October, 2004.

Rakshit, A. & Krishnamurthy, N. & Yu, G. 1996, System *Operations Advisor: A Real-Time Decision Support System for Managing Airline Operations at United Airlines*, Institute for Operations Research and the Management Sciences, INTERFACES 26:2, pp50-58, March-April 1996.

Richetta, O. & Odoni, A. 1993, *Solving Optimally the Static Ground-Holding Policy Problem in Air Traffic Control.* Transportation Science 27(3), 228-238.

Rifkin, R. 1998, *The static stochastic ground-holding problem*. SM thesis, Massachusetts Institute of Technology.

Shane, J.N. 2006, *Towards More Golden Anniversaries*, AAPA Seaports Magazine, Fall 2006, Vol. 8, pp25-26.

Teodorovic, D. & Guberinic, S., 1984, *Optimal Dispatching Strategy on an airline network after a schedule perturbation,* Europe Journal of Operational Research, Vol. 15, pp178-182, 1984.

Teodorovic, D. & Stojkovic G. 1990, *Model for Operational Daily Airline Scheduling*, Transportation Planning and Technology, Vol. 14, pp273-285, 1990.

Vetrovsky, D. & Kanafani, A. 1994, *The Potential Role of Airports as Inter-modal Terminals: Lessons from International and Domestic Experiences*, Research report, UCB-ITS-RR-94-3, March 1994, Institute of Transportation Study, University of California at Berkeley, USA.

Vranas, P. & Bertsimas, D. & Odoni, A. 1994. *Dynamic Ground-Holding Policies for a Network of Airports*, Transportation Science, Vol. 28, No. 4, Nov. 1994.

Willemain, T.R. 2001, *Contingencies and Cancellations in Ground-Holding Programs*, Working paper, NEXTOR wp-01-3, May 30, 2001.

Yan, S. & Yang, D.-H. 1996, *A Decision Support Framework for Handling Schedule Perturbation*, Transportation Research, Part B, Vol. 30, No. 6, pp405-419, 1996.

Zhang, Y., Menendez M., & Hansen M. 2004, *Analysis of De-peaking Strategies Implemented by American Airlines: Causes and Effects*, the proceeding of annual meeting of Transportation Research Board, 2004.