Mobile Ad Hoc Networks for Oceanic Aircraft Communications

by

Ryan W. Kingsbury

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science in Aeronautics and Astronautics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2009

Author .....
Department of Aeronautics and Astronautics
August 20, 2009

Certified by .....
Prof. Eytan Modiano
Associate Professor
Thesis Supervisor

Accepted by .....
Prof. David L. Darmofal
Associate Department Head
Chair, Committee on Graduate Students

Mobile Ad Hoc Networks for Oceanic Aircraft Communications

by

Ryan W. Kingsbury

Submitted to the Department of Aeronautics and Astronautics
on August 20, 2009, in partial fulfillment of the
requirements for the degree of
Master of Science in Aeronautics and Astronautics

Abstract

Research into mobile ad hoc networks (MANET's) has exploded in recent years. MANET's are a type of wireless network in which independently mobile nodes are capable of *self-forming* and *maintaining* a connected network, even in the face of topology changes. Surprisingly, there has been minimal work in applying these techniques to a problem that thousands of people face each day: lack of connectivity during transoceanic air travel.

This work investigates the technical feasibility of using MANET techniques to provide connectivity from aircraft to land-based communication infrastructure. In pursuit of this goal, a simulation has been developed which incorporates a wide range of system issues including aircraft mobility, communication link performance, and optimal network allocation.

At the center of this simulation lies an aircraft mobility model which fuses airline schedule data with probabilistic flight delay and cancellation events. An SNR-based link capacity model is used to predict feasible communication rates between aircraft and ground stations. Finally, an optimal max-min fair allocation algorithm is used to assess the capacity of this network.

Our results show that system connectivity percentages in excess of 90% are achievable. Furthermore, the network allocation results indicate that megabit class data rates can be supplied to the majority of users. We conclude this thesis by presenting some design trade-offs that are likely to be of interest to those implementing the system.

Thesis Supervisor: Prof. Eytan Modiano
Title: Associate Professor

Acknowledgments

Whew! What a rush the past two years have been. Times of incredible stress complemented by reprieves filled with a great feeling of accomplishment. It is with great pleasure that I can now acknowledge those who helped me get through this most-humbling of life experiences.

First, I would like to express my profoundest appreciation toward my adviser Prof. Eytan Modiano for his never-ending support, guidance, and encouragement. His willingness to let me explore the research project of my choice was what made this thesis come to life. His razor-sharp recollection of work in the field saved me countless hours of sifting through research papers trying to find "the one" that pertained to the problem at hand. Our weekly one-on-one meetings were exceedingly valuable in dislodging my brain from those ruts that are often encountered in research. Lastly, his never-going-to-give-up work ethic is something that I certainly hope has rubbed off on me.

Next, I would like to acknowledge MIT Lincoln Laboratory and the Lincoln Scholars Program for their generous financial support of my graduate education. I am also indebted to my many colleagues at Lincoln Laboratory who listened to never-ending gripes regarding classes and research with a amazing amount of patience.

Over the past two years, I have learned that going MIT is much more than taking classes and performing research; its also about meeting people. The friendships that I have formed here will no doubt last for a lifetime and I am very thankful to have been able to share these fleeting years with such an incredible group of people. MIT may have provided the education and degree but you all are what made the experience fun!

In closing, I wish to express the sincerest of gratitude toward my parents and family for always being there for me. Their selfless sacrifice towards making sure that I had what I needed to succeed is something that I'll always cherish. Thank you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Research into mobile ad hoc networks (MANET's) has exploded in recent years. MANET's are a type of wireless network in which independently mobile nodes are capable of *self-forming* and *maintaining* a connected network, even in the face of topology changes. The inherent versatility of MANET's makes them well suited for applications where little is known a priori about the topology of the network. Despite all of the research in this field, it is surprising that there has been minimal work in applying these techniques to a problem that thousands of people face each day: lack of connectivity during *transoceanic* air travel.

This work investigates the technical feasibility of using MANET techniques to provide "reach-back" from commercial aircraft to land-based communication infrastructure. In particular, we are interested in exploring the *reliability* and *capacity* merits of this application of MANET's.

1.1 Motivation and Prior Work

We live in an increasingly connected world. Whether it be at work, at home, or during the daily commute; there is almost always a solution for connectivity. Indeed, connectivity in ubiquitous in our everyday lives, however, one common exception to this rule is during commercial air travel.

There is great commercial interest in conquering this final frontier of "last mile"

connectivity. In fact, a number of domestic airlines are starting to offer on-board Internet access. A product called GoGo by Aircell is being rolled out on major carriers across the U.S. which leverages existing cellular telephone technologies [21]. Since link distances are relative small (as compared to a satellite system), Aircell has been able to make their equipment relatively lightweight and inexpensive. Unfortunately, since this system uses air-to-ground connections for all aircraft, it is not a feasible solution for oceanic coverage.

To avoid this problem, some companies have turned to satellite technology. The best known examples of this approach include Connexion by Boeing and, more recently, a system developed by a company called Row 44 [17, 5]. The Boeing product was a commercial failure due to heavy and expensive equipment requirements as well as poor market timing. Though it was deployed on multiple carriers, the program was eventually canceled due to insufficient growth [3].

The ongoing Row 44 project is seen by many as a risky endeavor given the fate of Boeing's Connexion system. While Row 44 claims to have made major improvements on per-aircraft retrofitting costs and equipment weight, they are still bound by the expensive nature of satellite bandwidth. Additionally, there are concerns about the scalability of this system since all aircraft must communicate with a single satellite or constellation of satellites[1].

We are motivated to study the use of MANET's in this application because they have the potential to capture the benefits associated with short-range links while at the same time offering high-capacity and good scalability in the oceanic regime.

In addition to private industry's efforts to tackle this problem, a number of recent academic studies consider this application of MANET's. In [19], Sakhaee describes the basic concept of using MANET's to supplement an existing satellite system. He models aircraft locations as a two-dimensional Poisson process and completes a simple connectivity assessment. Him analysis is limited by the fact that he doesn't incorporate actual airline operating schedules nor does he consider the dynamic behavior of

[1] An additional limitation of many satellite systems is poor availability in the polar region which is traversed by many intercontinental airline routes.

the system.

In [4], Campos considers the business challenges associated with deploying a MANET on the North Atlantic market. Her conclusions are based upon surveys conducted with various industry stake-holders (e.g. airlines, air traffic control groups, etc.). The main questions she tries to answer are who benefits from such a system, how much they must invest in order to get the system operational and whether they will be able to recover their investment.

Her work differs from this thesis in that she is more interested in the airline operational benefits of the idea. Specifically, she suggests that this communication system could be used to improve the efficiency of air traffic control over the North Atlantic. By providing more reliable communications to aircraft operators, it should be possible to relax spacing requirements and fly along more energy-efficient paths. While we approach this problem from the angle of improving passenger connectivity, it could certainly be used for airline operations as Campos suggests, however, a much more rigorous reliability analysis is needed. Overall, Campos' work nicely complements the technical contributions of this thesis by assessing the financial viability of the idea.

1.2 Our Approach

In pursuit of our goal to understand the technical feasibility of using MANET's in this application, we have developed a system model which incorporates ideas from both the aviation and communication disciplines. Our system model fuses aircraft mobility predictions with a SNR-based communications model to produce an evolving picture of potential communication links between aircraft. By analyzing this model, we have gained insight into the connectivity behaviors of the system.

In addition to raw connectivity, we are keenly interested in the capacity or throughput that the system can provide. To satisfy these interests, we use linear programming techniques to solve for the optimal max-min fair capacity of the system. These optimal results provide an important baseline for comparison to (sub-optimal) system designs which can be realized with actual hardware.

Our final contribution is to consider some of the implementation challenges and design trade-offs for this system. For example, there are unique physical limitations associated with mounting antenna on aircraft due to aerodynamic constraints. We consider how these constraints impact antenna system capabilities and hence the overall performance of the system.

1.3 Case Study

In this thesis, we focus on the North Atlantic oceanic airspace. In terms of number of flights per day, this is the most heavily-traveled oceanic airspace in the world. For this reason, it is a leading candidate for the application of the MANET techniques since they require the cooperation of many nodes. It would also be interesting to study other popular markets such as the North Pacific or Arctic regions, however, we leave these to future research.

1.4 Organization

The remainder of this thesis in structured in the following manner. In Chapter 2, we describe our system model, the justifications for our assumptions, and the known limitations of the model. In Chapter 3 we complete a connectivity analysis of this system. We present results relating to both the overall system connectivity as well as connectivity seen from the user's perspective. Chapter 4 is devoted to our analysis of the capacity of the system. Again, we consider both system-wide performance measures such as average node throughput as well as geographically localized metrics such as expected capacity versus aircraft position. Finally, we discuss some system design trade-offs in Chapter 5 that would likely be of interest to engineers implementing this system.

Chapter 2

System Model

In this chapter, we describe our model of the communication system. We start by giving a brief introduction of commercial flight operations over the North Atlantic. This is followed by a description of the mobility model which is used to determine physical node positions. We then present the communications model which uses node position information to determine which point-to-point communication links can be established as well as their capacity.

2.1 Oceanic Flight Operations

In this section, a brief overview of oceanic flight operations is provided. This overview starts with a short discussion of the airline scheduling patterns observed in the North Atlantic market. The air traffic control system is described and the differences from the ATC system used for domestic flights are described. We close by discussing the specific ATC system currently being used in the North Atlantic.

Nearly all aircraft operating in the North Atlantic market exhibit a similar scheduling pattern: they generally fly westbound during daylight hours and then return eastbound overnight. These patterns are a direct result of the airline's desire to highly utilize their aircraft as well as to allow passengers to efficiently connect to other flights on either side of the ocean. These scheduling patterns produce two large "waves" of traffic over the North Atlantic each day. As we will show in this thesis, these peak

travel times are of great benefit to cooperative systems such as our MANET.

Commercial aircraft are operated in a vastly different manner when flying through oceanic airspaces than when flying over populated areas. In the United States, Canada, and Europe a large system of ground-based radars have been deployed in order to monitor the movement of all aircraft. In these areas, a human air traffic controller monitors all aircraft in realtime and works to arbitrate conflicts where necessary. This type of human-in-the-loop control is known as *positive control*.

Unfortunately, it is too costly to extend coverage of these radar systems to very remote regions or over the world's oceans. Due to this lack of radar coverage, a different form of ATC known as *procedural control* is utilized. In procedural control, all aircraft obey predetermined separation and routing rules in order to avoid conflicts.

In the North Atlantic, most aircraft utilize the Organized Track System (OTS) for their procedural control needs. The OTS includes a set of predetermined non-intesecting paths across the ocean. Flights are guided under positive control by ATC from their origin airport to the entry point of the OTS. Once a flight reaches the entry point of the OTS, it is "slotted" into a specific track at which point it is required to travel at a predetermined speed along the track waypoints. At the terminus of the track on the far side of the ocean, positive control is resumed to guide the flight into its destination [24]. The OTS is discussed in more detail in later sections relating to our mobility model.

2.2 Mobility Model

A critical issue in the design and analysis of any MANET is understanding the mobility characteristics of the nodes. For this study, it would be ideal to use actual node position data for this purpose. Unfortunately, the lack of radar coverage over the oceans means that this data is not available for our study. As a compromise, we have chosen to approximate aircraft movement using published flight schedules and great circle extrapolation.

2.2.1 Airline Schedule Data

The primary data source for the mobility model are published airline schedules. This data was kindly provided by ITA Software and represents a filtered version of published flight schedules [25]. The only flights included are those with scheduled passenger service between the US/Canada and Europe.

To capture seasonal variations in flight schedules, we have obtain a separate set of data for the "low season" (February 2008) and the "high season" (July 2008). Flight volume in the North Atlantic market varies by nearly a factor of two between seasons so it is important to consider both situations. The low and high season schedule data used in this thesis contains 517 and 861 unique flights, respectively.

The schedule data was provided in an electronic format and each entry included a flight number, the location and time of departure and arrival, and the operating frequency (days of the week that the flight is operated). The schedule data was filtered to remove code-share flights (one aircraft carrying two flight numbers) so that there weren't erroneous extra nodes in the simulation. We have also opted to remove all flights which are not operated on a daily basis.

2.2.2 Trajectory Extrapolation

We assume that all aircraft travel along the great circle, or shortest path, between their origin and destination. This is clearly a simplification that ignores the impact of ATC, weather, and restricted airspace's. That said, we contend that it is a good first-order approximation since we are considering long-haul routes where there are cost and time incentives to closely follow the shortest route. In section 2.2.4 we discuss this simplification, and its limitations, in more detail.

The airline schedule data provides us with a source and destination airport for each flight. These airport codes were translated into latitude and longitude values by using the databases provided by [22] and [16].

Next, we then turn to a cartographic projection library known as PROJ.4, which has been provided by the US Geological Survey, to compute the great circle paths

between airports [6]. Though it is likely an insignificant detail, the PROJ4 projection library has been configured to use the WGS 84 map coordinate system which matches the system used for the aforementioned airport database. This coordinate system is used consistently throughout the study.

Given the great circle path for each flight route, we sample it at intervals which correspond to the average speed of the aircraft to determine the position of the flight. We ignore prevailing winds and assume that all aircraft cruise at Mach 0.84 (approximately 900 km/hr) which is typical for most modern commercial jet aircraft.

2.2.3 Delay Model

Up to this point, we have described a model which is entirely deterministic. That is, all flights depart exactly at their scheduled time, fly precisely on the great circle route, and always arrive at the same time. Clearly real-world flight operations are not this predictable or consistent.

To enhance our model, we have included a stochastic delay component which is used to perturb each flight's scheduled departure time. The idea here is to randomize our aircraft departure times and hence their relative positions. In doing so, we ensure that we are not misled by schedule data which happens to produce node positions which yield unusually good (or unusually bad) network behavior.

Each flight randomly and *independently* experiences one of the following events: on-time departure ($P = 0.75$), delayed departure ($P = 0.20$), and cancellation ($P = 0.05$). Delayed flights have their departure times shifted by an exponentially distributed random variable with a mean of 30 minutes. Canceled flights are removed from the simulation altogether.

The parameters of this delay model are loosely based on delay statistics provided by the Bureau of Transportation Statistics [15], however, no attempt was made at matching them to the exact flights and routes used in this study. Nevertheless, their purpose is to provide a source of randomness in the model not to precisely characterize the delay behavior of real flights.

2.2.4 Mobility Model Limitations

As with any modeling task, one is constantly faced with the trade-off between improving the model's fidelity and keeping it simple for analysis. The mobility model used in this project is no exception and in this subsection we disclose some of the known limitations of the model.

First and foremost, we are entirely neglecting the potential impact of weather in our model. While the delay model does provide us with some chance of flight cancellations, it does so in an *independent* manner. Circumstances such as a weather systems can cause widespread *correlated* delays throughout the air transport system. For example, a major weather system over the north-eastern US could cause the shutdown of major airports such as BOS, JFK, PHL, etc. This would lead to cancellation of a large number of Trans-Atlantic flights. An area of future work is to investigate how susceptible the system is to these correlated failures.

Second, the the flight trajectories produced by our simulation always follow the great circle path between origin and destination. In reality, most flights use the Organized Track System (OTS) as mentioned in the first part of the chapter. Flights that use the OTS are slotted in to one of a handful of parallel tracks at the start of their oceanic crossing. The flights follow their assigned track until they reach the terminus point where ATC resumes positive control of the aircraft.

The OTS paths, which are adjusted daily, often deviate from great circle positions since they are optimized to allow the airlines to take advantage of (or to avoid) prevailing winds. Despite these deviations, we claim that the Great Circle trajectories are still a reasonable approximation since it would perturb all flight trajectories in a similar manner[1]

Lastly, we have chosen to ignore ground delays. Specifically, we directly use the scheduled departure time (a.k.a. "block time") as our take-off time. In reality, this block time is followed by push-back and taxi before the flight takes to the sky. Taxi

[1]Those familiar with OTS operations will note that east and westbound tracks are generally routed through vastly different latitudes. We claim that this isn't of consequence to our simplification since the eastbound and westbound flights generally operate in different times of the day. In other words, there isn't a strong dependence between these flights in our communications system.

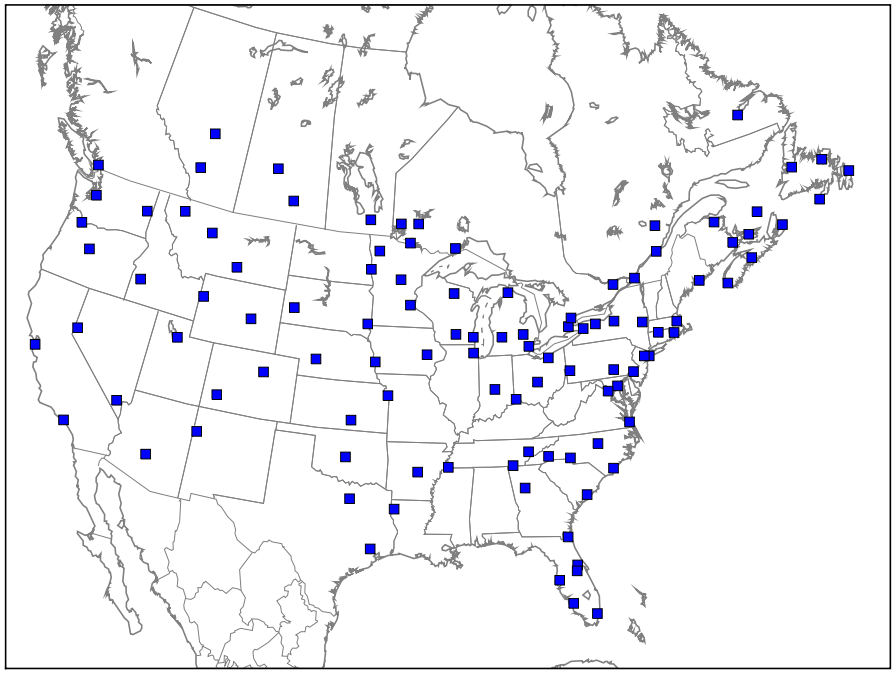


Figure 2-1: Ground Station Locations - North America

times do vary widely from airport to airport and depending on local conditions. While statistics on these issues are available, we have chosen to omit them from this study.

2.3 Ground Segment

In addition to mobile airborne nodes, our system also has a network of ground stations. These nodes serve as the gateways between the airborne network and ground infrastructure such as the Internet. Our system model includes 202 such ground stations.

We place these ground stations at major population centers across North America and Europe as well as at some strategic locations (such as Gander in Eastern Canada). A requirement for these ground stations is that they all must be located in places with well-established terrestrial communication infrastructure so that they can forward traffic from the aircraft to the Internet. The location of our ground stations are shown in Figures 2-1 and 2-2.

We assume that all ground stations are located at sea-level. This provides a con-

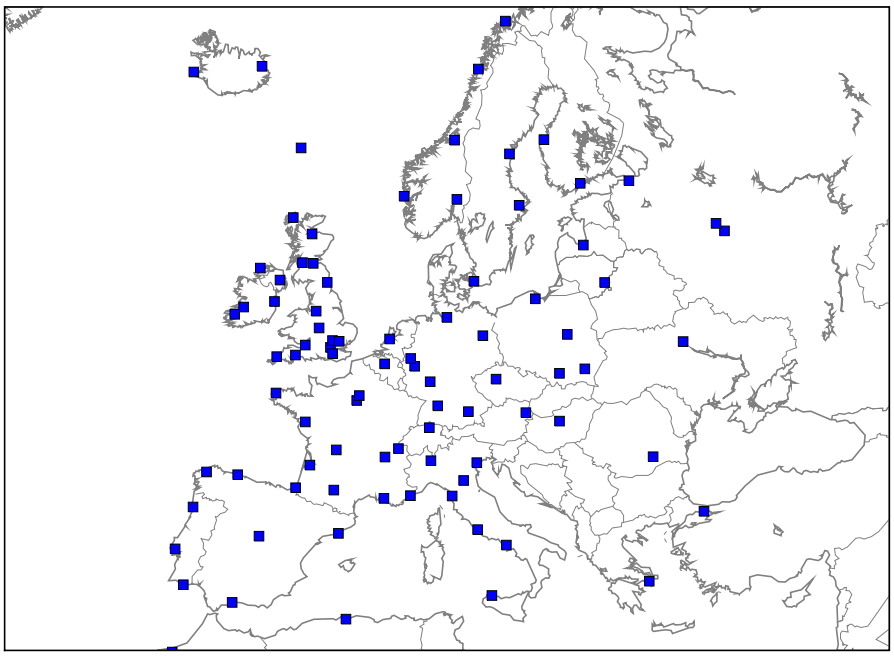


Figure 2-2: Ground Station Locations - Europe

servative estimate of their line-of-sight communications range. We also assume that they have omnidirectional communications capability and can support an unlimited number of simultaneous users.

2.4 Communications Model

In this section we discuss the model which is used to predict the existence and capacity of point-to-point communication links in the network. The input to this model are the aircraft positions, the ground station positions and a set of system parameters. The output of this model is a graph with nodes that represent either aircraft or ground stations and weighted edges which represent communication links with specified capacities.

2.4.1 Frequency Selection

We assume that a VHF, UHF, or microwave frequency is used for all communication links in the network. Lower frequencies (such as HF) tend to have longer albeit less-reliable propagation capabilities, however, the bandwidth available tends to be very limited. Since the goal for this system is to provide high-capacity access (on the order of megabits per second), we focus on higher frequencies where spectrum is more plentiful.

Selecting a precise frequency (or range thereof) for use in this application is a complicated matter since the system must operate in an *international* context. While the International Telecommunications Union (ITU) has centralized many decisions about spectrum allocation, many of the finer aspects of allocations are left up to the constituent ITU regions. Finding internationally available spectrum for this system is a complicated regulatory problem on its own.

Since our main interest is in examining the technical feasibility of this system, we have selected a channel frequency of 25.3 GHz with a bandwidth of 50 MHz. This choice is not entirely arbitrary. First, it lies within a portion of the spectrum that has been assigned to mobile services in ITU Regions 1 and 2 which cover the Americas, Europe and Africa [10, 14]. Second, it avoids any major oxygen or water absorption "lines" which lead to heavy attenuation. Third, the short wavelength allows for physically compact, high-gain antennas. Lastly, a bandwidth of 50MHz was chosen because it is well within the capabilities of modern amplifiers (such as TWTA's) that are used at these frequencies.

2.4.2 Link Geometry

At microwave frequencies, radio signals primarily propagate in a line-of-sight manner. These propagation characteristics have led us to require that a "line-of-sight" exists between two nodes in order for them to communicate. Computing whether or not a line-of-sight exists is a simple exercise in trigonometry given that one knows the position of the two endpoints and the radius of the earth.

We assume that all aircraft maintain a constant altitude of 30,000ft. This serves as a conservative estimate as typical cruise altitudes are above 30,000ft where line-of-sight distances would be even further. Obviously this assumption does not hold up when aircraft are taking off and landing. To account for this situation, we "crop" the first and last 30 minutes of each flight. When a particular flight is cropped its position is still tracked by the mobility model, however, it is assumed to be unavailable for communications. This 30 minute period is sufficient for ascent and descent as well as any airport area maneuvers.

We calculate a ray-traced line-of-sight between the two endpoints assuming that there aren't any obstructions (such as mountains) along the communications path. Given our assumptions on aircraft altitude, this implies an air-to-air communication range of approximately 680km and a air-to-ground range of 340km. The latter is lower because one node (the ground station, located at sea level) doesn't have the altitude advantage that an aircraft would have.

In the case of a 680km link, the signal traveling along the link grazes the earth in a tangential fashion at exactly the link midpoint. A question that arises when this occurs is whether or not diffraction is occurring which could produce signal fading. These fading effects can be estimated by considering the size of the Fresnel zones for the link. Given the link length and signal frequency, a quick calculation shows that the size of the first Fresnel zone is a mere 50m. This is insignificant as compared to our node altitudes of 30,000ft (9,144m) which could be adjusted slightly to compensate [20, p. 399].

2.4.3 Link Budget

We use a simple link budget in order to estimate the received signal power of our communication links. This budget factors in transmit power, antenna gains, and path loss and is given by Equation 2.1. The last term L_{margin} represents the "link margin" needed for the receiver to properly decode the received signal. We assumed

a fixed -6dB margin for all links.

$$(P_{RX})_{dBm} = (P_{TX})_{dBm} + (G_{TX})_{dB} + (L_{path}(d))_{dB} + (G_{RX})_{dB} + (L_{margin})_{dB} \quad (2.1)$$

For the purposes of this simulation the transmit power P_{TX} and antenna gains G_{TX} and G_{RX} will be constant. The path loss $L_{path}(d)$ is a function of the link distance.

Transmit Power

At the microwave frequencies we are considering, traveling wave tube amplifiers (TWTA's) represent the state of the art in efficient power amplification. These amplifiers are commonly used in SATCOM applications and consequently have been adapted for aerospace applications. We assume that our nodes have TWTA's capable of generating a 100 W signal. This is well within the capabilities of modern TWTA's.

$$P_{TX} = 100 \text{ W} = +50 \text{ dBm}$$

Antenna Gains

The antennas used in our simulation are patterned after a simple circular parabolic dish design. For antennas of this type, the maximum gain is given by

$$G \approx \eta \left(\frac{\pi^2 D^2}{\lambda^2} \right)$$

where D is the reflector diameter in the same units as the wavelength λ and η is the illumination efficiency factor (typically $0.5 \leq \eta \leq 0.6$) [18].

We assume that all links use an antenna of this type which is 15cm in diameter and which has an illumination factor of 0.5. At 25.3 GHz, an antenna with these parameters has a maximum gain of approximately 29dB. We assume that both receive and transmit antenna have identical capabilities.

$$G_{TX} = G_{RX} = +29 \text{ dBm}$$

Path Loss

At the frequencies proposed for this system, many effects impact the attenuation will be observed on the communication links. Some of the largest players include water absorption (both vapor and precipitous forms like rain or fog), oxygen absorption, and multi-path fading. Indeed, hundreds of books, thesis, and papers have been written on the topic and many a career have been devoted to understanding these physical effects.

We considered a number of complex models for path loss at these frequencies including [12] and [13]. Unfortunately, these models are not well-suited for this application. Their main shortcoming is that they assume a homogeneous medium between the transmitter and the receiver. In our system, the parameters of the medium such as temperature, pressure, and humidity vary widely along the link. This is because the link path between nodes traverses a wide range of altitudes in the atmosphere.

Since our main interest in this study is to understand the performance of this system as a whole, we have opted to use the simpler free-space path loss equation. The free-space path loss equation is given by

$$L_{path}(d) = \left(\frac{\lambda}{4\pi d}\right)^{\alpha}$$

where α is the path loss exponent, d is the distance and $\lambda = c/f$ is the wavelength [18]. We assume a path loss exponent of 2.5 for our simulation which seems to produce a conservative estimate of path loss as compared to some of the predictions listed in [12, 13].

Receiver Noise

The receiver in our model accepts both the signal of interest plus some fixed white Gaussian thermal noise. The noise power is computed using Boltzmann's constant ($k_B = 1.38 \times 10^{-23}$ W-s/K), the noise temperature ($T_0 = 300$ K), and the channel

bandwidth (B in Hz):

$$\begin{aligned} N &= B \cdot k_B \cdot T_0 \\ &= B \cdot 4.142 \times 10^{-21} \text{ W/Hz} \end{aligned}$$

In our simulation, we ignore any interference contributed from other sources (such as other nodes).

2.4.4 Channel Capacity

Using the link budget described above, we are able to calculate a signal-to-noise ratio (SNR) for each link. Given this SNR, we use the Shannon-Hartley theorem to calculate the capacity of the link. The Shannon-Hartley theorem [18] states that

$$C = B \times \log_2 \left(1 + \frac{S}{N}\right)$$

where C is the channel capacity in bits per second, B is the channel bandwidth in hertz, and S and N are signal and noise powers respectively given in watts.

2.4.5 Communication Model Limitations

The communications model presented above is admittedly simple, however, we have attempted compensate for its limitations by using conservative estimates where it was deemed important. Despite these conservative estimates, there are still some limitations which need to be disclosed.

First, our communications model uses a relatively simple link budget. An actual communication system involves many more terms that crop up due to the implementation losses in the physical design. These losses come from a variety of sources such as antenna imperfections, digital processing side-effects, and connector losses. Since we can't possibly account for all of these terms, we have lumped them into a single link margin term.

Second, we use the Shannon channel capacity in our predictions. Modern bandwidth-

efficient modulation techniques can come close to Shannon's bound, however, it is somewhat challenging to make a system that is capable of operating at a variety of points along the Shannon bound. In order to achieve these capabilities, we assume that the real system uses some form of adaptive modulation (both in symbol rates and possibly in modulation type).

Third, we completely ignore receiver design issues. In this system, there are many cases where a given aircraft must simultaneously decode multiple signals. These sort of issues can be solved using spectrum sharing techniques (such as CDMA), however, they come with their own set of losses which must be taken into account. An additional complication is created by the fact that a single receiver may have to decode signals from two nodes one of which is "near" and the other is "far." These situations require a large amount of receiver dynamic range which, again, must be accounted for in the receiver engineering.

Lastly, our communications model does not include any constraints relating to antenna steering. We feel that these issues are of great importance to this application, however, we chose to omit them from the basic model in order to establish a baseline on system performance. In Chapter 5, we refine our assumptions about antenna steering to better understand their effect on system performance.

Chapter 3

Connectivity Analysis

In this chapter, we examine the connectivity of the system. We define "connectivity" in the graph theoretic sense; we check for the existence of a path from a given aircraft to any ground station. This path can be either direct or via relaying aircraft. Throughput and capacity issues are ignored in this chapter but are covered in Chapter 4.

This chapter has been organized into various experiments each of which gives insight into a different aspect of system behavior. We start with the most obvious system metric: percentage of aircraft connected. The second experiment explores system connectivity from the user's perspective or, more precisely, how the probability of connectivity tends to vary during the progression of a flight. We conclude the chapter by investigating the dependence of flight routing on connectivity.

3.1 Overall System Connectivity

The goal of this experiment is to determine what fraction of the aircraft have ground connectivity either directly or indrectly over the course of a 24-hour period. A secondary goal of this experiement is to demonstrate the system's resilience to system schedule "jitter."

In this experiment we consider both the summer and winter datasets since they each have vastly different numbers of scheduled flights. Using each of these data sets,

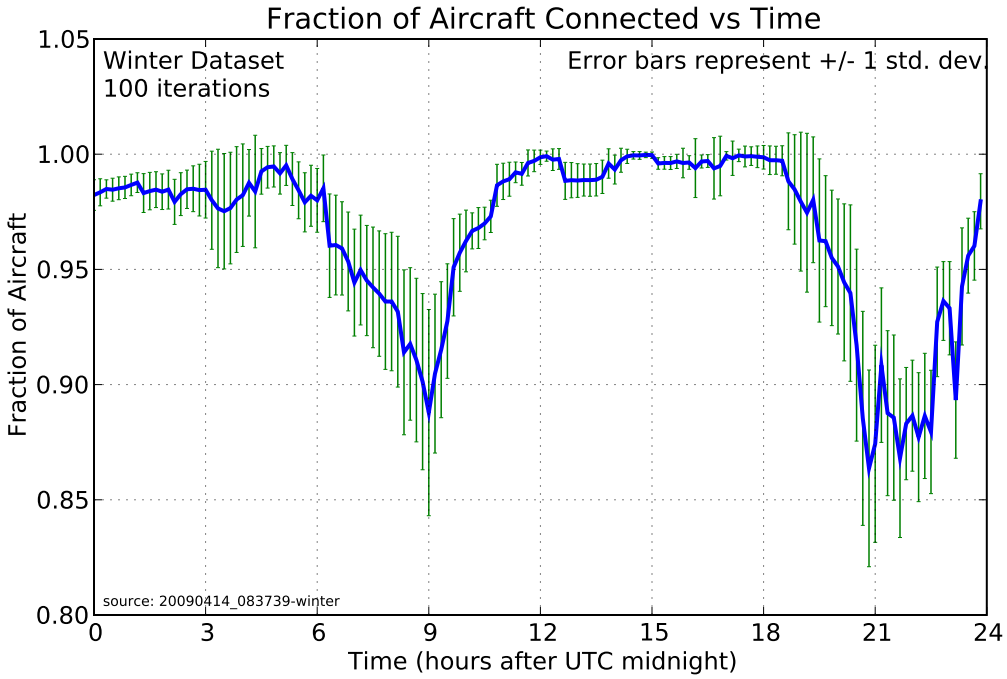


Figure 3-1: Plot of percent connectivity vs time - Winter Dataset

a total of 100 regressions with random schedule delays were executed. A "regression" is a single run of our 24-hour simulation with some set of delay values for the flights. The simulation time step was set at 10 minutes to strike a good balance between time resolution and reasonable simulation runtimes.

3.1.1 Results

Our simulation produces a connectivity graph at each time step. For each of these connectivity graphs, we use breadth-first search algorithm (detailed in [11]) to determine which nodes have ground connectivity. This quantity is divided by the total number of aircraft currently airbourne to determine a connectivity percentage. This process is repeated at each time step and in each of the delay regressions.

Figures 3-1 and 3-2 show the mean values for this percentage as averaged over all the regressions in the experiment. The error bars represent show plus/minus one standard deviation.

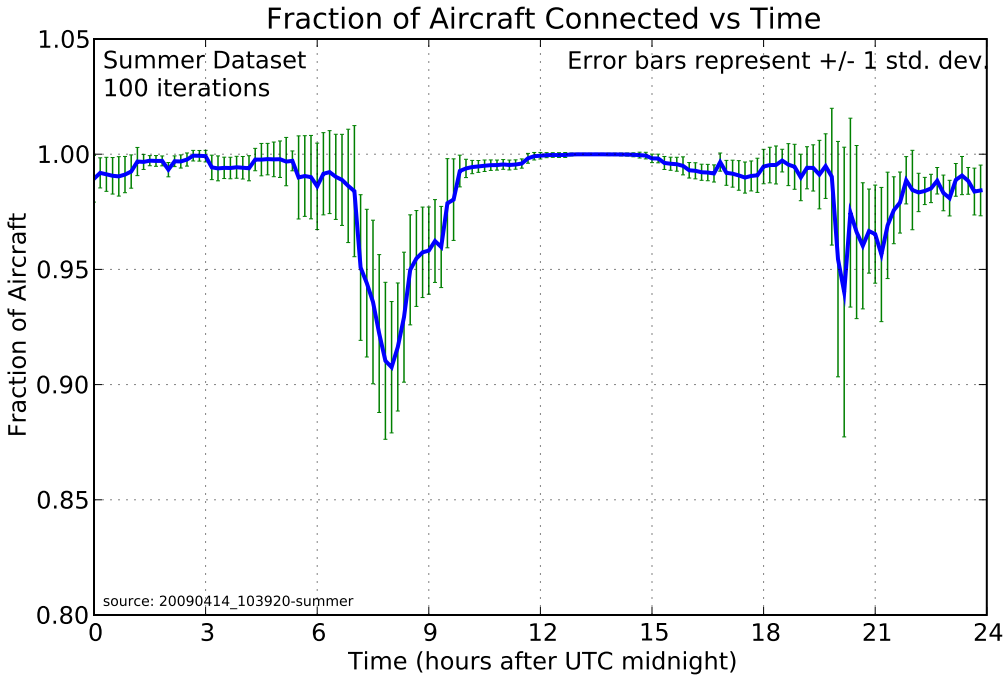


Figure 3-2: Plot of percent connectivity vs time - Summer Dataset

3.1.2 Discussion

In general, both plots show connecvity percentages exceeding 90% for the majority of the 24-hour simulation. As expected, the summer dataset yields better performance due to the larged number of scheduled flights.

A second issue of interest are the connectivity "drop-out" events observed in both datasets around 08:00UTC and 21:00UTC. These drop-outs are somewhat misleading in that they correspond to a time of day where there are very few aircraft in the air.

Lastly, the error-bars show that system connectivity is very consistent during peak travel times (e.g. 03:00UTC and 15:00UTC) despite the fluctuating aircraft departure times. At the same time, connectivity is quite erratic during the drop-outs mentioned earlier. This follows intuition in the sense that fewer aircraft lead to a more fragile MANET than one composed of many nodes.

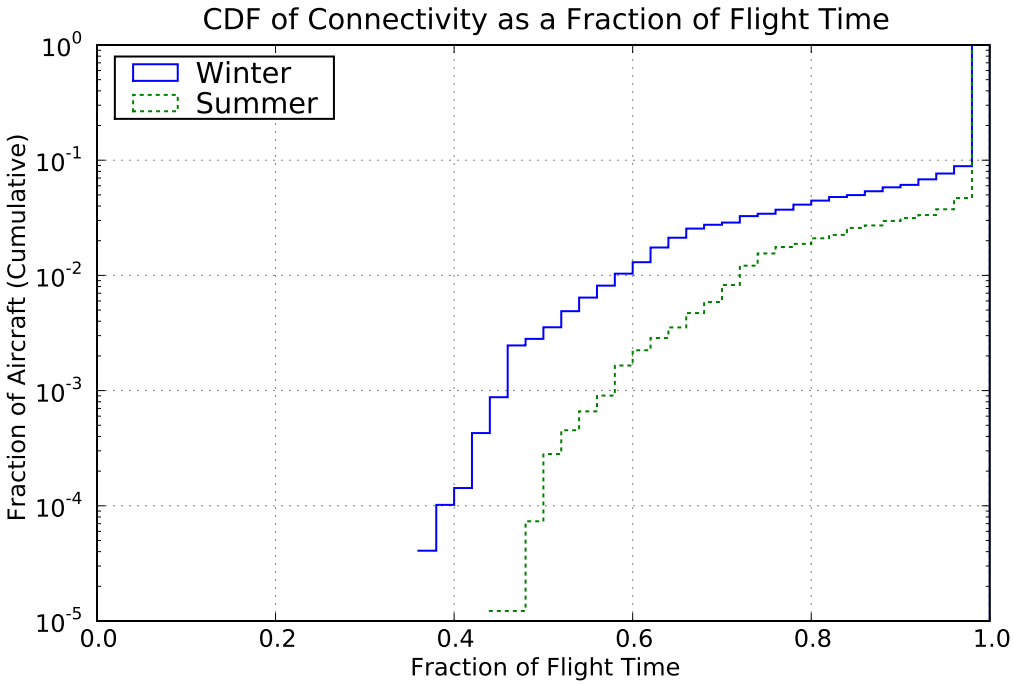


Figure 3-3: Fraction of Flight Time with Connectivity

3.2 Connected Fraction of Flight Time

In this experiment, we take a user perspective in quantifying the connectivity of the system. We measure what fraction of a typical flight has connectivity. Our notion of the fraction of flight time excludes the first and last 30 minutes of each flight when the node is considered to be unavailable for communication.

3.2.1 Results

For each delay regression of the simulation, we compute the fraction of each flight where ground connectivity is available. These fractions are aggregated into a single cumulative distribution function. The results for both the summer and winter datasets are included for comparison in Figure 3-3.

3.2.2 Discussion

As in the previous experiment, these results demonstrate that connectivity is better in the summer season due to higher flight volume. Additionally, the results show that 91.1% and 95.3% of flights in the winter and summer respectively received *uninterrupted* connectivity throughout the flight.

3.3 Connectivity's Route Dependency

In the first two experiements we established that over 90% of flights have continuous connectivity throughout their duration. In this experiment, we attempt to determine why the remaining 10% have poor connectivity.

Since our system's connectivity is dependent on the relative location of nodes, we consider the location of nodes when they are in a disconnected state. By plotting these outage events, we hope to see if these outages are correlated to flight routing.

3.3.1 Results

To generate these results, we have combined the connectivity information from our communications model with the node location information from the mobility model. If a given aircraft does not have a communications path to ground (either direct or indirect), then its location is retrieved and it is plotted. The process is repeated for all of the time steps in the simulation. The resulting plots are shown in Figures 3-4 and 3-5.

3.3.2 Discussion

The plots reveal that flights which service routes offset from the main North Atlantic corridor, which extends along the great circle between major east coast US cities and Western Europe, are more likely to suffer connectivity outages.

For example, flight that depart west coast airports such as LAX or SFO tend to take more of a polar route to Europe. Since there are far fewer flights servicing these

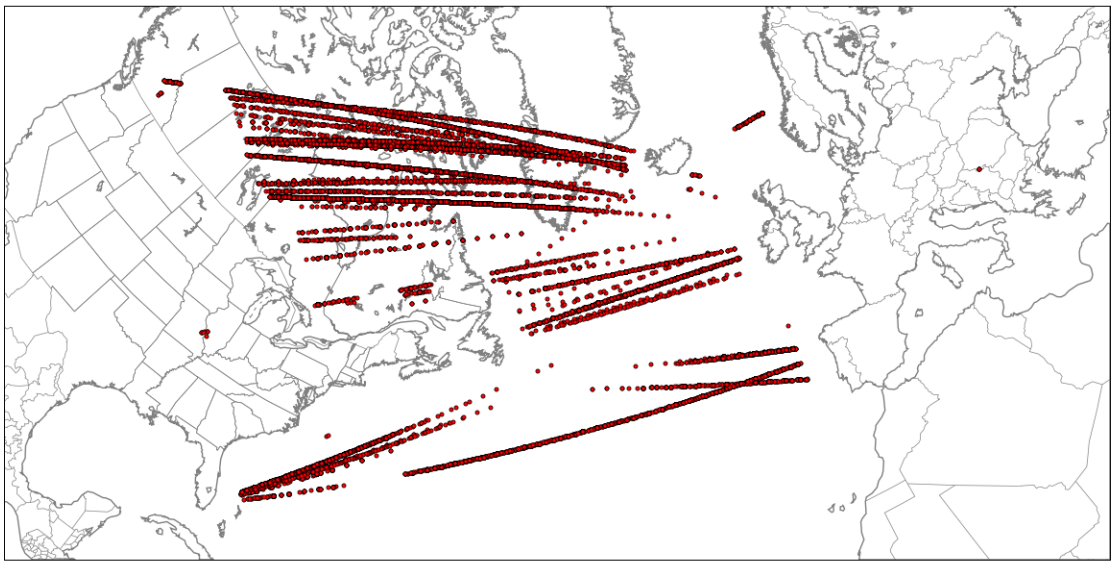


Figure 3-4: Map of Outage Events - Winter Dataset

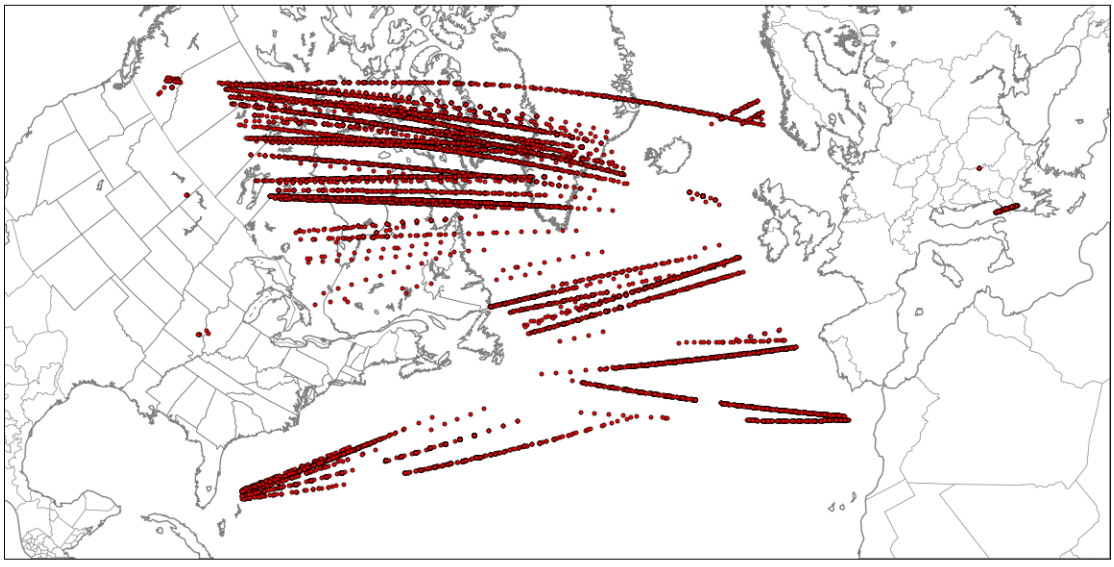


Figure 3-5: Map of Outage Events - Summer Dataset

routes, node density is low and as a consequence the MANET is fragile and prone to disconnects. Similar behavior is seen for routes across the southern portion of the North Atlantic region (e.g. Miami to Madrid).

A second point for discussion are the outages that are observed directly along the main corridor across the Atlantic. While these events are far sparser than the aforementioned route-based outages, they are still of interest. It turns out that these outages are occuring on flights which have abnormal departure times that cause them to travel well outside of the high-volume hours. Abnormal departure times can arise due to our delay model (recall, we use an expontentially distributed random variable for our delay), however, the majority of these events arise from regularly scheduled off-peak service. An example of this would be Air France flight 5 which is scheduled to depart New York bound for Paris at 1AM local time. This is a few hours after most aircraft depart the U.S. for Europe and, consequently, this flight is often positioned poorly relative to other aircraft.

Chapter 4

Network Design

Up to this point, our focus has been to understand the connectivity of the system. In Chapter 3 we showed that the system has robust connectivity in the face of randomly "jittering" flight departure times. In this chapter, we consider how the capacity resources of this network can be utilized in a *fair* and *optimal* manner. We are specifically interested in how much throughput can be sustained simultaneously to all aircraft in this system. After describing the theoretical framework for accomplishing this allocation task, we present a variety of capacity measurements from our system simulations.

4.1 Capacity Analysis Framework

In a MANET, the point-to-point links represent a *shared* resource that must be allocated to users in some way. In our system, we are concerned with dividing these resources amongst all the aircraft in a fair manner. To determine this network allocations, we start by defining the system constraints in the form of a linear program. Next, we mathematically describe our notion of fairness. We close the section by describing an algorithm which uses the linear program iteratively to determine the fair network allocation.

4.1.1 Allocation Problem Formulation

Recall from Chapter 2 that the output of our communications model is a graph which represents all feasible communication links in the system. We let $G = (V, E)$ represent this graph which is made of vertices (V: aircraft and ground stations) and weighted edges (E) representing finite-capacity point-to-point links. We assume that all links are capable of supporting *full-duplex* traffic up to the rate specified by the edge weight.

We have chosen to assess the network capacity in terms of how much throughput each aircraft can simultaneously downlink to ground stations[1]. We also assume that a given aircraft's downlink traffic can be routing to any ground station. Once this traffic reaches a ground station, it is reasonable to assume that it can be routed freely along existing ground infrastructure. Lastly, we do not constrain routing along a single path. Traffic carried by the network is free to bifurcate along as many paths as necessary.

Figure 4-1a shows a simple network consisting of five aircraft and four ground stations. The r_i variables represent the traffic being sourced by the individual aircraft. The g_i variables represent the traffic sunk by each ground station. Each edge on the graph is bidirectional and can support up to a rate of $c(u, v)$ bits per second in both directions simultaneously (i.e. full-duplex).

It is important to note that the network can be transformed into a *multi-source single-destination* network as shown in Figure 4-1b. Since all traffic is being routed to one (or more) ground nodes, they can be lumped into a single logical destination ($\triangle$) which sinks all network traffic. Similarly, we can treat the aircraft as a distributed source for network traffic. The edges from the logical source to the aircraft can be rate-constrained by the r_i variables which represent the allocation to each aircraft.

A second important property of the network is that the cost of transporting a bit of traffic on a given link is the same regardless of where the bit originated - this implies that our problem has a single commodity. If this were not true, we would be forced to formulate this problem as a multi-commodity flow problem. While these types of

[1]This is equivalent to measuring the uplink capacity due to the full-duplex link assumption.

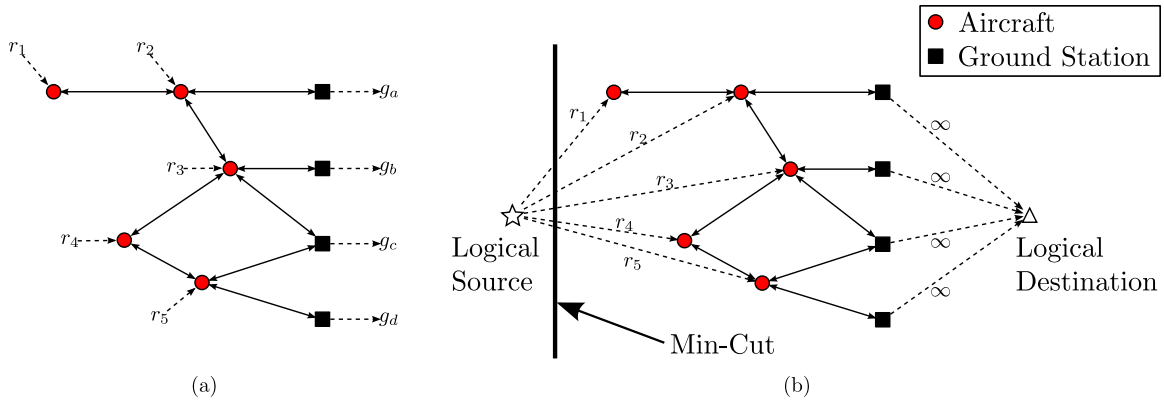


Figure 4-1: Sample network shown as (a) multiple-source multiple-destination problem and equivalent (b) multi-source, single-destination problem. Individual link capacities are not shown.

problems can be solved with LP techniques, they are more complex computationally.

Since we are faced with a multi-source single-destination flow problem, we can apply the max-flow min-cut theorem [7] to help us find the capacity of the network. Our interest is in maximizing the r_i values such that a min-cut exists between the logical source node and the aircraft nodes (shown in Figure 4-1b). Our method for maximizing the r_i's will be discussed in the next section on fairness. First, we define the linear program.

Linear Program (LP) Formulation

To mathematically constrain traffic allocations within the network, we turn to a linear programming approach. The constraints in this LP consist of link capacity constraints and flow conservation constraints. The variables in the LP represent the flow allocated on each link as well as the rate allocation for each aircraft.

We categorize each aircraft's rate allocation into one of two traffic classes: *guaranteed* or *concurrent*. Guaranteed flows represent fixed allocations to certain aircraft which *must* be met by the LP (an equality constraint). Concurrent flows represent the yet-to-be-determined aircraft flows - the goal of the LP is to *concurrently maximize* these flows. In will become clear in Section 4.1.3 why it is necessary to have two classes of flows.

LP Objective: The objective of the LP is to *maximize* λ, the concurrent flow rate of unsaturated flows.

$$\textbf{maximize}: \quad \lambda \tag{4.1}$$

Link Capacity Constraint: This constraint ensures that the total flow allocated on a given link $f(u,v)$ does not exceed the capacity $c(u,v)$ of that link.

$$0 \leq f(u,v) \leq c(u,v) \qquad \forall (u,v) \in E \tag{4.2}$$

Flow Conservation Constraint: These constraints ensure that the flow into a node is equal to the flow out of that node. We use a separate constraint for aircraft and ground nodes for clarity.

Given the set of all aircraft V_A, the following must hold:

$$\underbrace{\sum_{v \ s.t. \ (u,v) \in E} f(u,v)}_{\text{flow out of node}} - \underbrace{\sum_{v \ s.t. \ (v,u) \in E} f(v,u)}_{\text{flow in to node}} - \underbrace{r_u}_{\text{node } u\text{'s allocation}} = 0 \qquad \forall u \in V_A \tag{4.3}$$

Given the set of all ground stations V_G, the following must hold:

$$\underbrace{\sum_{v \ s.t. \ (v,u) \in E} f(v,u)}_{\text{flow to ground station}} - g_u = 0 \qquad \forall u \in V_G \tag{4.4}$$

Concurrent Flow Constraint: K_{unsat} is the set of variable flows which are to be maximized by the LP.

$$r_i \geq \lambda \qquad \forall i \in K_{\text{unsat}} \tag{4.5}$$

Guaranteed Flow Constraint: K_{sat} is the set of flows which have already been fixed (by a previous iteration of the allocation algorithm, to be described later). The i^{th} element of vector ϕ represents the guaranteed rate for flow i.

$$r_i = \phi_i \qquad \forall i \in K_{\text{sat}} \tag{4.6}$$

To summarize, the LP performs three tasks. First, it ensures that the rates allocated on each edge represent a feasible solution (i.e. a solution for which capacity and conservation constraints are met). Second, it allocates the exact rates specified in $\vec{\phi}$ to the guaranteed flows. Third, it maximizes the concurrent rate λ of all the unsaturated flows (those in K_{unsat}).

4.1.2 Fairness

One of the most popular definitions of max-min fairness was given by Bertsekas and Gallager in [2]. The essence of their fairness definition is that no flow is allocated additional capacity if that incremental allocation has a detrimental affect on other equal, or lower rate flows. We use this definition of max-min fairness.

To state this definition mathematically, we define ϕ_i as the guaranteed rate allocated to aircraft i. Thus, for a given solution to the allocation problem, we have a vector $\vec{\phi}$ which is comprised of an allocated rate for each aircraft. A given allocation vector is max-min fair if each of its elements ϕ_i can't be increased without decreasing any other node k's rate ϕ_k for which $\phi_k \leq \phi_i$. In other words, the globally max-min fair rate vector $\vec{\phi}^*$ is the lexicographically largest feasible rate vector amongst all possible values for $\vec{\phi}$. This definition was inspired by [1].

4.1.3 Max-Min Fair Algorithm

Next we describe the algorithm which is used in conjunction with the LP to determine the globally optimal max-min fair allocation vector $\vec{\phi}^*$. In [1], a centralized algorithm is presented for finding the globally optimal, max-min fair allocation vector in a weighted multi-commodity flow problem. They call this algorithm `OPT_WMMF` . As we have discussed, our problem is a degenerate case of the multi-commodity flow problem so we use their algorithm for our task.

The basic idea behind `OPT_WMMF` is to iteratively utilize an LP to slowly build up the optimal rate vector starting from the lowest rate nodes and working up to the highest rate nodes. It is somewhat similar to "water filling" algorithm that was proposed by

Bertsekas and Gallager in [2], however, it does not require a priori non-bifurcating routes for each commodity.

The `OPT_WMMF` algorithm has been customized for our application by setting all the commodity weights to one (we treat all aircraft equally) and by replacing their maximum-concurrent multi-commodity flow LP with the more simplistic single-commodity LP described in Section 4.1.1. The simplified LP does not track internal routing variables (i.e. per-edge allocations for each flow) as is necessary with general multi-commodity flow problems. These simplifications were crucial in reducing runtime for the scale of networks being considered in this work. The modified algorithm, which we call `OPT_MMF` , is shown in Algorithm 1.

Detailed Algorithm Description

The algorithm initializes by defining unsaturated flows at each aircraft (lines 2-4). Given this initial state, the algorithm enters its main loop at line 5. The condition for this main loop is that their exists flows which have not yet been saturated.

Within the main loop, the first task is to determine the maximum concurrent flow for all unsaturated commodities. Doing so, also involves making sure that all guaranteed (i.e. saturated) flows are satisfied (this is ensured by the constraints in the LP). The LP is executed in line 7 and both the concurrent flow rate λ_{iter} as well as the edge flow allocations (the $f(u, v)$'s) are returned. The next step is to determine which flows are saturated and which are not (recall, the LP does not attempt to maximize total flow, rather just the concurrent flow). The saturation test is performed in two phases. The first phase is a "cheap" test involving an easy graph connectivity test that is used to identify candidate flows which may be saturated. This is followed by a more "expensive" test based on our LP that determines whether each of these flows is indeed saturated.

The phase 1 saturation test is given in line 9. This test subtracts the edge flow allocations from the edge capacities to produce a "residual" connectivity graph. Nodes which no longer have a path to a ground station in the residual graph may be saturated. Unfortunately, this connectivity test is not bullet-proof. It is possible that

Algorithm 1 Optimal Max-Min Fair Allocation Algorithm (`OPT_MMF`)

1: *// Initialization*
2: $\Gamma_{\text{unsat}} = V_A$ *// Start with unsaturated flows at all aircraft*
3: $\Gamma_{\text{sat}} = \{\text{empty set}\}$
4: $\vec{\phi}^{iter} = \{\text{empty set}\}$
5: **while** Γ_{unsat} is non-empty **do** *// Main Loop*
6: *// Find maximum concurrent rate for unsaturated flows*
7: **perform:** $\texttt{LP}(K_{\texttt{unsat}} = \Gamma_{\texttt{unsat}}, K_{\texttt{sat}} = \Gamma_{\texttt{sat}}, \vec{\phi} = \vec{\phi}^{iter})$, yields: λ_{iter}, $f(u,v)$'s
8: *// Phase 1 ("Cheap") Saturation Test*
9: **compute** $K_{\text{tmp\_sat}}$: Subtract $f(u,v)$'s from $c(u,v)$'s, identify disconnected nodes.
// Phase 2 ("Expensive") Saturation Test
10: **for all** $k \in K_{\text{tmp\_sat}}$ **do**
11: **for all** $j \in V_A$ **do** *// Build temporary guaranteed rate vector*
12: **if** $j \in K_{\text{sat}}$ **then** *// Previously saturated flows*
13: $\phi_j^{tmp} = \phi_j^{iter}$
14: **else** *// Unsaturated flows get most recent concurrent rate*
15: $\phi_j^{tmp} = \lambda_{iter}$
16: **end if**
17: **end for**
18: **perform:** $\texttt{LP}(K_{\texttt{unsat}} = \{k\}, K_{\texttt{sat}} = V_A \setminus \{k\}, \vec{\phi} = \vec{\phi}^{tmp})$, yields: λ_{temp}
19: **if** $\lambda_{temp} = \lambda_{iter}$ **then** *// If the flow k couldn't be increased, it is saturated.*
20: $\Gamma_{\text{sat}} = \Gamma_{\text{sat}} \cup \{k\}$
21: $\Gamma_{\text{unsat}} = \Gamma_{\text{unsat}} \setminus \{k\}$
22: $\phi_j^{iter} = \lambda_{iter}$
23: **end if**
24: **end for**
25: **end while**
26: **return** $\vec{\phi}^* = \vec{\phi}^{iter}$

the LP's solution has happened to allocate flows in a manner which cause there to be zero residual capacity for this particular node. Stated differently, an alternate routing solution in the graph may be able to give this node more capacity. The node is added to a temporary set called $K_{\text{tmp\_sat}}$ and the phase 1 saturation test is repeated for all nodes.

The phase 2 saturation test is run against all flows $k \in K_{\text{tmp\_sat}}$. This test uses the LP to determine if it is possible to increase flow k's allocation by rerouting all other flows. This is completed under the assumption that all other flows are *saturated* either at their rate specified in $\vec{\phi}^{iter}$ or, if not, the most recently found concurrent rate λ_{iter}. This temporary saturated rate vector $\vec{\phi}^{temp}$ is constructed in lines 11-17 and the LP is executed with the new parameters in line 18. Since k is the only unsaturated flow in this run of the LP, the resulting value for maximum concurrent flow (λ_{temp}) is the maximum achievable rate found for flow k. If $\lambda_{temp} > \lambda_{iter}$ then the LP was able to find an alternate routing with higher rate for flow k and thus it is not saturated. Alternatively, if the value of $\lambda_{temp} = \lambda_{iter}$ then flow k is saturated under all possible network routings; it is moved into set Γ_{sat} and the saturated rate vector is updated accordingly (lines 20-22). The process continues until Γ_{unsat} is empty at which point $\vec{\phi}^{iter}$ contains max-min fair allocated rates for all flows in the network.

Correctness Justification

The key to proving correctness of this algorithm lies in showing that the two-phase saturation test is correct and that the algorithm is guaranteed to complete in finite time. Despite our simplification of the weighted max-min fair allocation algorithm `OPT_WMMF` given in [1], their correctness proof still holds for our `OPT_MMF` algorithm. The primary difference between our `OPT_MMF` algorithm and the more general `OPT_WMMF` algorithm is that we have applied the problem to a single-commodity flow problem as described in Section 4.1.1. This simplification was necessary in order to reduce the runtime of the algorithm so that it would work well with our large scale networks.

The correctness proof that has been provided in [1] is split into two lemmas. First, they show by contradiction that a flow identified as saturated can't increase its

allocation under any routing combinations [1, Lemma 1]. Since both algorithms share the same LP optimization goal (to maximizing the concurrent flow of unsaturated commodities), this lemma applies to both algorithms.

Second, they use induction to show that the algorithm yields a allocation vector which meets the max-min fair definition [1, Lemma 2]. Again, the fact that both algorithms share the same optimization goal implies that Allalouf's inductive proof also holds for our simplified version of the algorithm.

Lastly, the finite completion time requirement can be shown with an additional proof by contradiction. During each iteration of the main loop (lines 5-25), at least one flow is identified as being saturated. If this doesn't occur, it implies that the LP has not reached its objective of concurrently maximizing the unsaturated flows. Hence, the algorithm will complete in at most $\mid V_A \mid$ iterations of the main loop.

Implementation Details

We have implemented the `OPT_MMF` algorithm using a combination of the Python programming language and a commercial LP solver (AMPL/CPLEX). The Python script preprocess the connectivity graph before it is handed off to the `OPT_MMF` algorithm. This preprocessing involves removing aircraft which are disconnected (i.e. those which do not have a path to ground) as well as converting the connectivity graph from a undirected to a directed graph. Next, the script interacts with the AMPL/CPLEX to repeatedly feed it LP model data (graph structure, saturated flow values, etc.) and retrieve LP solution data. This algorithm runs in approximately 10 minutes on our largest connectivity graph (summer dataset, peak flight volume time of day). The algorithm is applied to each time "snapshot" in our simulation across all delay scenario regressions.

4.2 Experimental Results

In this section a variety of capacity-based experimental results are presented. As with the connectivity analysis, we consider both overall system performance as well

as perceived performance for specific flights. When interpreting this data it is very important to keep in mind that the absolute capacity values are a direct result of the assumptions made for the physical link design (antenna gains, channel BW, etc). For this reason, they are most valuable for observing the dynamic behaviors of the system and for comparative purposes. If a real system were being designed, the system link assumptions could be adjusted to best suit the needs of the system designer and the end user.

4.2.1 Rate versus Time

In this experiment we measure the max-min fair downlink rate that can be supplied to a particular flight during its voyage from origin to destination. We have computed the max-min fair allocations for all aircraft across a 24-hour period with 10 minute time steps. From this database of allocation data, we have chosen a few representative flights to show the capacity capabilities of the network. The vantage point on the allocation data represents an "end user perspective" just as was done for connectivity in Section 3.2.

Results

For these results, we make use of box plots to display the data. This type of plot has been chosen because it conveys more information than simple arithmetic means and unit standard deviation error bars. This is especially important for our capacity values which tend vary across a few orders of magnitude.

We use a box plot convention that is similar to Tukey's original definition in [23]. Our plots have a box which extends from the first to the third quartile (the interquartile range). The red line dividing the box represents the median of the data. Data points which lie outside of the interquartile range are plotted individually as outliers.

In Figures 4-2a and 4-2b we show a typical flight in both the summer and winter seasons. This flight, United 950 from Washington, DC to Brussels was selected because it has a departure time which, on average, places it in the middle of the "pack"

of eastbound aircraft.

The opposite situation is given in Figure 4-3. This flight from Paris to New York which has a relatively early departure time which places it on the leading edge of the westbound flow of traffic.

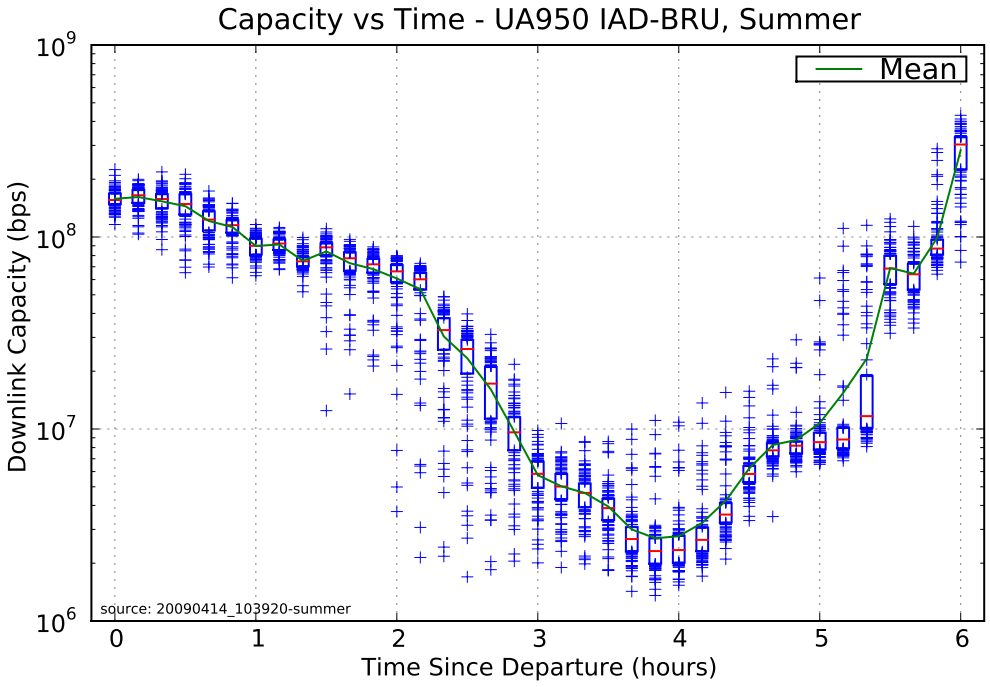
Discussion

The most obvious feature of all three capacity plots is the decrease in capacity during the middle portion of each flight. This corresponds to the times when the aircraft is in the oceanic phase of its flight, out of range of any ground infrastructure. The two plots for United 950 (one from winter, one from summer) show that the downlink capacity is roughly the same for both seasons. This is significant because the number of aircraft being supported by the network is drastically different for the two seasons: there are approximately twice as many aircraft in the air during the summer season than during the winter season. This provides evidence that the achieved downlink rates are not sensitive to the number of aircraft participating in the system.

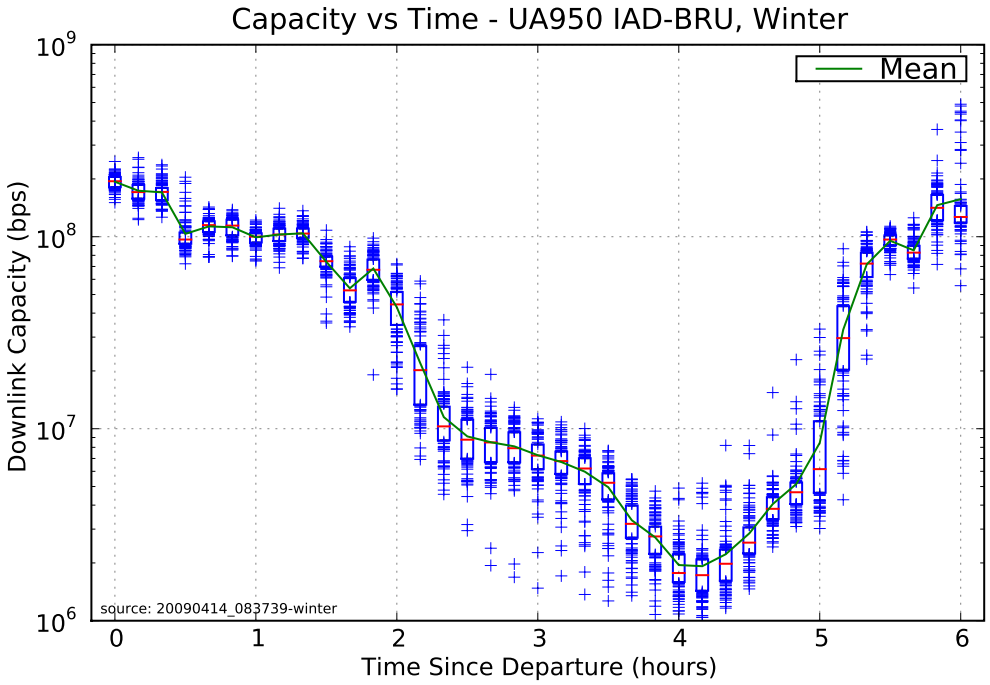
As mentioned before, United 950's departure time and route imply that it is generally well-positioned in the middle of the large group of eastbound flights each night. The capacity plot for AF22 has been provided to show the capacity available to an aircraft which has less-fortunate positioning due to its early departure time. Despite this disadvantage, the flight still tends to have megabit-class downlink rates for the duration of its journey. The sharp increase in allocated rate approximately 4 hours into the flight coincides with the time at which that aircraft comes within range of ground stations on the other side of the ocean.

4.2.2 Rate versus Position

In this experiment, we consider the geographic dependence of capacity. In the previous experiment, we showed plots which indicate that flight downlink capacity is at its lowest during the middle of the flight (when it aircraft is located over the ocean). The goal of this experiment is to make this observation more concrete by plotting



(a) Summer schedule data



(b) Winter schedule data

Figure 4-2: Downlink capacity versus time for UA950 averaged over 100 delay scenarios. The box represents the interquartile data points, the red line inside of the box represents the median. The plus (+) markers are outliers falling outside of the interquartile range. The arithmetic mean is also plotted.

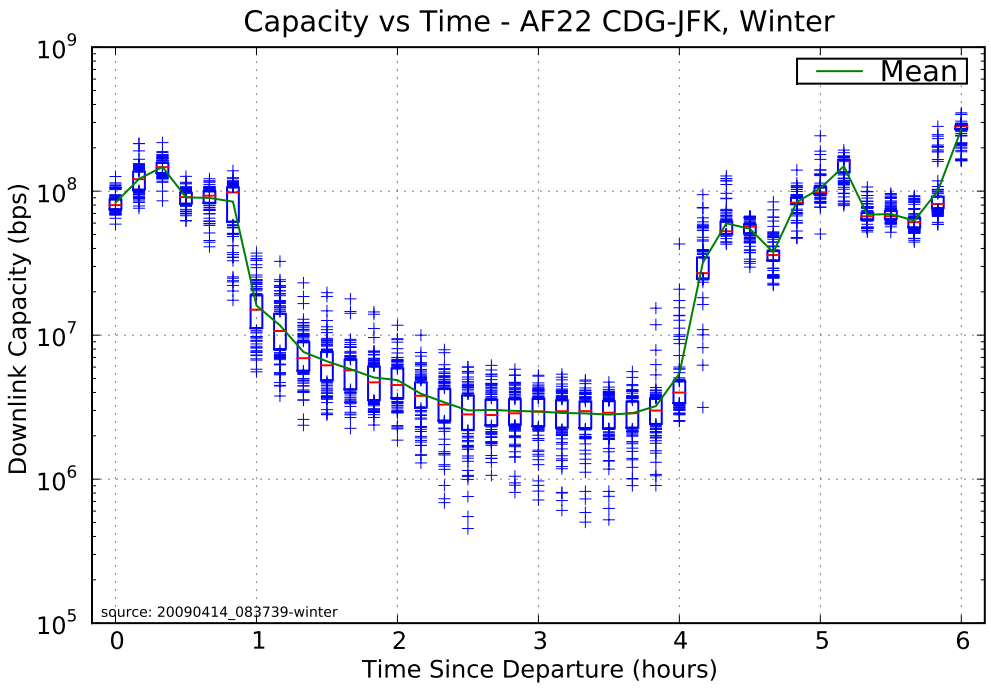


Figure 4-3: Downlink capacity versus time for AF22, winter data set. This flight is on the leading edge of westbound flights. The large increase at approximately 4 hours is when this flight comes within range of North American ground stations.

Figure 4-4: Downlink capacity contour plot for winter data set at 5:00UTC. Averaged over all simulation runs. Capacity measured in Mbps.

aircraft downlink rates versus position for certain times of day.

Results

One method for displaying this data would be to make a scatter plot which shows the capacity allocated to each aircraft at its location. This technique yields plots which are very busy and hard to interpret especially when multiple regressions of the simulation are combined on one plot. As an alternative, we have turned to contour plots which show lines of equal capacity and represent the allocation data from 100 delay scenario regressions of the simulation.

In Figures 4-4 and 4-5, we show two of these contour plots, one of each season. Both plots are from 5:00UTC which roughly aligns with one of the high-volume times of day (the east-bound flow). The plots represent the aggregation of flight position and capacity data for all simulation regressions (various delay scenarios).

Discussion

As expected, the results confirm that capacity tends to be lower during the oceanic phase of the flight. The plots also indicate that flights which travel outside of the main

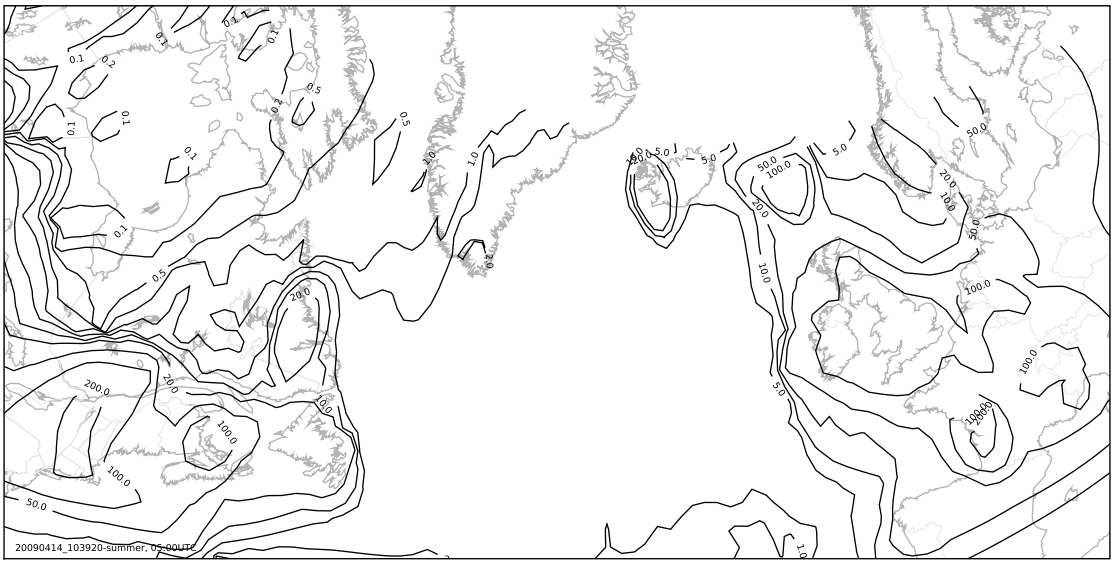


Figure 4-5: Downlink capacity contour plot for summer data set at 5:00UTC. Averaged over all simulation runs. Capacity measured in Mbps.

great circle corridor from the Northeast US to Europe generally have low downlink rates. For example, flights which traverse Greenland are those which travel to and from western cities in North America. Since there are relatively few flights on these routes, node spatial density tends to be low and capacity suffers.

An additional qualitative result comes from comparing the two plots (summer and winter). In general the per-node capacity of the system is more or less the same for both seasons despite the nearly two-fold difference in the number of flights. This is a result of the scalable nature of MANET's.

4.2.3 Mid-Ocean Downlink Capacity

In this final capacity experiment, we further characterize the expected downlink capacity that can be provided to aircraft when they are positioned at the middle of the ocean. We use the same simulation data as before, however, we filter it for values which correspond to when flights are crossing the 35°W meridian. This meridian roughly corresponds to the midpoint of each flight's oceanic crossing.

Results

In Figure 4-6 we give the PDF of aircraft downlink capacity when each flight crosses the 35°W meridian. This PDF was generated from 100 independent delay scenarios for both the winter and summer flight schedules.

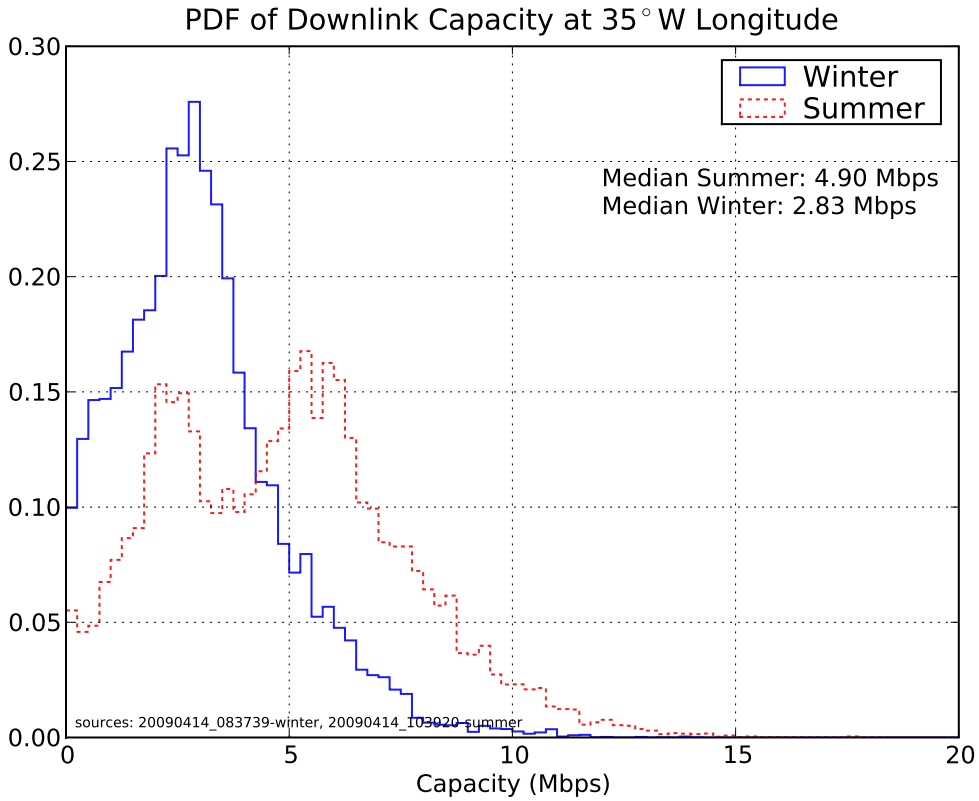


Figure 4-6: PDF of mid-ocean downlink capacity for all flights when crossing 35°W. Includes data from 100 delay scenario.

Discussion

The results in Figure 4-6 show that both seasons tend to yield mid-ocean data rates of approximately 1-10 Mbps just as was observed in the earlier experiments. These results also show that the median capacity is higher for the summer schedule despite the two-fold increase in the number of scheduled flights.

The fact that the PDF for the summer dataset is bimodal is quite interesting. Further investigation reveals that the peak around 3 Mbps is comprised mainly of

eastbound aircraft whereas the peak at 6 Mbps is made up mostly of westbound aircraft. Figure 4-7 gives a histogram that shows the breakdown of these two components. We suspect that this is being caused by differences in the *relative* departure times of eastbound versus westbound flights. It is likely that the westbound flight schedules tend to result in a better node placement (from a capacity sense) of aircraft.

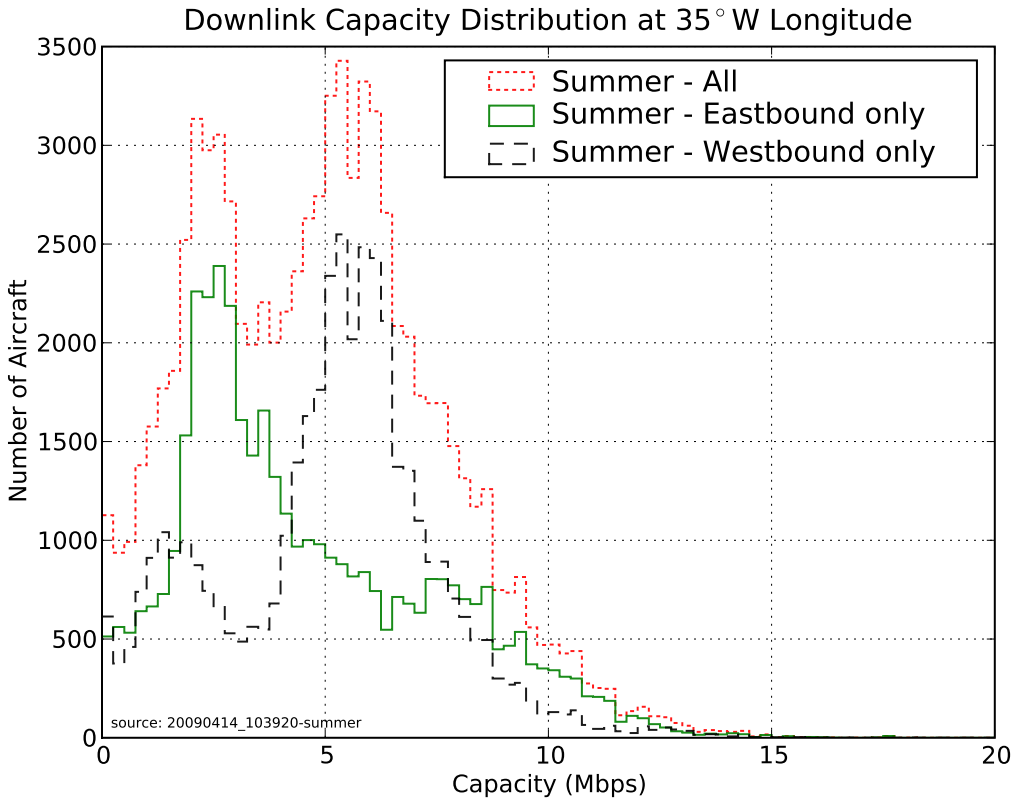


Figure 4-7: Histogram showing differences between eastbound and westbound flights in the summer dataset.

Chapter 5

Implementation Study

In this chapter, we delve into some implementation trade-offs which we feel would be of great interest to those attempting to build the system. We start by quantifying the number of simultaneous connections each aircraft must be capable of supporting so as to not compromise system connectivity. The second section describes a new antenna model which would likely be easier to implement on an aircraft platform. We provide simulation results which quantify the impact of these antenna restrictions on system capacity and connectivity. We close the chapter by simulating a hypothetical deployment scenario and show how the system capabilities evolve as it is deployed.

5.1 Concurrent Connection Requirements

When designing MANET's, it is important to understand how many point-to-point links each node in the network must be capable of supporting concurrently. This system parameter directly impacts the design of the physical layer (e.g. simultaneous beams to different neighbors). In our simulations we have seen that the degree of nodes in the connectivity graph can be quite high - often around 30. These densely-connected graphs have resilient connectivity properties and ample capacity, however, it is important to realize that the system can be fully connected by only using a subset of these links.

In this experiment we wish to determine the minimum number of beams each

aircraft must support in order to maintain the overall connectivity of the system. Given our full connectivity graph $G = (V, E)$ from the communications model, it is possible to compute a simple spanning tree which achieves this connectivity goal. Unfortunately, standard spanning tree algorithms place no constraint on the degree of nodes within the resulting tree.

A special type of spanning tree known as a minimum-degree spanning tree (MDST) incorporates the desired degree restrictions. In the MDST problem, the goal is to find a spanning tree which achieves the following goal:

$$\textbf{Minimize:} \quad k = \max\{d_i\} \quad \forall i \in V. \tag{5.1}$$

where d_i is the degree of node i in the tree. Although standard spanning trees (such as the minimum spanning tree) are trivial to calculate in polynomial time, the MDST problem is NP-complete [9].

Fortunately, approximation algorithms have been devised to solve this class of problem within a proven performance bound. In [8], Fürer and Raghavanachari describe an algorithm which finds a spanning tree with maximal degree $k \leq \Delta^* + 1$ where Δ^* is the maximum degree of the optimal degree-minimized spanning tree.

Their Δ^*+1 algorithm starts by constructing an arbitrary spanning tree T from the full graph G. Next, the algorithm iteratively rewires T by removing edges incident to high-degree nodes and replaces them with other edges from G. This process repeats until degree reduction is no longer possible given the edges available in G. This algorithm runs in $O(\Delta^* \log n)$ time.

5.1.1 Results

The MDST approximation algorithm from [8] was implemented in our simulation environment and was run against the system connectivity graphs at each simulation time point. This was completed for each of the 100 delay scenario regressions. The degree distribution (i.e. the fraction of nodes of each degree) was measured from each of these spanning trees and the distributions were aggregated into Table 5.1.

Node Degree	Fraction of Nodes
0	0.0069
1	0.0993
2	0.8470
3	0.0468

(a) Winter Dataset

Node Degree	Fraction of Nodes
0	0.0027
1	0.0687
2	0.8880
3	0.0406

(b) Summer Dataset

Table 5.1: Node degree distribution of MDST spanning trees.

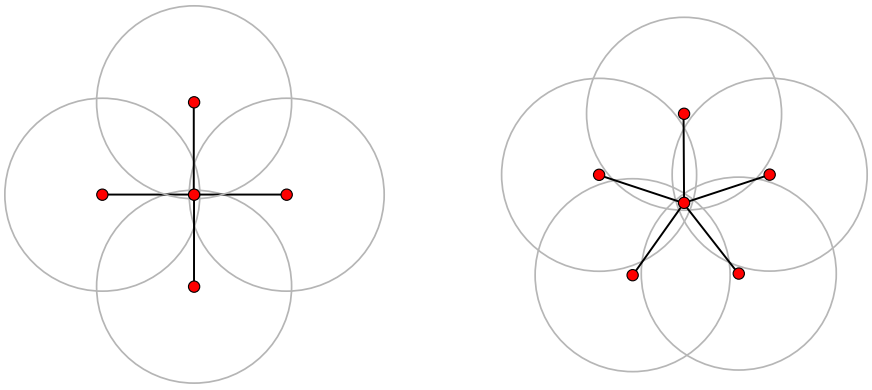
5.1.2 Discussion

The results given in Table 5.1 show that it is possible to maintain connectivity with a few as three beams per aircraft. Considering the omnidirectional antenna coverage and uniform range capabilities assumed thus far, this isn't entirely surprising.

Omnidirectional antenna coverage implies that groups of node located near one another will tend to be strongly connected (i.e. the connectivity graph will have an edge between all pairs of nodes in the vicinity). In strongly connected graphs, it is trivial to construct low-degree spanning trees. That said, there are still pathological node placement cases which result in spanning trees with maximum degree higher than three.

Figure 5-1 shows some of these obscure cases. The figure shows a central node surrounded by either 4 or 5 neighbors. The circles represent the maximum communications range of the neighbors. Note that they all have links to the central node, however, they can't establish links with one another directly. This forces the degree of the central node to be k in the optimal minimum-degree spanning tree.

By considering the geometry of the problem, it is easy to see that it is not possible to add another in order to increase the degree of the central node above $k = 5$. Adding an additional node that is within range of the central node, will also be within range of at least one of the neighboring nodes. This extra neighbor-to-neighbor edge relaxes the high degree requirement on the central node. A formal proof of this claim is beyond the scope of this section, however, it is simple to construct using the triangle inequality.



(a) k=4 case (b) k=5 case

Figure 5-1: Node positioning cases which result in high degree spanning trees. The circles delimit the communication range of the nodes.

5.2 Antenna Restrictions

In this section, we address some of the difficulties associated with designing an antenna system for the application. We consider the impact of these restrictions on the overall connectivity and capacity of the system. The goal of this section is to determine whether or not full omnidirectional antenna coverage is necessary for a successful implementation.

In Section 2.4 we described a simple antenna model based on a modestly dimensioned parabolic antenna. While this antenna is highly directional, we assumed that it could be pointed in any direction in order to achieve maximum gain with the link endpoint. On an aircraft, it is impossible to achieve fully omnidirectional coverage with such an antenna due to shadowing from the fuselage. Instead, it is necessary to either employ a system of antennas or to give up some portion of the omnidirectional coverage.

Over the course of the system simulations, we observed some general trends in how links are allocated within the network. Specifically, we noticed that heavily allocated links are generally those which are directed along an aircraft's direction of travel (either forward, or rearward). Intuitively, this makes sense since the aircraft are trying to communicate with the ground stations along the coast during their ocean crossing.

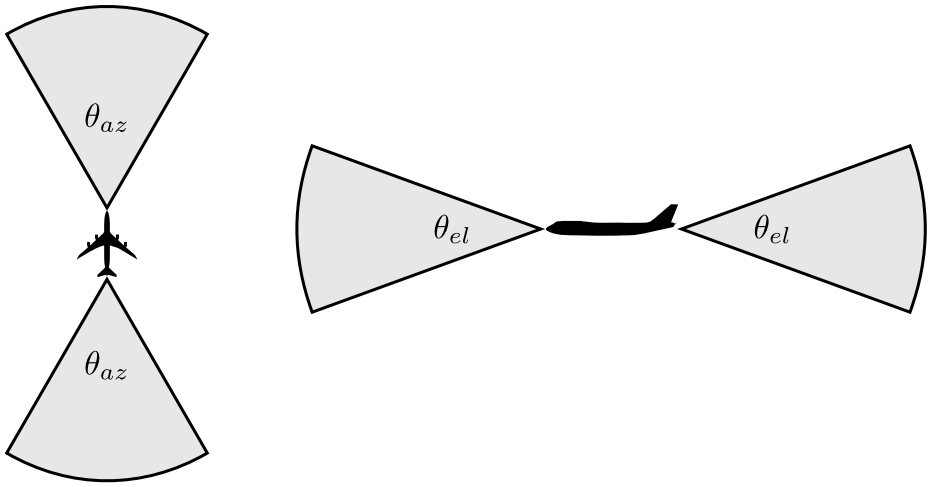


Figure 5-2: Diagram of antenna steering restrictions.

This observation was the inspiration for a restricted antenna steering model which only allows forward and rearward facing connections. These antenna steering restrictions are shown in Figure 5-2. The model is parametrized by an azimuth beam width, θ_{az}, and a elevation beam width, θ_{el}. The forward and rearward beams have identical parameters.

Given the locations of all aircraft, it is straightforward to calculate the elevation and azimuth link relative to each aircraft's frame of reference. Using these link geometry parameters, we filter the connectivity graph to remove links which do not meet the antenna restriction requirements. The resulting subgraph is then run through the connectivity and capacity analysis.

5.2.1 Results

We have chosen three different antenna restriction scenarios where $\theta_{el} = 30°$ and $\theta_{az} = \{30°, 60°, 90°\}$. Figure 5-3 shows the overall system connectivity fractions for these configurations plus the unrestricted scenario. Table 5.2 shows the fraction of flights which enjoy *uninterrupted connectivity* from takeoff to landing. Finally, a comparison of the mid-ocean downlink capacities is given in Figure 5-4. All results represent the aggregation of 100 different delay scenarios.

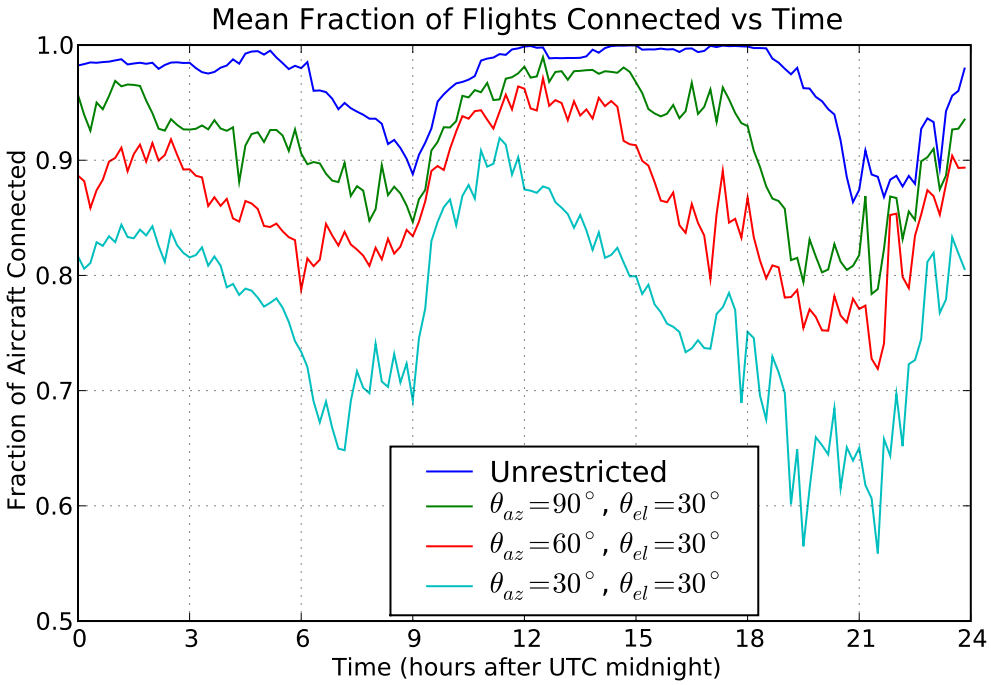


Figure 5-3: Impact of antenna steering restrictions on overall system connectivity. Each trace represents the mean connectivity fraction over 100 delay scenarios. Winter schedule data is used. The vertical ordering of traces in legend match the ordering in the plot.

θ_{az}	**Fraction of Flights**
30°	0.5926
60°	0.7223
90°	0.7809
Unrestricted	0.9113

Table 5.2: Fraction of flights with uninterrupted connectivity for various antenna restriction levels. Measured across 100 independent delay scenarios. $\theta_{el} = 30°$ for all cases.

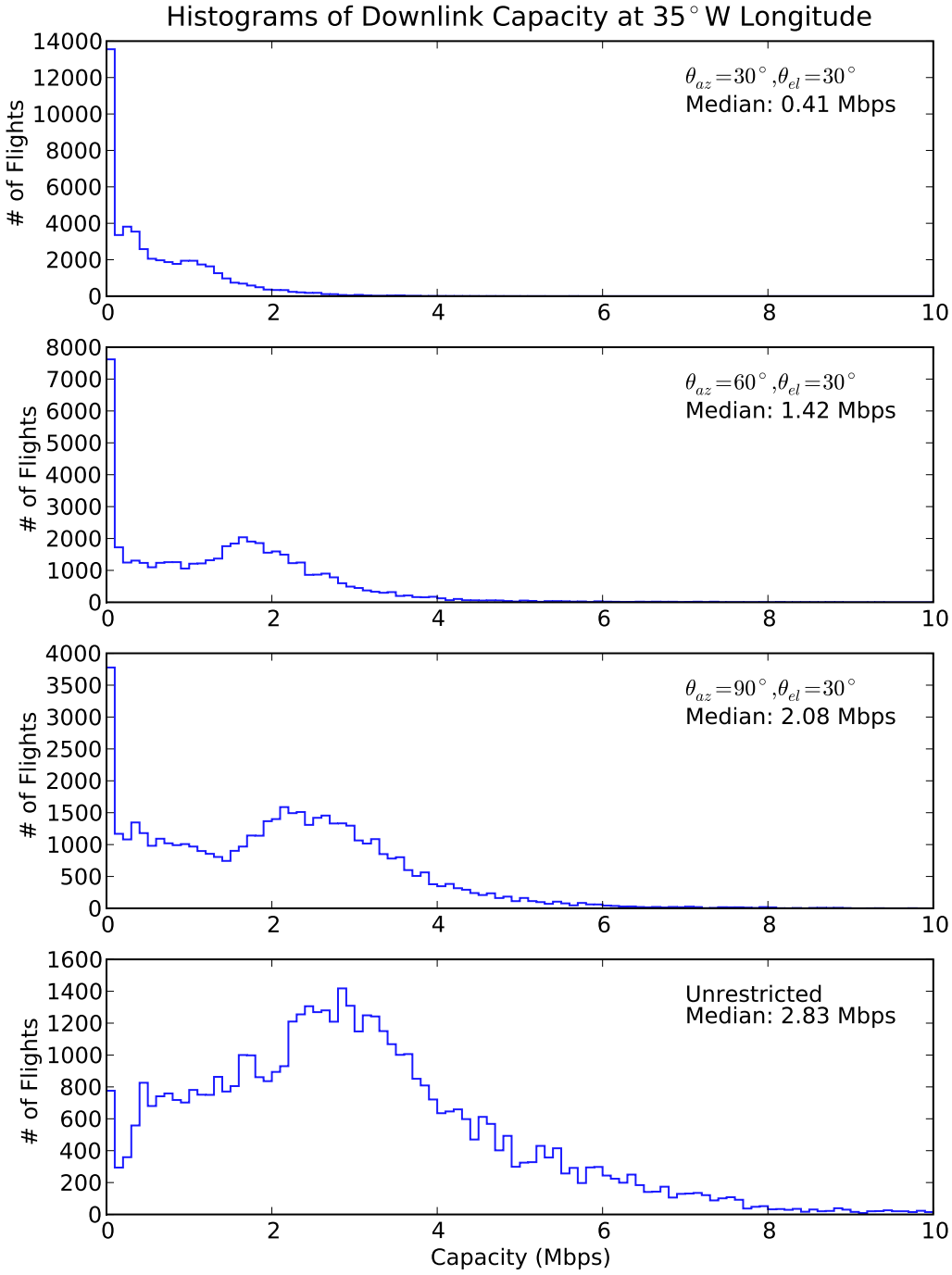


Figure 5-4: Mid-ocean downlink capacity for the antenna restriction scenarios. Each histogram represents the distribution of aircraft downlink capacity over 100 delay scenarios. Winter schedule data is used.

5.2.2 Discussion

In this experiment we have drastically reduced the steering range of the antenna system. As expected, the system's connectivity and capacity metrics have decreased in response to the restrictions.

In the most restrictive case where $\theta_{az} = 30°$, we can see that overall system connectivity fluctuates between 0.6 and 0.9 which is much poorer than the unrestricted scenario. That being said, Table 5.2 shows that 59% of the flights in this scenario enjoy uninterrupted connectivity. These connectivity properties are very encouraging, however, the downlink capacities in this very restrictive case are quite low - generally less than 1 Mbps. The $\theta_{az} = 90°$ scenario shows much improved capacity while still providing a substantial reduction in the antenna steering requirements.

Further investigation of this trade between antenna capabilities and system performance is an area of future work. An additional issue that must be investigated is how the antenna is steered. Due to the fact that each aircraft must simultaneously support multiple connections, it is unlikely that a mechanically steered antenna will suffice. Technologies such as phased-array antennas, which have fast steering capabilities, may be well-suited for this application.

5.3 System Deployment

A significant challenge with implementing this communication system lies in formulating a deployment strategy. The downside of a cooperative system such as our MANET is that they have limited utility in the early stages of deployment - there simply aren't enough nodes to form a cohesive network. In this section, we measure the connectivity and capacity properties of the system as it is deployed.

Thus far, our system model has incorporated all regularly scheduled commercial flights in the North Atlantic market. During the initial roll-out of the system only a subset of these flights will be retrofitted with the communication system. Perhaps the system will only be supported by a single airline, or alliance of airlines. To capture this deployment process, we have modified the simulation so that it only includes

Deployment	Fraction of Flights
10%	0.2504
20%	0.5877
40%	0.7670
60%	0.8297
80%	0.8783
100%	0.9113

Table 5.3: Fraction of flights with uninterrupted connectivity. These values were measured across 100 independent delay scenarios at each deployment percentage.

a specified percentage of the scheduled flights. In each of our 24-hour simulation regressions, we randomly select this percentage of flights from the full flight database. Flights not selected are omitted from the simulation. Once the flights have been selected, the simulation is run just as in Chapters 3 and 4. To guard against biasing the results with a lucky selection of flights, we repeat this random selection of deployed flights for each delay scenario.

5.3.1 Results

We measured system performance at six different deployment percentages: 10%, 20%, 40%, 60%, 80%, and 100%. The 100% deployment case is identical to results presented earlier in the thesis, however, it is presented again for comparative purposes. In Figure 5-5 we show the system connectivity fraction averaged over 100 representative delay scenarios. Error bars have been omitted for clarity, however, the general trend is that connectivity variance increases for the lower deployment percentages. The bottom and top-most traces correspond to the 10% and 100% deployment scenarios, respectively; the intermediate cases lie in the expected order in between.

As we discussed in Section 3.2, it is also important to consider the "user perspective" of connectivity in addition to the overall system connectivity. In Table 5.3 we show the fraction of flights which have *uninterrupted connectivity* from takeoff to landing.

Figure 5-6 encompasses the capacity results for the various deployment scenarios. We have opted to only present the distribution of mid-ocean downlink capacity as

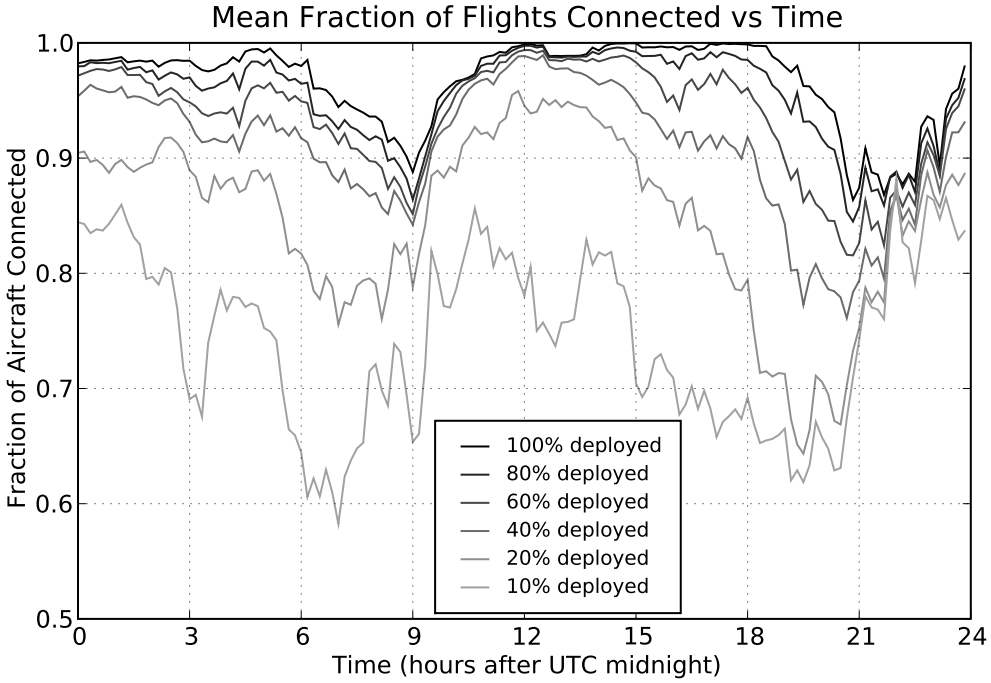


Figure 5-5: Overall system connectivity at various deployment percentages. Each trace represents the mean connectivity fraction over 100 delay scenarios. Winter schedule data is used. The vertical ordering of traces in legend match the ordering in the plot.

this tends to be the most stressing case in the simulation. A separate histogram is provided for each of the deployment percentages, the median downlink capacity is also provided on the plots.

5.3.2 Discussion

The system connectivity and capacity results respond in the expected manner as the deployment percentage is changed. That is, both connectivity and capacity improve as the number of retrofitted aircraft is increased.

The connectivity results are encouraging since they show that system connectivity fractions in excess of 0.8 are achieved even if a little as 20% of scheduled flights are retrofitted. Additionally, we can see from Table 5.3 that even though the 20% deployment scenario never reaches full system connectivity, 58% of the flights enjoy uninterrupted connectivity during their journey across the ocean. If flights on sparely-traveled routes are omitted (e.g. flights to/from the west coast of the US) then this percentage would improve even more.

The mid-ocean capacity results are mixed. The positive result is that as deployment progresses, the capacity of the system certainly increases. This result corroborates the winter/summer season capacity comparisons that were made in Chapter 4. Unfortunately, the deployment capacity results also show that capacity is quite poor at low deployment percentages. An open question is whether this initial limitation is significant enough to deter airlines and aircraft manufacturers from invested in such a system.

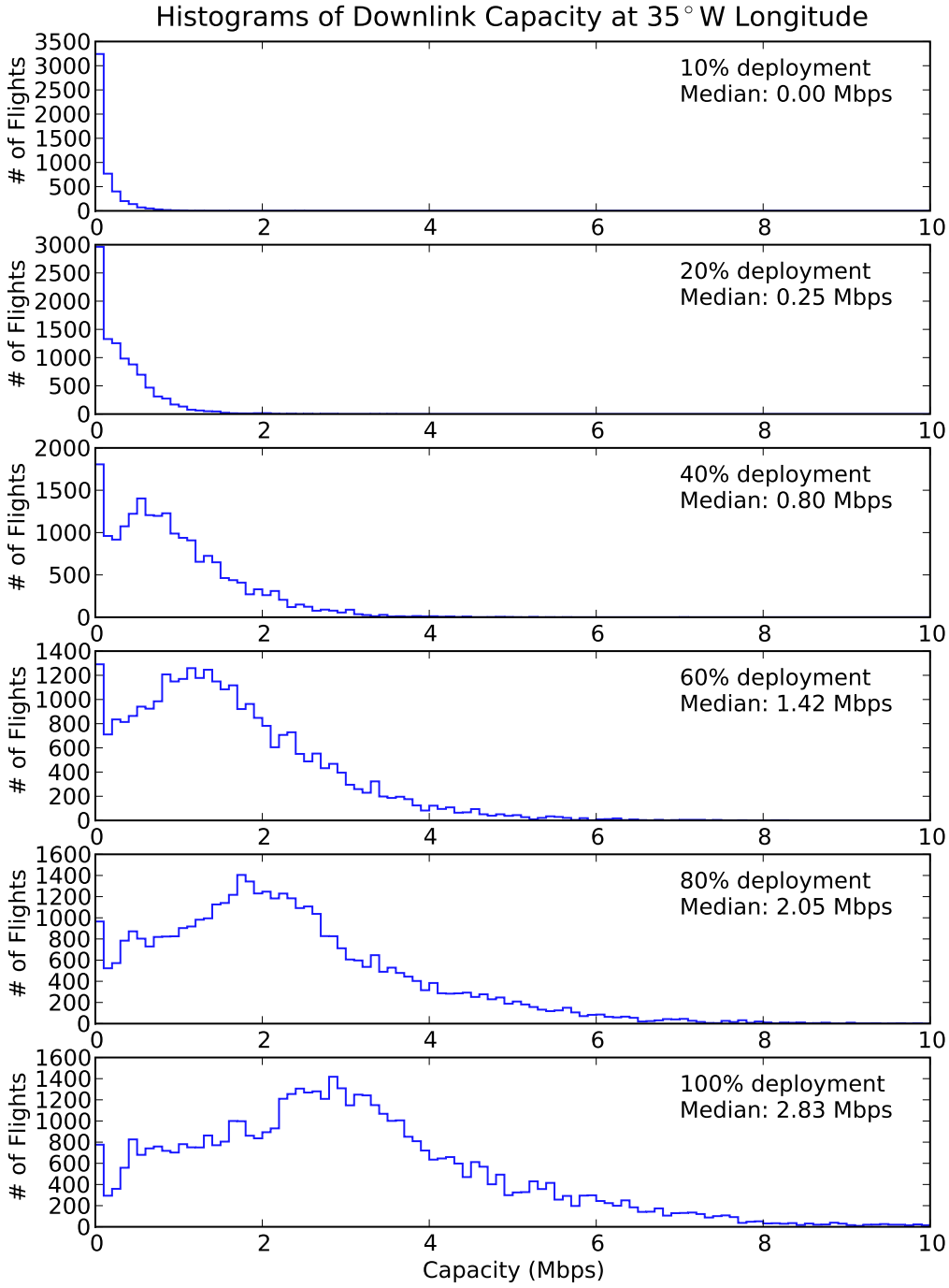


Figure 5-6: Mid-ocean downlink capacity during deployment. Each histogram represents the distribution of aircraft downlink capacity over 100 delay scenarios. Winter schedule data is used.

Chapter 6

Conclusions & Future Work

In this thesis we have completed a broad investigation into the feasibility of using MANET's to provide communications infrastructure for aircraft on transoceanic routes. We first formulated a model of the system which included mobility prediction based on airline schedule data as well as a SNR-based communications link model. This model allowed us to predict the position of all aircraft as well as all feasible communication links between both aircraft and ground stations.

Using these communication graphs, we conducted an analysis of the connectivity of the system. We presented connectivity metrics which show the overall capabilities of the system as well as a "user perspective" of connectivity behavior. Most importantly, we have shown that the connectivity of the system is robust in the face of random flight delays.

The second group of results focus on the capacity of the network. We have applied an optimal max-min fair allocation algorithm in order to establish an upper bound on network capacity. The resulting capacity values show that megabit-class data rates are achievable under our link assumptions. One caveat of these results is that they are an upper bound - actual system performance may be lower due to implementation losses. Despite this caveat, the results are still valuable in that they give insight into a the dynamic behavior of the system, namely, the capacity results show that capacity scales well as additional aircraft are added to the system.

The final group of results explore some fundamental implementation issues. We

establish a bound on the minimum number of simultaneous connections each aircraft must support using in order to maintain system connectivity. We also consider the impact of antenna steering restrictions on the performance of the system. A final implementation issue that is investigated is the behavior of the system as it is deployed.

There are many avenues for future work that can grow from this thesis. From the networking perspective, we have yet to consider how traffic can be routed in a distributed manner. There is a wealth of research in this field, however, it is unclear what techniques are best suited for this application. From an aircraft integration standpoint, it would also be interesting to determine what sort of steerable antenna designs are available and how they could fulfill the requirements of this system. Lastly, we have not touched the design of the physical layer. Issues such as adaptive modulation design, and medium sharing have not been resolved and some of these may be quite challenging.

In closing, we'd like to point out that while our focus was to create a system for passenger communications, it is likely that such a system would be of great interest to both airline operators and air traffic control organizations. Existing communication infrastructure within the oceanic regime is extremely limited so a robust high-capacity communications framework would be of great value to these groups. Having such a system could allow for more efficient routing of aircraft and better use of available airspace. That said, using the system for any "safety of flight" purpose will require a much more rigorous reliability analysis; this too is an avenue for future research.

Bibliography

[1] M. Allalouf and Y. Shavitt. Centralized and distributed algorithms for routing and weighted max-min fair bandwidth allocation. *Networking, IEEE/ACM Transactions on*, 16(5):1015–1024, Oct. 2008.

[2] D.P. Bertsekas, R.G. Gallager, P. Humblet, Center for Advanced Engineering Study, and Massachusetts Institute of Technology. *Data networks*. Prentice-hall Englewood Cliffs, NJ, 1987.

[3] Boring. Boeing to discontinue connexion by boeing service. Press Release, August 17, 2006. [Online; accessed 12-Dec-2008].

[4] N.V. Campos. Encouraging Technology Transition through Value Creation, Capture and Delivery Strategies: The Case of Data Link in the North Atlantic Airspace. Master's thesis, Massachusetts Institute of Technology Cambridge, MA 02139 USA, 2008.

[5] Boeing Corporation. Connexion by Boeing: History, 2007. [Online; accessed 12-Dec-2008].

[6] G.I. Evenden. *Cartographic Projection Procedures for the UNIX Environment: A User's Manual*. U.S. Geological Survey, 1990.

[7] L.R. Ford and D.R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8(3):399–404, 1956.

[8] M. Fürer and B. Raghavachari. Approximating the minimum degree spanning tree to within one from the optimal degree. In *Proceedings of the third annual ACM-SIAM symposium on Discrete algorithms*, pages 317–324. Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 1992.

[9] MR Garey, D.S. Johnson, R.C. Backhouse, G. von Bochmann, D. Harel, CJ van Rijsbergen, J.E. Hopcroft, J.D. Ullman, A.W. Marshall, I. Olkin, et al. *Computers and Intractability: A Guide to the Theory of*. Springer, 1979.

[10] National Frequency Planning Group. United Kingdom Frequency Allocation Table. Technical report, Cabinet Official Committee on UK Spectrum Strategy, 2008.

[11] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, Pasadena, CA USA, August 2008.

[12] G. D. Gierhart & M.E. Johnson. The IF-77 electromagnetic wave propagation model. Technical report, U.S. Department of Transportation, Sept 1983.

[13] H.J. Liebe. MPM: An atmospheric millimeter-wave propagation model. *International Journal of Infrared and Millimeter Waves*, 10(6):631–650, 1989.

[14] Office of Spectrum Management. U.S. Frequency Allocation Chart. Technical report, National Telecommunications and Information Administration, 2003.

[15] Bureau of Transportation Statistics. National transportation statistics. Technical report, U.S. Department of Transportation, 2008. [Online; accessed 03-February-2009].

[16] Jani Patokallio. Open flights: Airport and airline data, 2008. [Online; accessed 20-March-2009].

[17] Robert Poe. Row 44 Brings In-Flight Wifi to Alaska Airlines. *DailyWireless*, September 25, 2007.

[18] J.G. Proakis and M. Salehi. *Digital communications*. McGraw-Hill New York, 2001.

[19] E. Sakhaee, A. Jamalipour, and N. Kato. Aeronautical ad hoc networks. In *IEEE Wireless Communications and Networking Conference*, volume 1, 2006.

[20] C.W. Sayre. *Complete wireless design*. McGraw-Hill Professional, 2008.

[21] Brad Stone. Using the Web at 40,000 Feet. *New York Times*, March 17, 2008.

[22] Karl Swartz. Great circle mapper, 2008. [Online; accessed 15-December-2008].

[23] J.W. Tukey. Exploratory data analysis. 1977.

[24] C. Whelan. Operating in the North Atlantic MNPS Airspace. *Journal of Navigation*, 52(01):11–27, 1999.

[25] Todd Williamson, August 2008. Personal communications.