Virtual Hubs: An Airline Schedule Recovery Concept and Model

By
Michelle J. Karow
B.S. Civil Engineering
University of Illinois at Urbana-Champaign, 2001

SUBMITTED TO THE DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN TRANSPORTATION

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2003

© 2003 Massachusetts Institute of Technology. All rights reserved.

Signature of Author: ____________________________
Department of Civil and Environmental Engineering
May 9, 2003

Certified by: ____________________________
John-Paul Clarke
Associate Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by: ____________________________
Cynthia Barnhart
Professor of Civil and Environmental Engineering
Co-Director, Center for Transportation and Logistics

Accepted by: ____________________________
Oral Buyukozturk
Chairman, Departmental Committee on Graduate Studies
Virtual Hubs: An Airline Schedule Recovery Concept and Model

By
Michelle J. Karow

Submitted to the Department of Civil and Environmental Engineering on May 9, 2003
in partial fulfillment of the requirements for the Degree of
Master of Science in Transportation

ABSTRACT

Inclement weather at an airline’s hub airport can be devastating to that airline’s schedule. The repercussions resonate throughout the airline’s network as capacity is reduced, connections are missed, and passengers are delayed on a larger scale than during irregular operations at a spoke airport. The main hypothesis behind the work presented in this thesis is that by shifting a small fraction of a connecting bank to strategically located, under-utilized airports during irregular operations, an airline can reduce costs and aircraft delays relative to current industry rescheduling practices. These proposed “virtual hubs” would, in addition to hosting selected connecting traffic that is shifted from the original hub in order to maximize passenger flow through the network, also reduce the demand on the nominal hub airport.

The primary goal of this research project was to develop methods for the implementation of a virtual hub network and evaluate the potential benefits to the airline industry. To that end, a mathematical formulation is presented along with a case study of the benefits of a virtual hub to a major US airline. The actual recovered schedule and delay statistics for a day of irregular operations was compared to the results from the virtual hub network. Results indicate that significant passenger delays are reduced 94% and flight cancellations are reduced by 15% when a virtual hub network is implemented.

Thesis Supervisor: John-Paul B. Clarke
Title: Associate Professor
Department of Aeronautics and Astronautics
Acknowledgments

I would first like to thank my advisor, Dr. John-Paul Clarke for all of his help in bringing the virtual hubs project together. J-P. was an endless source of ideas, enthusiasm, and frequent flier miles. Thanks, J-P., for bringing me to MIT, making me an ‘angel’, and all of the amazing opportunities you have given me. I am truly grateful.

Thank you to everyone in the International Center for Air Transportation that helped me with this thesis (whether they knew it or not). Everyone was always willing to answer questions or help me dig up what I needed. I wish you all the best in your future endeavors.

I would also like to thank the members of the Not-So-Secret ICAT Social Society (in alphabetical order, Hayley, Lixia, Nathan, and Tom). I wouldn’t have made it through without your support and friendship. I will really miss our morning chats and nights on the town. Thank you for taking such good care of me.

And of course, I have to say thanks to the rest of J-P.’s Angels, Flora and Laura. Thank you for all of your pep-talks and research advice. I will miss all of our great trips together; it’s been fun to see the world with you.

A special thank you to my roommate Bryna for being a great friend and listener. I will miss our battleship games and all of our evening chats.

My family’s support, love, and blessing has brought me to where I am today. Thank you for everything; I love you all.

Finally, I would like to thank Brian for sticking by me through thick and thin. You are my best friend, and I cannot think of anyone I would rather take this incredible journey with.
# Table of Contents

1 Introduction .......................................................................................................................... 11  
1.1 Motivation ....................................................................................................................... 11  
1.2 Problem Statement ........................................................................................................... 12  
1.3 Previous Work in Schedule Recovery ............................................................................ 13  
1.4 Scope and Goal of the Thesis ......................................................................................... 18  

2 Methodology ......................................................................................................................... 19  
2.1 The Virtual Hub Network ............................................................................................... 19  
2.2 Selecting a Virtual Hub .................................................................................................... 22  
2.3 Problem Formulation ........................................................................................................ 27  
2.3.1 Assumptions .................................................................................................................. 28  
2.3.2 Notation ........................................................................................................................ 29  
2.3.3 Input Data ....................................................................................................................... 30  
2.3.4 The Virtual Hub Model ................................................................................................ 33  
2.3.5 Passenger Re-accommodation Module ....................................................................... 35  
2.4 Summary ........................................................................................................................... 39  

3 Application of the Virtual Hub Network to a Major US Carrier ........................................... 41  
3.1 Understanding the Airline ............................................................................................... 41  
3.2 Selecting a Day of Operations ........................................................................................ 41  
3.3 Input Data ........................................................................................................................ 44  
3.4 Implementation and Results ............................................................................................. 49  
3.5 Discussion .......................................................................................................................... 52  
3.6 Limitations ........................................................................................................................ 54  
3.7 Summary ............................................................................................................................ 56  

4 Conclusion ............................................................................................................................. 57  
4.1 Concluding Remarks ......................................................................................................... 57  
4.2 Areas for Future Research .............................................................................................. 58  

Bibliography ............................................................................................................................ 61
List of Figures

Figure 2.1: A Virtual Hub Network ................................................................. 20
Figure 2.2: The Virtual Hub Decision Process Over Time ...............................21
Figure 2.3: The Passenger Re-accommodation Module ....................................22
Figure 2.4: The Virtual Hub Selection Process ............................................... 23
Figure 2.5: Delay statistics for the hub airports and virtual hub candidates at American and United Airlines .......................................................... 24
Figure 2.6: Average Daily Delay at RDU for July 2000 .....................................25
Figure 2.7: Actual Aircraft Operations at RDU on July 23, 2000 .......................26
Figure 2.8: Number of Aircraft on the Ground at RDU on July 23, 2000 ..........26
Figure 2.9: Excess Capacity, in Number of Aircraft, at RDU on July 23, 2000 ...27
Figure 2.10: Best Virtual Hub Candidates for American and United Airlines ....27
Figure 2.11: Passenger Re-accommodation within the PRM ............................38
Figure 3.1: Average Daily Departure Delay at the Original Hub and Virtual Hub Airports for March 2002 ................................................................. 42
Figure 3.2: Distribution of Flight Delays at the Hub Airports on March 9, 2002 ....43
Figure 3.3: Passenger Itineraries During the Period of Inclement Weather........44
Figure 3.4: Number of Flights for the Airline at the Original Hub Airport ..........46
Figure 3.5: The Number of Aircraft on the Ground Throughout the Day at the Virtual Hub ...... 48
Figure 3.6: Excess Capacity Throughout the Day at the Virtual Hub...............48
Figure 3.7: Excess Capacity at the Virtual Hub During the Period of Irregular Operations ...... 49
Figure 3.8: Scheduled Flights Through the Original Hub and Hub Airport Capacities ....... 49
Figure 3.9: Problem Size and the Optimal Objective Function Value ..................50
Figure 3.10: Allocation of Scheduled Flights to the Original and Virtual Hubs ........50
Figure 3.11: Passenger Re-Accommodation .................................................51
Figure 3.12: Actual Recovery versus Virtual Hub Network ..............................52
1 Introduction

1.1 Motivation

Since the 1950's, the main focus of airline cost reduction has been the optimization of daily operations to maximize their resource utilization. For many years, airlines manually solved the sequential decision making processes of schedule design, fleet assignment, maintenance routing, and crew scheduling. As airlines expanded to the mega-carriers of today and the field of applied optimization progressed with the increase in computer processing capabilities, each of the aforementioned 'steps' were mathematically formulated and implemented at many airlines. More efficient crew pairings, higher aircraft utilization, and improved overall passenger load factors were achieved through innovative optimization techniques and algorithms. While researchers continue to search for a solution integrating all of the aspects of the airline scheduling problem, the state-of-the-practice models have already minimized operating costs and maximized revenues at levels greatly surpassing the previous manual solutions.

The current optimized schedules leave little slack to accommodate the irregularities common to a large, complex system such as an airline. On a daily basis, airlines are confronted with bad weather, maintenance problems, and a variety of other factors that cause the original schedule to breakdown. Before September 11, 2001, over 25% of aircraft operations (arrivals and departures) were delayed, leading to overtime for crew and ground staff, missed passenger connections, and large passenger delays [Bureau of Transportation Statistics, 2001]. During these periods of irregular operations, schedules that were originally optimized are obsolete and the airlines are forced to resort to a combination of manual and first-generation computerized decision support to delay and cancel flights, re-accommodate passengers, and rebuild complex crew and aircraft schedules.

With irregular operations costing a single, major US carrier up to $440 million a year in lost revenue, crew overtime pay, and passenger re-accommodation costs (according to a January 21, 1997 article in the New York Times), researchers and industry are aware of the large cost
savings associated with uncovering an optimal recovery technique. Although emphasis on enhancing schedule recovery has grown over the past decade, researchers and industry have failed to form a consensus on a general approach or determine a dominant method for optimal results. Successes have been noted in individual areas of the problem, such as crew recovery or aircraft routing, however, solutions addressing all aspects, including cancellations, delays, passenger re-accommodation, and operational and crew scheduling, are far from common. While these complicated systems are available from airline solutions software providers, they often require large amounts of processing power and are met with opposition from experienced controllers who are not comfortable with a ‘black box’ approach to re-adjusting the schedule. These models are also highly dependent on the elusive passenger delay costs (the cost to the airline of delaying a passenger) in addition to other cost coefficients that are difficult to define in practice. Although significant contributions have been made in the area of optimal schedule recovery methods, the enormous potential cost savings dictates continued efforts in designing faster, more effective methods.

1.2 Problem Statement

One of the most devastating events to an airline schedule occurs when inclement weather, the number one cause of delay, affects a hub airport. Capacity is reduced, connections are missed, passengers are delayed on a larger scale than during irregular operations at a spoke airport, and the repercussions resonate throughout the network. Because five of the ten airports having the highest number of delays are hub airports for major U.S. carriers, there is an evident need to provide immediate recovery solutions for these airports [Federal Aviation Administration, 2001]. Although the complex recovery models currently available can be used for a variety of situations, providing a solution for the single devastating scenario of reduced capacity at a hub airport can yield substantial annual cost savings. By reducing the scope of the model to this unique yet prevalent case, the problem becomes more tractable and still has the potential to significantly impact the airlines’ bottom line.

The fundamental hypothesis of this thesis is that during these periods of bad weather at the hub airports, airlines can reduce delays and cancellations by rerouting entire connecting banks of
traffic to another airport with excess capacity. This predetermined alternative airport, or virtual hub, will then host connection complexes to maximize passenger flow through the network during irregular operations at the original hub. Shifting the connecting demand over two hubs can decrease the strain on the original hub and capitalize on under-utilized airports. In addition, the continuity of passenger flow from origins to destinations through one of two hub airports ensures a reduction in total passenger delay. This thesis explores the potential benefits of redirecting flights through a virtual hub and presents a framework for airlines to implement this recovery procedure.

1.3 Previous Work in Schedule Recovery

Schedule recovery has been a fertile research area over the past ten years, transitioning from simple frameworks to intricate optimization techniques. As airlines began to optimize their crew, maintenance, and flight schedules, the increased negative impact of weather, surprise maintenance issues, and other unexpected delays motivated researchers to develop algorithms for optimal schedule recovery. Deriving methodologies from optimal scheduling solutions, researchers in both academia and industry sought a quick and inexpensive recovery plan to bring schedules back to their optimized operations.

In 1993, Jarrah et al. published one of the initial papers on airline decision support. Two separate minimum cost network flow models are introduced for flight delays and cancellations and are solved using Busacker-Gowen’s dual algorithm, where the shortest path is solved repeatedly. The models require a disutility be assigned to each flight with correctness relative to all of the other flights. Factors such as passenger ill-will and delay costs per minute are required to calculate the disutility function, representing the value lost if the flight is cancelled or delayed. Therefore, the results of the model are highly variable and dependent on the inexact calculation of these disutility functions. While the models consider multiple delays and cancellations, aircraft swapping, and spare aircraft, they do not consider a combination of delays and cancellations, nor do they address crew and maintenance considerations or aircraft substitution across fleets, leaving room for additional research.
Teodorovic et al. published several papers on airline schedule reliability and recovery. Teodorovic and Guberinic (1984) published one of the first efforts in daily operational airline scheduling. The paper discusses a methodology to design a new airline schedule and aircraft rotation when one or more aircraft experiences a technical failure. The model aims to minimize total passenger delay throughout the network and is solved using the branch-and-bound method. Only a small network example was considered and the model assumed uniform capacity among fleet types. In addition, the methodology did not address crew requirements, maintenance requirements, or airport operating hours.

Teodorovic and Guberinic (1990) introduced a lexicographic optimization problem considering aircraft scheduling and routing to minimize the total number of canceled flights. The solution method is based on dynamic programming, assigning flights to aircraft in sequences. When multiple solutions are found, the schedule that minimizes total passenger delay on non-cancelled flights is flown. The model does not consider crew planning requirements and therefore, often generates infeasible solutions. Teodorovic and Stojkovic (1995) built on the previous research while considering all operational requirements (airport operating hours, legal and company rules regarding crew working hours, and maintenance requirements). Crew rotation is decided first using a first-in, first-out policy and a sequential approach based on dynamic programming. The aircraft rotation is decided afterwards to reduce computational time. These algorithms require an active role by the dispatcher and rely heavily on their intuition and experience to select the final solution.

Mathaisel (1996) presents a systematic approach to integrating computer science and operations research for schedule recovery problems. Despite the value of previous airline schedule recovery algorithms, the individual solutions are cumbersome, not integrated with each other, and cannot account for all of the underlying issues complicating operations control simultaneously, such as aircraft routings, weather, crews, maintenance, gates, and marketing needs of the customer. A systematic interaction environment is proposed, Airline Scheduling Control (ASC), to facilitate communication between humans, standardized databases across the airline, powerful workstations for decision support equipped with the a suite of optimization tools, and a standardized graphical user interface for schedule editing. By providing a common interface to
the various planning systems (scheduling, crew scheduling, maintenance routing, airport management, marketing, etc), the approach is designed to improve the efficiency of operations through a seamless method of communication to all involved in the decision making process. The system includes a variety of real-time, graphical user displays of schedule information to accommodate each of the groups supporting operations control (crew management, aircraft maintenance, airport operations, and system operations control), in addition to “what-if” scenario capabilities and a rule system to check for the violations of operational constraints. The integrated environment is tested with a small schedule and application of a network flow algorithm to the disruption problem.

Yan and Yang (1996) were the first to introduce a single model that incorporates delays, cancellations, and ferry flights. The model is constructed as a network flow problem that minimizes the schedule-perturbed period after an incident and obtains the most profitable schedule for that period. The network simplex method and Lagrangian relaxation with subgradient methods (for the NP-hard network flow problem with side constraints) are utilized to solve the problem. With a basic dynamic (time-space) representation of the network, a computational example from China Airlines is presented. Only a small, single fleet is considered, indicating that more research is needed before the model can be applied to larger fleets or multiple fleet types. Yan and Lin (1997) build on this initial research by considering temporary station closures and including modification of multi-stop flights and aircraft swapping. Yan and Tu (1997) consider multiple fleets, but none of the models consider aircraft maintenance or crew scheduling.

Clarke (1998) presents an extensive review of the state of the industry in Airline Operations Control Centers (AOCC) and discusses a new decision framework for irregular operations. The mathematical formulation, presented initially in Clarke (1997), is a time-space network flow problem that utilizes an efficient tree-searching algorithm to solve the aircraft routing subproblem. The model simultaneously solves the fleet assignment problem and the aircraft routing problem, implicitly satisfying maintenance requirements through the implemented algorithms. Delays, cancellations, multiple fleet type swapping, air traffic control restrictions, and crew availability are all incorporated into the model. The objective is to minimize the costs associated
with rescheduling. Consequently, solutions are highly dependent on accurate, real-time cost data and predetermined ‘spill’ costs that account for the financial impact of losing passengers on each flight.

Arguello et al. (1997) present a greedy randomized adaptive search procedure (GRASP) to rebuild aircraft routings during irregular operations. The objective is to minimize flight cancellation and delay costs associated with the new routings. The resource assignment is initially formulated as a general integer program and utilizes a randomized neighborhood search technique (GRASP) to generate feasible aircraft routings in polynomial time (excluding maintenance and crew restrictions). Results are highly dependant on the delay and cancellation costs, which are difficult to quantify. A computational example from Continental Airline’s 757 fleet is presented with generalized costs from Jarrah (1993), demonstrating near optimal results in real-time.

Bard et al. (2001) solve the same problem as Arguello (1997) by using a time-band optimization model. By transforming the routing problem into a time-based network with a sectioned time horizon, the resulting formulation is an integral minimum cost flow problem with side constraints. Both the lower bound and solution to the original problem are generated from solving the network flow problem as a linear program, or an integer program if necessary. A simple linear relaxation of this time-band model also provided a lower bound for Arguello (1997). Solution quality is gauged by comparison to the lower bound and the data set from Arguello (1997) is used to provide computational results demonstrating improved solutions over GRASP. The user-specified band lengths directly affect the solution quality and computation time.

Thengvall et al. (2000) present a flexible model allowing decision makers to evaluate the trade-offs between minimizing delays, cancellations, and deviance from the original schedule. The objective function maximizes the modified profits associated with the disrupted schedule, but users are encouraged to evaluate solutions based on the amount of delays, cancellations, and schedule modifications. Modeled as a network flow problem with side constraints, the majority of scenarios can be solved sufficiently using the linear relaxation of the original problem. An
adaptive rounding heuristic to provide near-optimal solutions is presented for use when integrality is not achieved. The model does not address the feasibility of passenger connections or maintenance constraints in the revised schedule.

Thengvall et al. (2001) expand on their initial framework by considering large-scale disruptions resulting from hub closures. Three multi-commodity network-type models are presented: a profit maximization model with incentive to minimize deviation from the original schedule, a generalized profit maximization model with adapted solution algorithms, and the time-band model presented by Bard et al. (2001) modified to include multiple fleets. Results are presented for a Continental Airlines schedule including over 300 aircraft from 12 different fleets for 9 different irregular operations scenarios. The first profit maximization model outperformed the other two models in both solution time and the percentage of cancelled and delayed flights. The authors noted that solutions are highly dependant on the cost parameters defined by the user, especially in longer recovery periods.

Golany et al. (2002) present an interactive goal programming approach to operational recovery decision making. In this example, goal programming sets the original schedule as goals and allows partial solutions by permitting violations from the original constraints. The flexibility of the model enables the decision maker to accept and make small adjustments to a solution found in real-time, having a slight infeasibility. The procedure outlined utilizes the acceptance of a non-global optimum that is considered reasonably good and found within the time constraints. The techniques presented are applicable to a variety of industries and two examples are used to demonstrate the proposed procedure: an abstract application to a minimum spanning tree problem and a practical example of a production-inventory problem. The relevance to airline schedule recovery is discussed although a computational example is not provided.

Rosenberger et al. (2002) explore a stochastic modeling approach to evaluating airline operations. The stochastic model is a discrete event semi-Markov process, described in terms of both random and deterministic states and transitions. Original and current schedule information, recovery policies, and randomly generated ground time, block time, and unscheduled maintenance delays are input into the simulation implementation of the model, SimAir. The
model can employ all of the major recovery components, including delays, cancellations, ferried and swapped aircraft, deadhead crews, reserve crews, and passenger and crew re-routing. These recovery components are combined to form recovery policies, including schedule pushback (delaying a flight until the scheduled plane and crew are ready), passenger pushback (delaying flights so passengers will not miss connecting flights), compensatory crew rest delays, reserve crews for planning violations, and short cycle cancellation. A variety of performance metrics are generated by the simulation, allowing the user to determine the trade-off. SimAir is capable of evaluating a multitude of recovery policies during operations, including individual policies and metrics not included in the publication.

A computational example tests varying deterministic and probabilistic crew scheduling policies with several different recovery policies to accommodate the randomly generated delay events. Results indicate the model provides a more realistic environment to evaluate the performance of an airline plan in operations. The results also suggested that considering delay and disruption probability disruptions in constructing crew schedules might provide better operational performance than the current state-of-the-art deterministic models. Large scale, hub disruptions were not considered in the computational examples.

1.4 Scope and Goal of the Thesis

In this thesis, the methodology and implementation of virtual hubs is explored along with a mathematical model for solving the recovery problem. The effectiveness of the model is evaluated for a major U.S. carrier's airline schedule during a thunderstorm at their hub airport. A comparison with the actual recovered schedule for the airline is also presented to demonstrate the models benefits.

The document is divided into four chapters:

- Chapter 2 presents the methodology and model formulation
- Chapter 3 presents the application of the virtual hub model at a major US carrier
- Chapter 4 summarizes the findings of the thesis and suggests areas for future research
2 Methodology

The methodology behind the virtual hub network for schedule recovery is described below. A description of the network is presented, along with a discussion of the process by which a virtual hub is selected. The chapter concludes with the mathematical formulation of the virtual hub problem and the accompanying Passenger Re-accommodation Module.

2.1 The Virtual Hub Network

A virtual hub is a predetermined alternative airport that hosts part of a connection complex when the scheduled operations at a hub airport are delayed due to weather. Using a virtual hub network maximizes passenger flow through the network by shifting just enough traffic from the original hub to lower the demand-capacity imbalance (thus reducing delay at the original hub) and to direct passengers who are not going to the original hub through an alternative path. For example, consider a thunderstorm at O'Hare International Airport in Chicago. During the inclement weather, actual arrival rates can be reduced by as much as 50% relative to the scheduled arrival rate or the airport can be shut down periodically throughout the day. As an alternative to canceling and delaying flights in response to the reduction in capacity at O'Hare, an airline can reroute a combination of arrivals and departures through an unaffected, underutilized airport. These diverted flights represent a subset of the original connecting bank scheduled through O'Hare, and the optimal combination of arrival and departure cities maximizes the number of passengers able to maintain their connections at either hub airport. By shifting traffic to the virtual hub and decreasing the flights sent through O'Hare, cancellations are minimized and aircraft are correctly positioned for rapid recovery to the original schedule soon after the airport capacity is increased. Figure 2.1 shows the structure of a virtual hub network.
Figure 2.1: A Virtual Hub Network

The virtual hub network would be implemented in the hours before the weather is predicted to impact the operations at the original hub, as outlined in Figure 2.2. Typically, a ground delay program (GDP) is issued by the FAA control tower to provide adjusted aircraft arrival rates when bad weather limits the visibility at the airport. The arrival rates dictated by the GDP fluctuate with the airport conditions and are updated in each time window in the model. The excess capacity for the virtual hub, the reduced arrival rate for the original hub, the scheduled flights through the original hub, and the passenger origins and destinations within the time window are all necessary inputs to the decision making process. From this information, flights are selected for diversion to the virtual hub, service through the original hub, delay until a later time window, or cancellation to maximize the number of passengers accommodated by the network in the present time window. After the initial iteration, the variables are updated and the process is repeated until the schedule is recovered and operations return to normal at the original hub airport.
After the scheduling decisions are made for a time window, some passengers will be disrupted and require re-accommodation. A disrupted passenger is a passenger that cannot fly one or more of the originally scheduled leg(s) of their trip. For a time window within a virtual hub network, a disrupted passenger can be any of the following:

- A connecting passenger with their original flight from their origin serviced by the virtual hub and their original flight to their destination serviced by the original hub.
- A connecting passenger with their original flight from their origin serviced by the original hub and their original flight to their destination serviced by the virtual hub.
- A non-stop passenger with their original flight either to or from the original hub serviced by the virtual hub.

Disrupted passengers are re-accommodated through a heuristic-based Passenger Re-accommodation Module (PRM) that explores the possibilities of accommodating passengers on a
combination of flights sent to or from the virtual hub and flights already scheduled through either of the hub airports in later time windows. Once all of the passengers from one time window have been accommodated, the next time window begins and the decision making process is repeated. Figure 2.3 provides a high-level overview of the PRM:

![Diagram of Virtual Hub Model and Passenger Re-accommodation Module](image)

**Figure 2.3: The Passenger Re-accommodation Module**

### 2.2 Selecting a Virtual Hub

An airline can identify candidate airports for their virtual hub network through the several important characteristics shown in Figure 2.4. First, the candidate airports must be in the same geographic region to ensure relatively similar aircraft utilization and flight times. In addition, the virtual hub candidates must have low average daily delays, indicating they can handle extra traffic. Finally, the excess capacities of the candidates satisfying the two initial criteria are measured to determine if the airports can accommodate the diverted flights from the original hub. Airports with all three of these attributes represent good virtual hub candidates.
To demonstrate the selection process, the virtual hub candidates for the two largest domestic U.S. carriers were examined. The first two aspects of good candidacy were combined to find airports relatively close to the original hub with low average daily delay. The Federal Aviation Administration’s Airport Capacity Benchmark Report of 2001 was used to determine the delay statistics for some of the candidate airports. The report contains delay information for 31 of the largest airports in the country using the number of delays per 1,000 arrival and departure operations from the FAA Operations Network (OPSNET) database. Each of the airports is ranked relative to the other airports represented in the report. Some of the potential candidate airports were not shown in the FAA study, but were still considered based on their geographical location. Candidate airports, delay statistics and the number of gates owned by the airline are shown in Figure 2.5. Three candidate airports were selected for American Airlines and four for United Airlines, each with lower delays than their corresponding original hub and situated in a favorable location to act as a virtual hub airport. Although the delay and location criteria are satisfied, the low number of available gates indicates some of the airports did not have enough excess capacity to become a virtual hub.
Excess capacity was used as the final criteria to determine the best virtual hub airports. High delays at airports can result from reduced capacity during bad weather or being scheduled over capacity during regular operations. Since the low delay criteria for the virtual hub candidates are satisfied, it is assumed these airports are not scheduled over their capacity during regular operations. This assumption suggests the virtual hub candidates are either below or meeting their airspace capacity requirements and therefore, the excess capacity at the virtual hub is a measure of the airline’s ability to accommodate diverted flights. The number of total gates and the number of free gates throughout the course of the day was used to measure the airline’s excess capacity at the virtual hub candidates. By applying the excess capacity criteria to the candidates, the best virtual hub options for an airline are identified.
The calculations for excess capacity at a virtual hub candidate airport are illustrated by the example case at Raleigh-Durham International Airport. First, a representative day with relatively low departure delays (according to the Airline Service Quality Performance (ASQP) database) during the typically high travel month of July was chosen (Figure 2.6). From the figure, Wednesday, July 26, 2000 was selected because of the low level of delay and the resulting representative picture of operations at the virtual hub candidate. A plot of the actual number of aircraft arriving and departing the airport in 30-minute intervals was constructed over the course of the day (Figure 2.7). Starting with the number of aircraft at the airport from the previous day, the number of arrival and departure flights were added and subtracted to keep a running total of the aircraft on the ground throughout the day (Figure 2.8). The number of aircraft on the ground during the 30-minute time intervals is then subtracted from the number of available gates at the airport to determine the excess capacity (Figure 2.9). This process was repeated for all of the candidate airports. The best virtual hub candidates for the two largest domestic carriers are shown in Figure 2.10.

![Figure 2.6: Average Daily Delay at RDU for July 2000](image-url)
Figure 2.7: Actual Aircraft Operations at RDU on July 23, 2000

Figure 2.8: Number of Aircraft on the Ground at RDU on July 23, 2000
2.3 Problem Formulation

The virtual hub problem is formulated as a mixed integer network flow problem. The model is implemented when inclement weather is predicted to affect the original hub and the virtual hub is predicted to have relatively normal operations. Given the original flight schedule, aircraft
capacities, passenger itineraries, and airport capacities during a time window, the model suggests the flights to be diverted, cancelled, delayed, or flown as scheduled in order to maximize the passenger flow through the network. The process is repeated over time windows of arbitrary size until the irregularities at the original hub are resolved and the schedule is recovered.

Because additional passengers may be accommodated on flights previously scheduled to and from the virtual hub airport within the given time window, the model is formulated with three distinct hub airports: the original hub (OH), the virtual hub (VH), and the virtual hub as a normally scheduled airport (VHs). Splitting the virtual hub into two hub airports for modeling purposes ensures previously scheduled flights through the virtual hub remain unchanged while flights scheduled to the original hub can be diverted to the virtual hub. Modeling the fixed virtual hub flights in the formulation also provides more opportunities to divert passengers to pre-existing flights and thus, arrive at their destination in their originally scheduled time window.

After an iteration of the virtual hub model, the disrupted passengers are re-accommodated using the heuristic-based Passenger Re-accommodation Module (PRM). The PRM is a greedy heuristic that searches all options for a passenger’s re-accommodation through the original and virtual hubs over the course of the day. From the complete list of possibilities, the module selects the new itinerary with the earliest scheduled arrival time. Re-accommodated passengers are then added to the time window that corresponds to their new itinerary, and the virtual hub model proceeds to the next iteration. The PRM is an important component in the process of accommodating passengers on a virtual hub network and works in series with the virtual hub model to maximize the number of passengers arriving at their destination on time.

2.3.1 Assumptions

To ensure model tractability and efficiency, the virtual hub formulation was developed with the following assumptions.

**Ground Resource Availability**

It is assumed that the ground resources are in place at the virtual hub to accommodate the diverted flights. Although the excess aircraft capacity is derived solely from the number of
available gates during the period of irregular operations, it is assumed the corresponding ground staff, gate agents, baggage resources, maintenance crew, etc. are also available to accommodate the extra flights at the airport.

**Crew and Maintenance Flexibility**

It is assumed that the flight crew can be diverted to the virtual hub airport and that maintenance procedures can occur close to their originally scheduled time. While the initial virtual hub solution might pose difficulty in meeting the crew and maintenance requirements, it is assumed that crew and maintenance schedules can be easily altered to repair crew pairings and accommodate the maintenance needs of the aircraft.

**Passenger Connections Within A Time Window**

It is assumed that passengers can make their connection at either the original or virtual hub airports if both of their flights are contained within the same time window. Although the 2nd leg departure time of an itinerary could be scheduled at the beginning of the time window and the 1st leg arrival time at to the end of the time window, it is assumed that controllers can shift the schedules to accommodate connections and properly space the flights to match the reduced arrival rates.

**Passenger Consent**

It is assumed that passengers would prefer to be re-routed through the virtual hub than experience extended delays at the original hub. While passengers are scheduled to travel through the original hub, it is assumed that passengers do not have a strong preference towards their connecting airport, especially when compared to their value of time. The model re-accommodates passengers through either hub, trying to provide the earliest arrival time for the passenger.

2.3.2 Notation

**Sets:**

\( O \): the set of all origin airports indexed by \( i \)
$D$: the set of all destination airports indexed by $j$

$H$: the set of all hub airports, including the original hub (OH), the virtual hub (VH), and the scheduled virtual hub (VHs), indexed by $k$

**Decision Variables:**

\[ x_{ik} = \begin{cases} 
1 & \text{if the flight leg from origin } i \in O \text{ is selected to fly to hub } k \in H; \\
0 & \text{otherwise.} 
\end{cases} \]

\[ y_{kj} = \begin{cases} 
1 & \text{if the flight leg from hub } k \in H \text{ is selected to fly to destination } j \in D; \\
0 & \text{otherwise.} 
\end{cases} \]

\[ w_{ijk} = \begin{cases} 
1 & \text{if a path exists from origin } i \in O \text{ to destination } j \in D \text{ through hub } k \in H; \\
0 & \text{otherwise.} 
\end{cases} \]

$z_{jk}$: the fraction of passengers that are accommodated from origin $i \in O$ to destination $j \in D$ through hub $k \in H$.

**Parameters/Data:**

\[ d_{ij} \]: the number of passengers scheduled to travel from origin $i \in O$ to destination $j \in D$.

\[ c_k \]: the aircraft capacity of the airline at hub $k \in H$.

\[ b_k \]: the number of aircraft on the ground from the previous time window at hub $k \in H$.

\[ p_i \]: the capacity of the aircraft scheduled to fly from origin $i \in O$ to hub $k \in H$.

\[ q_j \]: the capacity of the aircraft scheduled to fly from hub $k \in H$ to destination $j \in D$.

\[ f_i \]: the excess capacity on the aircraft scheduled to fly from origin $i \in O$ to the virtual hub $k = VH_s$.

\[ g_j \]: the excess capacity on the aircraft scheduled to fly from the virtual hub $k = VH_s$ to destination $j \in D$.

**2.3.3 Input Data**

The virtual hub model requires five types of input data:

1. Size of the time window
2. Passenger itineraries
3. Original flight schedules
4. Airport capacities
5. Aircraft capacities

Size of the Time Window

The exactness of the flight scheduling and the feasibility of the passenger connections are dictated by the size of the time window. Because specific flight numbers and exact schedule timings are not input into the model, the virtual hub model is formulated such that an origin or destination represents a flight to or from the original hub within a time window. With larger time windows, it is likely that there will be more than one flight to an origin or destination within the time window. The model is not formulated to schedule these flights separately; each origin and destination is considered as one flight through the original hub, regardless of the scheduled number of flights. Therefore, time windows that are smaller than the time between flights in the most frequently served market more accurately depict the number of flights scheduled. Smaller time windows also reduce the variability of the scheduled flight times and decrease the chances for passengers to be disrupted on their re-accommodated schedules. The virtual hub model relies on controllers to shift flights within a time window to accommodate passengers in the instance their departure from the hub is scheduled before their arrival, however smaller time windows also reduce this potential flight overlap and increase the accuracy of the model’s solutions. Although smaller time windows represent better modeling of flights within the model, they also limit the number of passengers that connect within a distinct time window (i.e., arrive on the first leg and depart on the second leg within a time window). The formulation of the model requires passengers to be assigned to a distinct time window and smaller time windows often do not accommodate the average passenger connection time. Both large and small time windows bring advantages and disadvantages to the modeling of the virtual hub network.

The size of the time window represents a trade-off between the number of connections included exclusively in a time-window and the number of flights per time window. Larger time windows consider more passengers and their destinations while smaller time windows provide greater flight scheduling accuracy. The average passenger connection time and the markets served with the highest frequency thus provide the two boundaries of the time window decision. The
decision maker must weigh the aforementioned trade-offs to select the time window size in between the two limits.

**Passenger Itineraries**

In order to re-accommodate passengers on a virtual hub network, it is necessary to have the itineraries for passengers traveling through the original hub during the period of disruption. An itinerary consists of the passenger’s origin, final destination, and flight leg information, where a flight leg is an aircraft flight taking off from an origin and landing at a destination. For each itinerary flight leg originating at or departing from the original hub airport, the flight number and scheduled arrival and departure times are needed to ensure passengers are considered in the objective function during the appropriate time window. For each time window, all of the passengers traveling through the original hub are grouped by their origin-destination pair, regardless of their individual itineraries.

**Original Flight Schedule**

The original flight schedule for all arrivals and departures from the hub airports is required input for the model. For each time window, the origins and destinations of the flights scheduled by the airline through the original hub enumerate the sets of origins $O$ and destinations $D$, respectively. After obtaining the sets of origins and destinations, the flights scheduled to and from these cities via the virtual hub are set to fixed values in the model, with origins in set $O$, destinations in set $D$, and the previously scheduled virtual hub $VH_s$. By comparing the number of flights scheduled to arrive with the number of flights scheduled to depart in the time window at a hub, the number of aircraft on the ground from the previous time window is obtained.

**Airport Capacities**

The capacities for the hub airports are needed to ensure the restrictions at the original hub are satisfied and the virtual hub is not over-burdened during recovery. To calculate the reduced capacity at the original hub for the airline in question, the number of flights for that airline that are scheduled to arrive during the time window are adjusted by the reduction in the arrival rate from the ground delay program (GDP) as follows:

\[
c_k = \frac{\text{scheduled # of arrivals by the airline} \times \text{adjusted airport arrival rate}}{\text{scheduled airport arrival rate}}
\]  

(2.1)
The capacity at the virtual hub is calculated as described in section 2.2.

**Aircraft Capacities**

Because the originally scheduled aircraft is used if the flight leg is flown, the aircraft capacity per the original flight schedule is used in the model for flights traveling through the original or virtual hub. For the scheduled flights traveling through the virtual hub (VHs), the number of passengers booked on the flights is subtracted from the aircraft capacity to obtain the excess capacity on these flights.

2.3.4 The Virtual Hub Model

The virtual hub model can be described as follows:

Maximize: passenger flow

Subject to:

- a path exists from origin to destination through a hub,
- capacity of the hub airports cannot be exceeded,
- aircraft flow balance,
- passengers assigned to an aircraft cannot exceed aircraft capacity, and
- all origins and destinations are flown to or from exactly one hub

Or mathematically as:

\[
\text{Maximize } \sum_{i \in O} \sum_{j \in D} \sum_{k \in H} d_{ij} z_{ijk}
\]

(2.2)

Subject to:

\[
z_{ijk} \leq w_{ijk} \quad \forall i \in O, j \in D, k \in H
\]

(2.3)

\[
w_{ijk} \leq x_{ik} \quad \forall i \in O, j \in D, k \in H
\]

(2.4)

\[
w_{ijk} \leq y_{ik} \quad \forall i \in O, j \in D, k \in H
\]

(2.5)

\[
w_{ijk} \geq x_{ik} + y_{ij} - 1 \quad \forall i \in O, j \in D, k \in H
\]

(2.6)

\[
\sum_{k \in H} z_{ijk} \leq 1 \quad \forall i \in O, j \in D
\]

(2.7)
\[ \sum_{k \in O} x_{ik} \leq c_k \quad \forall k \in H \]  
(2.8)

\[ \sum_{k \in O} x_{ik} - \sum_{j \in D} y_{jk} + b_k = 0 \quad \forall k \]  
(2.9)

\[ \sum_{j \in D} \sum_{k \in \{OH, VH\}} d_{ij} z_{ijk} \leq p_j \quad \forall i \in O \]  
(2.10)

\[ \sum_{i \in O} \sum_{k \in \{OH, VH\}} d_{ij} z_{ijk} \leq q_j \quad \forall j \in D \]  
(2.11)

\[ \sum_{k \in O} d_{ij} z_{ijk} \leq g_j \quad \forall j \in D, k = VH \]  
(2.12)

\[ \sum_{j \in D} d_{ij} z_{ijk} \leq f_i \quad \forall i \in O, k = VH \]  
(2.13)

\[ \sum_{k \in \{OH, VH\}} x_{ik} \leq 1 \quad \forall i \in O \]  
(2.14)

\[ \sum_{k \in \{OH, VH\}} y_{ijk} \leq 1 \quad \forall j \in D \]  
(2.15)

\[ x_{ik}, y_{ijk}, w_{ijk} \in \{0, 1\} \quad \forall i \in O, j \in D, k \in H \]  
(2.16)

\[ z_{ijk} \in \mathbb{R}^+ \]  
(2.17)

The virtual hub model is a network flow mixed integer program with constraints. Constraints 2.3 ensure the percentage of passengers traveling on a path from an origin to a destination through a hub is zero if the path does not exist. Constraints 2.4 ensure that a path cannot exist from an origin to a destination through a hub unless the path exists from the origin to the hub. Constraints 2.5 ensure that a path cannot exist from an origin to a destination through a hub unless the path exists from the hub to the destination. Constraints 2.6 ensure a path exists from an origin to a destination through a hub when both the origin and destination are serviced through the hub. Constraints 2.7 forces the total percentage of passengers served in the time window to be less than or equal to 100%. Constraints 2.8 are count constraints guaranteeing the number of aircraft sent to a hub airport will not exceed the capacity. Constraints 2.9 are conservation of flow constraints ensuring the number of planes sent from a hub does not exceed the number of aircraft arriving or on the ground at the hub airport. Constraints 2.10 and 2.11 are count constraints guaranteeing the number of passengers assigned to a flight leg does not exceed the capacity on the flight leg. Constraints 2.12 and 2.13 are count constraints guaranteeing the
number of passengers assigned to a previously scheduled path through the virtual hub does not exceed the excess capacity on the flight leg. Constraints 2.14 and 2.15 are *cover constraints* ensuring that each flight is sent to either the virtual hub or the original hub or not served in the time window. Constraints 2.16 and 2.17 define the variables as binary or positive real numbers.

### 2.3.5 Passenger Re-accommodation Module

The PRM is the second step in passenger accommodation for each time window. After the virtual hub model assigns flights to the original or virtual hub, the model generates a list of passengers that cannot be accommodated given the adjusted schedule (i.e., disrupted passengers). Once the disrupted passengers are identified, their disrupted itineraries are input into the PRM and all of the possible re-accommodation solutions through scheduled flights at the virtual or original hubs are found. The module then separates the passengers into three categories:

- Passengers re-accommodated later in the day
- Un-accommodated passengers
- Disrupted international passengers

The re-accommodated passengers are added to later time windows for consideration in the virtual hub model while the other disrupted passengers are documented separately.

#### Re-accommodated Passengers

The PRM utilizes eight scenarios to re-accommodate passengers by the end of the day. The first two scenarios accommodate passengers traveling on two-leg itineraries where one leg is rescheduled to or from the virtual hub and the other leg can be accommodated on a previously scheduled flight from or to the virtual hub, in the current time window. Since the virtual hub model treats the previously scheduled and diverted flights traveling through the virtual hub as traveling through two separate hubs, the re-accommodation module repairs the possible connections occurring at the same hub airport:

**Scenario 1:** The itinerary's 1st leg is diverted to the virtual hub and the 2nd leg is flown out of the original hub, however a previously scheduled flight from the virtual hub exists to the final destination in the current time window.
Scenario 2: The itinerary's 1st leg is sent to the original hub and the 2nd leg is diverted to the virtual hub, however a previously scheduled flight to the virtual hub exists in the current time window.

The next two scenarios re-accommodate disrupted passengers with two-leg itineraries where the 1st leg is sent to the original hub but the 2nd leg is rescheduled to depart from the virtual hub:

Scenario 3: The itinerary's 1st leg is sent to the original hub and the 2nd leg is rescheduled from the virtual hub, however a previously scheduled flight from the original hub exists later in the day.

Scenario 4: The itinerary's 1st leg is sent to the original hub and the 2nd leg is rescheduled from the virtual hub, however previously scheduled flights from the origin to the virtual hub and from the virtual hub to the destination exist later in the day.

The next two scenarios re-accommodate disrupted passengers with two-leg itineraries where the 1st leg is diverted to the virtual hub, but the 2nd leg departs from the original hub:

Scenario 5: The itinerary's 1st leg is diverted to the virtual hub and the 2nd leg departs from the original hub, however a previously scheduled flight to the destination from the virtual hub exists.

Scenario 6: The itinerary's 1st leg is diverted to the virtual hub and the 2nd leg departs from the original hub, however previously scheduled flights from the origin to the original hub and from the original hub to the destination exist later in the day.

The final two scenarios re-accommodate disrupted passengers originating from or destined to the original hub with their flight diverted through the virtual hub:

Scenario 7: The leg from the origin to the original hub is diverted to the virtual hub, however a previously scheduled flight to the original hub exists later in the day.

Scenario 8: The leg from the original hub to the destination is rescheduled to depart from the virtual hub, however a previously scheduled flight from the original hub exists later in the day.
The module finds all possible itineraries for each scenario to re-accommodate disrupted passengers over the course of the day. From all of the possibilities, the itinerary with the earliest arrival time is selected for each passenger. Passengers and their new itineraries are then added to input data for the appropriate time windows. If the virtual hub model reschedules the later flights that passengers have been re-accommodated on, passengers are resubmitted to the next iteration of the PRM. Figure 2.11 illustrates the re-accommodation heuristic in PRM.
Scenario 1
1st Leg diverted to VH + 2nd leg on VH

Scenario 2
1st Leg on VH + 2nd leg rescheduled from VH

2-leg itinerary

Scenario 3 & 4
1st Leg diverted to VH
Destined for OH
Accommodated on a later flight from VH
Accommodated on a later flights through OH

Scenario 5 & 6
2nd Leg rescheduled from VH
Originating at OH
Accommodated on a later flights through OH

Scenario 7
Accommodated on a later flight to VH
Accommodated on a later flight from OH

Scenario 8
Accommodated on a later flight to OH

Re-accommodated Passengers

Figure 2.11: Passenger Re-accommodation within the PRM
Un-accommodated Passengers

If the PRM cannot find another itinerary for a disrupted passenger by the end of the current day, the passenger is considered un-accommodated. The un-accommodated passengers from all of the time windows are combined for the total number of passengers potentially disrupted overnight. Although the PRM only considers the original and virtual hub airports for re-accommodation, it is possible for passengers to be re-accommodated through other airports during the day. Therefore, while is may not be possible to serve these un-accommodated passengers through the virtual or original hub, they can still arrive at their final destination via other hub airports or direct flights.

Disrupted International Passengers

Since major domestic carriers often partner with non-domestic carriers for international flights, passengers that are disrupted on international itineraries require more attention during the re-accommodation process. For itineraries scheduled to travel to the original hub and depart on an international 2\textsuperscript{nd} leg, it is assumed passengers must arrive at the original hub within four hours of their scheduled departure time to avoid being disrupted. International flights from code share partners will still depart from the original hub, however they will likely be delayed due to the inclement weather. Past the four-hour time frame, it is highly probable passengers will miss their international connection. Re-accommodating these passengers requires cooperation between both airlines and dictates that these passengers should be considered separately and manually during the re-accommodation process. Passengers are re-accommodated by the PRM within the arbitrary four-hour time frame, but passengers experiencing larger delays, and consequently disrupted, are combined to provide a complete list of disrupted international passengers.

2.4 Summary

The virtual hub network is a framework for schedule recovery during inclement weather at a hub airport. In this approach, airlines divert part of their connecting banks of flights to a virtual hub to minimize delays and maximize the passenger flow through their network. By using three key criteria, airlines can select the best virtual hub candidates to relieve the strain on the original hub.
airports during irregular operations. The virtual hub network flow model selects the optimal combination of flights to divert and assigns passengers accordingly. The Passenger Re-accommodation Module (PRM) addresses the disrupted passengers encountered by diverting the flights and disrupting connections. Working in series, the virtual hub model and the PRM can accommodate passengers with potentially fewer cancellations and smaller delays compared to state-of-the-practice recovery procedures.
3 Application of the Virtual Hub Network to a Major US Carrier

The scenario and data used to apply a virtual hub network to a major US carrier are presented below. The characteristics of the day of operations are first discussed, followed by a detailed look at the airline data. The results and a discussion of their impact and limitations are then presented.

3.1 Understanding the Airline

A major US domestic carrier was used for this case study of the virtual hub model. Typical to other major US airlines, this airline operates a hub and spoke network with three major hub airports located throughout the United States. The airline operates close to 4,000 domestic and international flights a day, not including code share flights operated by international partners. On a typical day, the airline serves over 99,000 passengers, traveling on over 38,000 distinct itineraries. Close to half of these itineraries travel through the original hub airport examined in this study, illustrating their dependence on their hub airport to operate as scheduled. The virtual hub selection process described in Section 2.2 was utilized to identify one virtual hub for the airline within 250 nautical miles of the original hub having relatively low annual delays and excess capacity. The case study explores the implementation and results for this airport in a virtual hub network during periods of irregular operations at the airline's dominant hub airport.

3.2 Selecting a Day of Operations

The virtual hub network is implemented when inclement weather is predicted for the original hub, and the virtual hub is predicted to be unaffected and underutilized. To select a day for the case study, the average flight delay for the airline at both hub airports was examined for March 2002. The average flight delay and the percent difference between the two airports were calculated using the Airline Service Quality Performance (ASQP) Data and the results are shown in Figure 3.1. The percent difference in average departure delay indicates the difference between the amount of delay at the original hub and the virtual hub, i.e., a positive percent difference indicates a larger amount of delay at the original hub and a negative percent difference indicates
a larger amount of delay at the virtual hub. Throughout the month of March 2002, the virtual hub experiences delay larger than the original hub only three times, demonstrating that it is indeed a good candidate for hosting diverted flights. After examining Figure 3.1, it is clear that the day with the greatest delay at the original hub and the largest percent difference was March 9. The significant delays coupled with the high percent difference indicate a good day for further examination, with inclement weather at the original hub and a relatively unaffected virtual hub.

![Figure 3.1: Average Daily Departure Delay at the Original Hub and Virtual Hub Airports for March 2002](image)

A closer look at the operations at both airports on March 9, 2002 supports the assumptions that this is a good opportunity for the airline to implement the virtual hub network. From radar images from the National Climactic Data Center (NCDC) on March 9, 2002, it is evident that a large thunderstorm affected the original hub airport while the virtual hub remained relatively unaffected. The periods of inclement weather began in the morning and lasted until the early evening. Examination of the Collaborative Decision Making (CDM) ground delay programs for
that day confirmed the original hub airport experienced a reduction in capacity from approximately 9 am until 6 pm while the virtual hub did not experience any ground delay programs (all times are given in the time zone of the hub airports). Figure 3.2 illustrates the arrival and departure delays experienced at both airports with flight data provided by the airline. The figure clearly illustrates that while both airports experience delay, the majority of the flight delay at the original hub is over one hour and the original hub experiences significantly more cancellations.

![Figure 3.2: Distribution of Flight Delays at the Hub Airports on March 9, 2002](image-url)
3.3 Input Data

The virtual hub model requires five types of input data:

1. Size of the Time Window
2. Passenger itineraries
3. Original flight schedules
4. Airport capacities
5. Aircraft capacities

Size of the Time Window

Two-hour time windows were selected for the case study. For March 9, 2002, the highest frequency markets are served once an hour while the average connection time for domestic travel is 151 minutes (2.5 hours). The two-hour time window was selected to accommodate both the need for high scheduling accuracy and a large percentage of passengers connecting in distinct time windows. The period of irregular operations was split into five two-hour time windows beginning at 8am and lasting until 6pm.

Passenger Itineraries

Detailed passenger itineraries were obtained from the airline in question for March 9, 2002. Each itinerary consisted of the origin and final destination, date of departure, flight leg information, and the number of passengers. For each flight leg in an itinerary, the departure and arrival airports and scheduled flight times were provided. Only the itineraries traveling through the original hub during the period of inclement weather were necessary for the model. Figure 3.3 provides itinerary information.

<table>
<thead>
<tr>
<th>Itineraries</th>
<th>Passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traveling through the original hub during the period of irregular operations</td>
<td>4,342</td>
</tr>
</tbody>
</table>

**Figure 3.3: Passenger Itineraries During the Period of Inclement Weather**

The itineraries input into the model must only contain the flight legs traveling through the original hub. A subset of the passenger itineraries scheduled through the original hub had more
than one connection (i.e., more than two flight legs). Since the virtual hub model is only concerned with accommodating the passengers through the original hub, the flight legs not scheduled through the original hub were assumed to be on time and are not input into the model. For itineraries with their initial flight leg traveling through other airports, it was assumed their initial flight arrives on time and they can make their connection for their departing flight to the original hub. In cases where the itineraries travel through the original hub in their initial flight leg, it was assumed they were accommodated to their final destination by the end of the day. After these assumptions are applied to itineraries with multiple connections, all itineraries for input into the virtual hub model contained no more than two flight legs.

Itineraries with international flight legs represent a special consideration in the input data for the virtual hub model. International flights typically require special ground resources (customs, duty-free personnel, etc.) and the majority are flown by the airline’s code share partners. A combination of these factors dictate that international flights remain scheduled through the original hub airport and therefore, the itineraries containing these flights were input into the model separately. Because international flights are often given higher arrival preference at airports during inclement weather, it was assumed that passengers traveling on an itinerary where the first leg is an international flight arrive at the original hub airport in time for their second flight leg. These passengers continuing on a domestic flight out of the hub airport were then added to the model as passengers originating the original hub airport. Next, itineraries with an international second leg and domestic initial flight leg were considered by the model to be destined for the original hub. Although international flights are not re-scheduled by the virtual hub model, adding passengers with international itineraries to the origin or destination demand of the original hub airport accurately represented the original domestic portions of the passengers’ itineraries. Finally, one-leg itineraries destined for the original hub on an international flight were assumed to be on time and not included in the model. All of the international itineraries were adjusted to represent the domestic portion of their trip and combined with the domestic itineraries for input into the model.

After the itineraries were adjusted to include only the domestic legs scheduled through the original hub, they were sorted into the appropriate two-hour time windows. All of the itineraries
scheduled to complete both flight legs within the ten-hour period of irregular operations were first identified. Next, two-leg itineraries were placed in a time window if they were scheduled to arrive on their first leg and depart on their second leg within the time window. Passengers traveling on one-leg itineraries were placed in the time windows according to their arrival to or departure from the original hub. Two-leg itineraries that did not fit into the two-hour time windows were split into two separate itineraries consisting of the first flight leg and second flight leg. Splitting the itineraries satisfied the model constraint that an itinerary must be accommodated in one distinct time. Finally, the adjusted itineraries were then placed into the appropriate time window as if they were one-leg itineraries. Regardless of their initial itinerary, all passengers scheduled to travel between each origin-destination pair were combined for each time window and input into the model.

**Original Flight Schedules**

The original flight schedule was used for arrivals and departures at the original hub to determine the origins and destinations served in each time window. The flights were separated into time windows according to their departure or arrival time at the original hub. The domestic cities served by these flights enumerated the sets of origins and destinations. Only the domestic flights are considered candidates for diversion in the virtual hub model. Figure 3.4 provides an overview of the flight information for the airline during the period of irregular operations.

<table>
<thead>
<tr>
<th>Flights between 8am and 6pm the original hub</th>
<th>Domestic</th>
<th>International</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>548</td>
<td>46</td>
</tr>
</tbody>
</table>

*Figure 3.4: Number of Flights for the Airline at the Original Hub Airport*

**Airport Capacities**

The capacity inputs for the original hub airport throughout the period of irregular operations are obtained from the CDM ground delay program reduced arrival rates. For this case study, the hourly arrival rate at the original hub was reduced to two-thirds of the scheduled hourly arrival rate. Given the number of flights originally scheduled for each hour during the period of irregular operations, the original hub capacity was calculated using equation 1.1 (Section 2.3).
For each two-hour time window, the capacities of the corresponding one-hour time windows were combined to obtain the original hub capacity, in aircraft arrivals, for the five time windows.

The virtual hub capacity was calculated using the procedure described in Section 2.1. For March 9, 2002 at the virtual hub, a plot of the actual number of aircraft on the ground in 30-minute intervals was constructed over the course of the day (Figure 3.5). The number of aircraft on the ground during the 30-minute time intervals was then subtracted from the number of available gates at the airport to determine the excess capacity over the course of the day (Figure 3.6). It is assumed flights have a one-hour turn time (i.e., require one-hour at the gate between arrival and departure) and therefore, each hour a gate remains empty represents capacity for one extra flight. With the number of gates available to the airline at the virtual hub equal to 45, the airline can accommodate 45 flights per hour or 90 flights per two-hours. The number of aircraft on the ground for each one-hour period is subtracted from the number of gates and the corresponding one-hour excess capacities are combined for the virtual hub excess capacity (Figure 3.7). For the virtual hub capacity inputs, the minimum number of gates required to reduce the original hub to the reduced arrival rate without canceling flights is used. While the excess capacity calculations show the capacity at the virtual hub can exceed this minimum value, reducing the capacity for diversion also reduces the necessary ground resources and other potential limitations at the virtual hub. It is also assumed that controllers can shift the flights accordingly within the two-hour time windows to match diverted flights to the excess capacity. Figure 3.8 provides the scheduled flights and capacities (aircraft arrivals per time window) at the hub airports during the period of irregular operations.
Figure 3.5: The Number of Aircraft on the Ground Throughout the Day at the Virtual Hub

Figure 3.6: Excess Capacity Throughout the Day at the Virtual Hub
3.4 Implementation and Results

The virtual hub model was implemented for the case study and converged to optimal solutions with reasonable solution times. The problem was solved with OPLStudio optimization software by ILOG on Unix-based Sun workstations. For each time window, the model converged to an
optimal solution with solution times ranging between five minutes and two hours, depending on
the problem size and sparsity of the data set. In each time window, the virtual hub model sent
the maximum number of flights to the original hub and diverted the remaining flights to the
virtual hub. Figure 3.9 and Figure 3.10 provide an overview of the implementation for each time
window.

<table>
<thead>
<tr>
<th>Time Window</th>
<th>Number of Passengers</th>
<th>Constraints</th>
<th>Variables</th>
<th>Passengers Served (Objective Function)</th>
</tr>
</thead>
<tbody>
<tr>
<td>800 to 1000</td>
<td>4,436</td>
<td>26,304</td>
<td>12,247</td>
<td>4,037</td>
</tr>
<tr>
<td>1001 to 1200</td>
<td>6,191</td>
<td>31,311</td>
<td>14,566</td>
<td>5,747</td>
</tr>
<tr>
<td>1201 to 1400</td>
<td>5,139</td>
<td>26,019</td>
<td>12,112</td>
<td>4,753</td>
</tr>
<tr>
<td>1401 to 1600</td>
<td>6,298</td>
<td>41,100</td>
<td>19,099</td>
<td>5,852</td>
</tr>
<tr>
<td>1601 to 1800</td>
<td>3,122</td>
<td>16,639</td>
<td>7,762</td>
<td>2,978</td>
</tr>
</tbody>
</table>

Figure 3.9: Problem Size and the Optimal Objective Function Value

<table>
<thead>
<tr>
<th>Time Window</th>
<th>Scheduled Arrivals at Original Hub</th>
<th>cOH: Original Hub Capacity</th>
<th>cVH: Virtual Hub Capacity</th>
<th>Flights Sent to the Original Hub</th>
<th>Flights Sent to the Virtual Hub</th>
</tr>
</thead>
<tbody>
<tr>
<td>800 to 1000</td>
<td>35</td>
<td>21</td>
<td>19</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>1001 to 1200</td>
<td>41</td>
<td>28</td>
<td>19</td>
<td>28</td>
<td>13</td>
</tr>
<tr>
<td>1201 to 1400</td>
<td>47</td>
<td>32</td>
<td>19</td>
<td>32</td>
<td>15</td>
</tr>
<tr>
<td>1401 to 1600</td>
<td>59</td>
<td>40</td>
<td>19</td>
<td>40</td>
<td>19</td>
</tr>
<tr>
<td>1601 to 1800</td>
<td>33</td>
<td>22</td>
<td>19</td>
<td>22</td>
<td>11</td>
</tr>
</tbody>
</table>

Figure 3.10: Allocation of Scheduled Flights to the Original and Virtual Hubs

After every iteration of the virtual hub model, the passengers that were not accommodated were
entered into the Passenger Re-Accommodation Module (PRM). The PRM was built in Microsoft
Access using Microsoft Visual Basic and a series of SQL scripts to perform the heuristic search
procedures. Figure 3.11 shows the results from the PRM for each time window.
In order to interpret the results from the virtual hub model, the results were compared to the actual recovery procedures for the airline on March 9, 2002. Actual flight times were substituted into the passenger itineraries to obtain the amount of passenger delay and the number of disrupted passengers. While the actual arrival and departure times for the airline’s flights were provided, the flight times for code share flights were not provided. In order to evaluate the actual airline schedule on the same criteria as the virtual hub model, it was assumed that itineraries with international flights were scheduled with a four-hour layover at the hub airport and the international flight was on time. These are the same assumptions used in the PRM for the virtual hub model, and they provide a relative comparison of the number of disrupted international passengers. Because airlines currently re-accommodate passengers during irregular operations, disrupted passengers from the actual schedule were also re-accommodated using the PRM. Passengers were re-accommodated on a combination of flights through the original hub and scheduled flights through the virtual hub, consistent with the virtual hub model re-accommodation. The results from the actual day of operations and the virtual hub model are presented in Figure 3.12.

### Figure 3.11: Passenger Re-Accommodation

<table>
<thead>
<tr>
<th>Time Window</th>
<th>Passengers Not Accommodated by Virtual Hub Model</th>
<th>Re-Accommodated Passengers</th>
<th>Disrupted International Passengers</th>
<th>Un-Accommodated Passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>800 to 1000</td>
<td>399</td>
<td>340</td>
<td>53</td>
<td>6</td>
</tr>
<tr>
<td>1001 to 1200</td>
<td>444</td>
<td>321</td>
<td>107</td>
<td>16</td>
</tr>
<tr>
<td>1201 to 1400</td>
<td>386</td>
<td>361</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>1401 to 1600</td>
<td>446</td>
<td>356</td>
<td>58</td>
<td>32</td>
</tr>
<tr>
<td>1601 to 1800</td>
<td>144</td>
<td>131</td>
<td>9</td>
<td>4</td>
</tr>
</tbody>
</table>
3.5 Discussion

In a virtual hub network, the airlines would likely prefer to minimize the deviations from the original schedule to preserve crew pairings, aircraft routings, and maintenance schedules. The itineraries input into the model have significant of origin-destination demand for the hub airport, in addition to the international itineraries and domestic itineraries outside of distinct time windows that are adjusted to arrive or depart from the hub airport. As indicated by the case study, this artificially high origin-destination demand for the original hub provides the desired effect of maximizing the number of flights flown as scheduled with minimal dependence on the virtual hub. With the current data input procedures, airlines can expect the virtual hub model to rely primarily on the original hub, utilizing the virtual hub only to alleviate the strain on the original hub.

By diverting flights to the virtual hub, the airline does not need to cancel any of the originally scheduled flights. On the actual day of operations, a total of 123 flights scheduled to arrive or depart from the original hub were cancelled. In this case study, the excess capacity at the virtual hub combined with the reduced capacity at the original hub to accommodate all of the scheduled flights. The capacity at the original hub airport fluctuated with the ground delay program, however the virtual hub had enough excess capacity throughout the day to host enough diverted flights for the entire schedule to be flown. The ability to fly all of the scheduled flights

<table>
<thead>
<tr>
<th></th>
<th>Virtual Hub Network</th>
<th>Actual Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Passengers</td>
<td>19,291</td>
<td>19,291</td>
</tr>
<tr>
<td>Passengers Requiring Re-Accommodation</td>
<td>1,665</td>
<td>774</td>
</tr>
<tr>
<td>Disrupted International Passengers</td>
<td>248</td>
<td>237</td>
</tr>
<tr>
<td>Un-accommodated Domestic Passengers</td>
<td>67</td>
<td>207</td>
</tr>
<tr>
<td>Passengers Delayed Over Two Hours</td>
<td>838</td>
<td>14,123</td>
</tr>
</tbody>
</table>

Figure 3.12: Actual Recovery versus Virtual Hub Network
intuitively provides the virtual hub model a clear advantage in reducing flight delays and the impact of inclement weather on the passengers.

Splitting a connection bank between two hub airports results in more disrupted passengers compared to what occurred on the actual day of operations. While over 700 passengers were disrupted by the 123 flights cancelled by the airline’s actual recovery plan, over twice as many passengers were disrupted by the virtual hub model. The larger amount of re-accommodation required in the virtual hub model does require flexibility from the passengers but is not necessarily indicative of their overall satisfaction. The amount of delay experienced by all of the passengers, including the disrupted passengers, provides a more rigorous comparison of the virtual hub model and the actual day of operations.

The virtual hub model accommodated more passengers within two hours of their scheduled arrival time than the airline’s actual recovered schedule. During the actual day of operations, the large number of flight delays and cancellations resulted in a large number of delayed and un-accommodated passengers. In contrast, all of the flights in the virtual hub model flew within their scheduled time window and therefore, only experienced a range of delay between zero and two hours (the size of the time window). The model assumes flight controllers can make minor schedule changes to adequately space all of the flights according to the GDP requirements, which can result in small delays. Although this assumption limits the precision of the recorded delay of less than the size of the time window (two hours), the delay greater than or equal to two hours is a more accurate indicator of the model’s performance. After comparing the delay for passengers in the virtual hub model and the actual recovery procedure, the virtual hub model reduced passenger delays of greater than two hours by 94%. The dramatic reduction in large passenger delays is a result of the virtual hub model’s ability to serve more aircraft and more passengers during periods of irregular operations.

The results of the virtual hub model also show an overall decrease in the number of un-accommodated passengers, despite the increased number of disrupted passengers. Canceling over 15% of the scheduled flights through the original hub airport reduced passenger options of getting to their final destination after their initial itinerary was disrupted. The combination of
maximizing the possible passenger flow through the network and having a zero cancellation rate allowed the virtual hub model to accommodate more passengers by the end of the day than current recovery procedures. Although the virtual hub model disrupts some passengers' connections, the results indicate the impact of these disrupted itineraries is less significant to the number of un-accommodated and delayed passengers than the large number of cancelled flights during current recovery procedures.

3.6 Limitations

Although implementing a virtual hub network in times of irregular operations can bring significant reductions in delays and cancellations at the original hub airport, the model is based on a few key assumptions that can limit its application in the industry.

**Excess Capacity at the Virtual Hub**

The selection process of the virtual hub candidates is deterministic and does not consider the fluctuation of capacity over time. While the number of gates at an airport is fixed in the short term, airlines can shed excess gates over extended periods of time. The current state of the industry demands that airlines trim costs whenever possible and, therefore, the excess capacity found at the virtual hub candidates will likely be reduced over time. Since the aforementioned capacity measurements only depict a snapshot of the airlines' position at the airport, airlines should consider the long-term capacity at their virtual hub candidates before implementing the model.

**Ground Resources**

While the virtual hub selection process considers the excess gate capacity at the airport, the availability of ground resources to service those gates is not considered. Airlines currently place a large emphasis on optimized and efficient work schedules, indicating that the amount of ground staff at the airport will directly reflect the number of scheduled arrivals and departures. Therefore, the proper staff may not be in place to accommodate the passengers despite the available gates at the airport. Other ground resources such as baggage services, maintenance staff, fuel, etc., might also be unavailable at the virtual hub. Despite these current limitations, provisions could be made to have 'on-call' ground staff or ask for over-time from the current
staff to accommodate the diverted flights pending the implementation of a virtual hub network. Airlines can also negotiate with other carriers at the airport for use of their baggage claim areas and other ground resources to handle the extra traffic. Although the ground resources at the virtual hub are currently not addressed and pose potential problems, airlines may be able to devise creative solutions to ensure the proper resources are in place for the successful implementation of a virtual hub network.

Crew Constraints

Another limiting factor in the implementation of a virtual hub network is crew legalities. The model currently does not consider crew pairings and the possible disruptions that would occur by diverting flights to the virtual hub. In addition, current contracts could make it difficult for airlines to shift crew schedules to other cities at the last minute. Although the current virtual hub model does not consider crew constraints, a next generation model can implement crew rules into the optimization process. Also, there is a possibility for airlines implementing a virtual hub network to work with labor unions to gain enough latitude to operate the network during irregular operations. While crew constraints currently pose potential limitations on diverting flights to a virtual hub, the airlines can expand the current formulation and labor contracts to overcome these difficulties.

Solving Over Time Windows

Solving the virtual hub problem iteratively over time windows limits the potential benefits and precision of the solution. For example, the current model does not give consideration to passengers with only one opportunity to make their connection during the day (i.e., their origin and/or their destination is served with a low frequency). By ignoring these passengers' potential for re-accommodation, the virtual hub model has more un-accommodated passengers than necessary. In addition, the current iterative approach considers each origin and destination appearing in a time window as one flight, when in actuality some cities have two flights through the original hub per time window. Treating these flights as one flight eliminates the potential to serve both hub airports within the time window and deliver all of the passengers to their final destination on time. The iterative approach also inaccurately assumes a connection will be possible if flights arrive and depart within the same two-hour time window. It is likely that some passengers will miss their scheduled connections within the two-hour time window, despite the
assumption that controllers can shift flights within the time window to accommodate all of the passengers and the GDP requirements. Although solving the virtual hub model with time windows provides real-time solutions with a relatively simple implementation, the formulations and results are limited in their ability for widespread use.

3.7 Summary

The virtual hub model was successfully applied to a major US airline, yielding large reductions in passenger delay and flight cancellations. A day with inclement weather at the original hub and a relatively normal day at the virtual hub were selected to apply the model. Actual passenger itineraries and schedule information from the airline were input into the model, producing optimal solutions in near real-time. In order to compare the results of the model to the actual recovered schedule, disrupted passengers were re-accommodated using the Passenger Re-Accommodation Module (PRM) to produce a complete picture of passenger delay and disruption for both scenarios. While the model produced encouraging results, the assumptions of the model, along with current industry conditions, present some limitations to immediate implementation at US airlines. Examination of the results and the simple implementation procedure indicate that overcoming current limitations could represent large savings opportunities for airlines and their passengers.
4 Conclusion

4.1 Concluding Remarks

With today’s airlines forced to cut costs and re-examine the efficiency and effectiveness of their operations, virtual hubs may provide a relatively simple solution to re-capturing the large amounts of money spent annually during irregular operations at the airlines. The major US carriers operate hub and spoke networks to exploit the cost savings and high load factors associated with condensing and redistributing passengers through a few select airports. While the hub and spoke system makes sense economically and creates a larger overall network for the airlines, the implications of irregular operations at the hub airport are overwhelming. Inclement weather at the hub airport affects all of the spoke airports, with delays and cancellations resonating network-wide. By creating a schedule recovery solution designed specifically for irregular operations at a hub airport, the problem of re-routing flights and passengers becomes more tractable and focused while retaining the opportunity for large-scale benefits.

The virtual hub network has demonstrated these potential benefits to hub and spoke airlines through a case study at a major US carrier. The examination of the airline’s operations during a thunderstorm at the original hub airport showed that over 93% of the passengers were delayed and over 73% of the passengers were delayed more than two hours. Application of the virtual hub network yielded a 94% reduction in delays over two hours and also reduced the number of passengers disrupted at the end of the day by 70%. While the assumptions regarding the iterative process of the model require closer examination of delays less than two hours, the drastic reduction in extended passenger delays can only have a positive impact on the airline. Reducing delays not only increases passenger satisfaction and confidence in an airline’s ability to arrive on-time, but it also decreases crew over-time pay and other costs associated with canceling and delaying large numbers of flights.

The reduction in delays for the major US carrier in the case study suggests other airlines should explore the implementation of a virtual hub network. The idea has limitations associated with
not accounting for the required amount of ground resources, crew scheduling constraints and varying amount of excess capacity at the virtual hub in modeling the problem, but the dramatic results indicate enough increased profits as incentive for airlines to address these limitations. It is possible that the airlines may not find good virtual hub candidates within their current network, however, it is also possible for the emphasis to shift towards building a virtual hub candidate as opposed to finding an existing one. The benefits may exist for airlines to maintain or strategically increase their resources at a virtual hub to decrease some or all of the strain on the original hub airport, especially for airlines with hubs experiencing large amounts of inclement weather. By overcoming their current limitations and building a virtual hub into their network, airlines and their passengers can receive the benefits of a virtual hub network.

4.2 Areas for Future Research

While the preliminary formulation and implementation for the virtual hub project is presented, further exploration to eliminate some of the underlying assumptions or limitations can increase the practical applications and potential benefits of the model.

Exploring Time Considerations

Expanding the current virtual hub formulation to address time considerations more precisely will provide more exact re-scheduling solutions and delay results. The current time window solution approach only provides a re-scheduling accuracy the size of the time window. Adding exact times to the scheduled flights would allow flight numbers and schedules to be directly matched with itineraries and ensure passengers are able to make their re-scheduled connections. The incorporation of the exact flight times would also assist controllers in making delay suggestions for proper flight spacing to satisfy the reduce capacity requirements. While the current use of time windows provides a tractable and quick solution, the accuracy of the solutions and passenger delay calculations will be increased by more precise time considerations.

Adding times to the variables also adjusts the model to be solved once for the entire period of irregular operations. Since the capacities and schedules of the flights are dynamic, modeling the problem over the entire day would require a non-linear or dynamic programming formulation. A linear formulation cannot correctly capture the necessary changes in the capacity and itinerary
variables required after each diversion decision is made. Solving the non-linear or dynamic optimization of the virtual hub network can provide the exact and optimal solution to the problem, but the tractability, speed, and efficiency of the current formulation will likely be sacrificed. Further research into comparing the linear and exact non-linear and dynamic formulations will provide an upper bound on the potential benefits for the airlines and offer suggestions on the best approach for widespread industry implementation.

**Crew and Maintenance Constraints**

Adding the proper crew and maintenance constraints can also enhance the current virtual hub formulation. With increased emphasis and research placed on schedule recovery from all perspectives of the airlines, the addition of current crew and maintenance recovery formulations will provide a more feasible virtual hub solution. Considering maintenance routings and rigorous crew constraints will restrict the current solution, but also ensure the solution is optimal and can be implemented. Further exploration into the amount of ground resources (including staff) at the virtual hub would also add to the robustness of the model. Adding ground resources, crew, and maintenance considerations eliminates some of the key limitations of the model and increases the potential for successful implementation.

**Cost-Benefit Analysis of Developing a Virtual Hub Network**

Performing a cost-benefit analysis for airlines planning to develop a virtual hub network will provide the necessary tools to make educated decisions regarding the investment. For airlines that do not currently have a good virtual hub candidate, a detailed analysis of the cost savings related to reducing the strain at the original hub and the expenses associated with increasing available ground resources at a virtual hub will give an actual dollar value for decreasing the current limitations of the virtual hub network. The costs of acquiring and maintaining gates, baggage equipment, ‘on-call’ ground staff, and catering resources should be considered in addition to the costs associated with not implementing a virtual hub network (i.e., passenger ill-will and spill, crew and ground staff overtime costs, repositioning costs, etc). Analysis of the financial aspects of virtual hub implementation will provide an important supplement to this initial research.
Bibliography


