NORTHWESTERN UNIVERSITY

Characterizing Broadband Services in a Broader Context – Vantage Points, Measurements, and Experimentation

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Computer Science

By

Zachary Scott Bischof

EVANSTON, ILLINOIS

December 2016

ABSTRACT

Characterizing Broadband Services in a Broader Context –

Vantage Points, Measurements, and Experimentation

Zachary Scott Bischof

Broadband networks are one of the most economically significant and fastest growing sectors of the Internet. Recent studies have shown that providing broadband Internet access is instrumental for social and economic development. Several governments, as well as the UN, have gone so far as to label broadband access a basic human right, similar to education and water. Motivated by the increased importance of broadband access, recent efforts are shedding light on the availability of broadband services. However, these works tend to focus on measuring service capacity. As a result, we still lack an understanding of how factors such as a link's capacity, quality, dependability, or cost affect user behavior and network demand.

We believe that realizing the full benefits of broadband access requires an understanding of how these services are being used by subscribers. The thesis of this work is that broadband service characterization must take a user-centric perspective, understanding how different aspects of the service impact its users' experiences, and thus should be done

in a broader context. It should include an analysis of factors such as link quality, service dependability, and market factors (e.g. monthly income and cost of broadband access) and an understanding of how each affects user behavior.

To achieve this, we need to look beyond controlled experiments and regression analysis, two techniques commonly used in the field of networking. Controlled experiments, where subjects in the study are assigned randomly to “treated” and “untreated” groups for comparisons, are not feasible for studying the effect of complex treatments such as market and economic factors at scale. On the other hand, regression analysis is insufficient for causal inference. A key contribution of this work is the application of natural experiments and related experiment designs, techniques common in fields such as epidemiology and the social sciences, in the context of broadband services.

In this dissertation, we study broadband services in this broader context. We present the results of our empirical study on the relationship between service characteristics (capacity, latency, loss rates, and reliability), price, time and user demand. We find a strong correlation between capacity and user demand, but note a law of diminishing returns with lower increases in relative demand as service capacity increases. We also find that subscribing to unreliable broadband services tends to result in users generating less network traffic, even during periods of normal operation. These findings suggest that service dependability is becoming more important to subscribers as service capacities increase globally.

We include a characterization of broadband services in terms of bandwidth, latency, and loss. For bandwidth, we find that a number of providers struggle to consistently meet their advertised throughput rates and identify multiple instances where service throughput is correlated with the time of day. We also show that access latencies can vary widely by region, even within the same provider. In terms of service reliability, we find that broadband service providers in the US are able to provide at most two nines (99%) of availability.

Motivated by our findings on both the importance and current state of service reliability, we present an approach to improving service reliability using broadband multihoming and describe our prototype implementation. Our evaluation shows that in many cases, users could add up to two nines to service availability (from 99% to 99.99%) by multihoming with a neighbor's connection. Due to the fact that an individual subscriber may experience a wide range of performance, we then explore the possibility of adopting broadband service level agreements (SLAs). We argue that the use of broadband SLAs could help service providers to better differentiate their retail services from competitors and better inform both customers and policymakers of the broadband services offered in their communities. Using four years of data collected from residential gateways, we show that many ISPs could offer meaningful service level agreements immediately at little to no cost.

Committee

Fabián Bustamante (Chair), *Northwestern University*

Yan Chen, *Northwestern University*

Peter Dinda, *Northwestern University*

Renata Teixeira, *Inria*

Acknowledgements

First and foremost, I would like to thank my advisor, Fabián Bustamante, for his unwavering support throughout my graduate studies. Countless times during the past seven years of working with him, I would go into his office for a meeting feeling frustrated, stressed, or lost and each time I'd walk out of the meeting feeling more confident, re-energized, and ready to tackle the problem. This pattern epitomizes his abilities as an advisor to mentor students. He has guided my growth as a researcher in many ways: helping me to identify new research directions, become a critical reader and thinker, and develop my voice as a writer. Fabián, I will always appreciate and be amazed by the numerous times that you went above and beyond for your students, myself included.

I would also like to thank my thesis committee for their valuable feedback, especially over the last few years of my graduate studies. I very much enjoyed discussing new research ideas with Renata Teixeira. Yan Chen was one of my first Computer Science professors and has always been approachable and helpful throughout studies at Northwestern University. Peter Dinda, with both his attention to detail and his ability to consider "out-there" ideas, pushed me to be better prepared and always kept me on my toes.

A special thanks to all the friends and lab mates with whom I collaborated during time as a graduate student. Thank you, David Choffnes, for helping me get my feet wet during my first year. John Otto and Mario Sánchez were both tremendously important throughout my entire graduate career as role models, advisors, and friends. My work would not have been possible without all of their accomplishments – the metaphor of "standing on the shoulders of giants" seems most apt for describing how crucial their work was to my research. I would also like to thank John Rula for being a collaborator, roommate, close friend, and frequent partner in crime. Your friendship and support over our years of working together helped to keep me motivated, to step back from time to time and "unplug", and most importantly, to stay sane.

I am also thankful to all my collaborators and internship hosts over the years. I've always enjoyed brainstorming with Rade Stanojevic and coming up with ideas that, at first, seem impossible to tackle but would lead to very interesting research. Working with Nick Feamster has also been tremendously helpful. I always appreciate his ability to look at our work with fresh eyes and identify areas that would benefit from further exploration. I am also grateful for my collaboration with Kenjiro Cho, working with you was a great deal of fun.

Last but certainly not least, I would like to thank my family. Without their love and support, I would not have grown into the person I am today. In particular my brother, Jared Bischof, was a constant source of encouragement during my graduate studies. No matter how many times I complained about feeling lost or frustrated, you have always been there, believing in me more than I believed in myself. I cannot thank you enough for your patience and support.

List of abbreviations

- CDN: Content delivery network
- DNS: Domain Name System
- DSL: Digital subscriber line
- ISP: Internet service provider
- LAN: Local area network
- NDT: Network Diagnostic Tool
- P2P: Peer-to-peer
- PPP: Purchasing power parity
- RTT: Round-trip time
- SLA: Service level agreement
- TCP: Transmission control protocol
- UDP: User datagram protocol
- UPnP: Universal Plug and Play
- VoD: Video on demand
- VoIP: Voice over Internet Protocol
- WAN: Wide area network

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

As one of the most economically significant and fastest growing sectors of the Internet [**51**, **96**], broadband networks have attracted interest from researchers, network operators, and policy makers. Over the past decade, the number of broadband networks has increased rapidly. The latest "State of Broadband" reports that there are over 60 countries where fixed or mobile broadband penetration is above 25% and more than 70 countries where the majority of the population is online [**14**, **15**]. Access speeds have also been increasing. According to Akamai's "State of the Internet" report, over the last four years the top four countries in terms of average connection speed (South Korea, Hong Kong, Romania, and Japan) have nearly doubled their average capacity [**27**]. Recent studies have also shown that providing broadband Internet access is instrumental in social and economic development [**96**]. Several governments (including France, Finland and Spain) and the UN have gone so far as to label broadband access a basic human right, similar to education and water [**1**, **53**, **64**].

While a number of recent and ongoing efforts have shed light on the bandwidth and availability of broadband services [**10**, **11**, **17**, **31**, **55**, **81**, **90**, **93**, **94**], many of these studies focus on measuring bandwidth, looking at performance at an aggregated level. We believe that realizing the full benefits of broadband access requires a thorough characterization of the services that subscribers use. The thesis of this work is that broadband service characterization must take a user-centric perspective, understanding how different aspects of the service impact its users experiences, and thus should be done in a broader context, including an analysis of factors such as link quality, service dependability, and market factors (e.g. monthly income and cost of broadband access). We look at how these services are being used and how this use is impacted by the particulars of the market. How much bandwidth do people actually need? Do users in developing countries impose different demands on their services than users with similar services in developed countries? What impact does price have on usage? How does the quality and reliability of a connection impact user demand?

In order to answer these questions, we must study broadband services while considering factors such as a link's capacity, quality, or cost and measuring their impact on user behavior and network demand. We study the complex interplay between user behavior and features of the broadband service market. Using each user's demand on the access link as an indicator of users' needs, we study how user demand is affected by both the pricing structure of available services and performance characteristics of the service, such as capacity, latency, and loss.

One of the biggest challenges for a study such as this, is the nature of experiments one is able to conduct. Controlled experiments, where subjects in the study are assigned randomly to "treated" and "untreated" groups for comparisons are clearly not feasible for studying the features of interest at scale. A key contribution of our work is a methodology for combining broadband measurements and retail price datasets along with the application of natural experiments to address a problem otherwise impossible to tackle. We use natural experiments to examine the interaction between price, the quality of services available, and users' demands. We show that higher broadband prices increase demand when comparing users of similar capacities across markets.

Our study offers several insights on the interplay between user demand and broadband market features that should be of value to the research community, network operators and policy makers. For network operators, an understanding of users behavior can better inform network planning and operation. For policy makers, the work provides a firm statistical footing for discussions on broadband incentives. In our analysis, we find a law of diminishing returns in the relationship between broadband capacity and the average/peak demand users put on their broadband link, implying that adding extra capacity on an already wide broadband line leads to a minor increment in the user demand. Further increases in capacity beyond 10 Mbps had little impact on usage and increases beyond 25 Mbps did not have a statistically significant impact on usage.

While access to such speeds remains a challenge in large parts of the world [**50**, **99**], our findings suggest that in many developed countries, broadband capacities are high enough for most users. This increased capacity has led to the adoption of over-the-Internet services, such as VoD (e.g. Netflix, Hulu) and VoIP services (e.g. Vonage, Skype). According to a recent study, nearly 60% of U.S. broadband households subscribe to an OTT video service [**59**]. Furthermore, compared to the average household in the US, those with a Netflix or Hulu subscription are about 2.8 times more likely to not have a cable subscription [**85**]. Many small businesses are also switching to VoIP as an economic alternative to landline phones [**45**, **92**].

This increased reliance on over-the-Internet services has highlighted the need for clear standards and a better understanding of the role that service reliability plays in shaping the behavior of broadband users [8, 40, 58]. A recent survey on consumer experience by the UK Office of Communication (Ofcom) ranks reliability first, above speed of connection, as the main reason for customer complaints. Despite its growing importance, both the reliability of broadband services as well as approaches to improve on it have received scant attention from the research community. Our belief is that as broadband capacities continue to improve, reliability will become the key differentiating factor between alternative services, a common pattern as a technology's market matures.

To this end, this dissertation includes the first comprehensive study of broadband service reliability. Our empirical study of access-ISP outages and user demand corroborates the Ofcom user survey on the increased importance of service reliability. We find that users that frequently experienced reliability issues tended to have lower network demand, even during periods of normal operation, suggesting that the service reliability is becoming paramount as service capacities increase globally.

Motivated by our findings on the importance of broadband service reliability, we present an approach to improving service reliability using broadband multihoming and describe our prototype implementation. Our evaluation shows that in many cases, users could add up to two nines to service availability (from 99% to 99.99%) by multihoming.

We also study reliability in terms of the consistency of performance. In today's broadband markets in North America and overseas, services are described in terms of their maximum potential throughput rate, often advertised as having speeds "up to X Mbps". Though such promises are often met, they are fairly limited in scope and, unfortunately, there is no basis for an appeal if a customer were to receive compromised quality of service. This 'best effort' model for broadband services was sufficient in the early days of the Internet. We argue that as broadband customers and the devices they use become more dependent on Internet connectivity, we will see an increased demand for more encompassing Service Level Agreements (SLA).

We study the design space of broadband SLAs and explore some of the trade-offs between the level of strictness of SLAs and the cost of delivering them. We argue that certain SLAs could be offered almost immediately with minimal impact on retail prices. The introduction of SLAs could enable broadband operators to personalize the service offerings down to the individual customer and improve their efficiency and overall user satisfaction. In addition, broadband SLAs could facilitate transparent competition in the broadband market, ultimately benefiting both consumers and service providers.

1.1. Summary of major findings and contributions

The following paragraphs summarize the major findings and contributions of this work.

Leveraging natural experiment study designs. Apply natural experiments and case-control studies in the field of networking. Using techniques such as propensity score matching, in particular, caliper matching, to compare similar users.

Longitudinal analysis of user demand. Looking at the longitudinal data, we find, somewhat surprisingly, that despite the fourfold increase in global IP traffic over the past five years [**25**], subscribers' demand in the same bandwidth capacity class remained constant, indicating that users "jump" to a higher service when their demand grows, rather than fill their (typically underutilized) pipe. We study in depth the service upgrade dynamics and report our findings.

Empirical study of the relationship between reliability and user behavior. We demonstrate that that poor reliability can affect user traffic demand well beyond periods of unavailability. For instance, we find that frequent periods of high packet loss (above 1%) can result in a decrease in traffic volume for 58% of users *even during periods of no packet loss.*

Characterization of broadband services. We include a characterization of broadband services in terms of bandwidth, latency, and loss. In terms of bandwidth, we find that a number of providers struggle to consistently meet their advertised throughput rates and identify cases where throughput varies by time of day. We show that access latencies can vary widely by region, even within the same provider. We present an approach to characterize broadband service reliability. We apply this approach to data collected from 7,998 residential gateways over four years (beginning in 2011) as part of the US FCC MBA deployment [**37**]. We show, among other findings, that current broadband services deliver an average availability of at most two nines (99%), with an average annual downtime of 17.8 hours.

Quantifying the potential benefits of broadband multihoming. Using the FCC MBA dataset and measurements collected by over 6,000 end-host vantage points in 75 countries [**71**], we show that multihoming the access link at the home gateway with two different providers adds two nines of service availability, matching the minimum four nines (99.99%) required by the FCC for the public switched telephone network (PSTN) [**57**].

AlwaysOn: a gateway for broadband multihoming. We present the design, implementation, and evaluation of *AlwaysOn*, a readily deployable system for multihoming broadband connections.

Design and evaluation of broadband SLAs. We analyze different QoS metrics for use in SLA and define a set of broadband SLAs. We find that, across all ISPs and access technologies, bandwidth is the most consistent of the three studied performance metrics (bandwidth, latency and loss rate). We evaluate the relationship between access technology, the SLA structure and the cost of having SLAs. We show that many of the studied ISPs could offer moderate SLAs with minimal impact on their existing business, but that SLAs with stringent constraints are much harder to deliver across the whole user-base.

Creating personalized broadband SLAs. We examine if SLAs could be tailored for each end-user individually and show that ISPs (or third parties) could accurately (with accuracy comparable to that in the car or credit insurance industry) infer the risk of offering SLA to individual customers and price the SLA service accordingly.

1.2. Organization

The rest of the paper is organized as follows. In Chapter 2, we explain our analysis methodology and describe the datasets we use throughout our analysis. The following chapter looks at how user demand on the network is affected by capacity, connection quality, service dependability, time and market factors (Chapter 3). We present a characterization of broadband services in terms of service capacity, latency, loss, and reliability (Chapter 4). Motivated by our findings on the importance of service reliability, we present and evaluate an approach to multihoming broadband services to improve service availability in Chapter 5. We then discuss the feasibility of applying SLAs to broadband services in Chapter 6. Finally, Chapter 7 provides an overview of the contributions of this work and discusses open research questions and future directions.

CHAPTER 2

Methodology

In the following paragraphs we describe the datasets used throughout our analysis, including a summary of key characteristics for the broadband connections they capture. We close the chapter with a brief discussion of the goals and methodology of our study.

2.1. Datasets

Our study builds on three types of datasets: (*i*) measurements from residential gateways in the US, and (*ii*) detailed end-host collected data on broadband connections from around the world, and (*iii*) a compilation of retail broadband connectivity plans made available by Google [**78**]. We describe each of these datasets in the following sections.

2.1.1. Residential gateway data

In 2010, the FCC, in collaboration with SamKnows, began to distribute residential gateways to broadband users around the US as part of the "Measuring Broadband America" effort [**37**]. Users that participate in this study were either selected to participate by their ISP or signed up through SamKnows' website. Measurement locations for these studies were chosen from a set of volunteers, and stratified into groups. For instance, the original 10,000 participants in the ongoing FCC study, shown in Figure 2.1, were selected from a pool of over 145,000 volunteers [**36**] based on features that, a priori, were thought to be

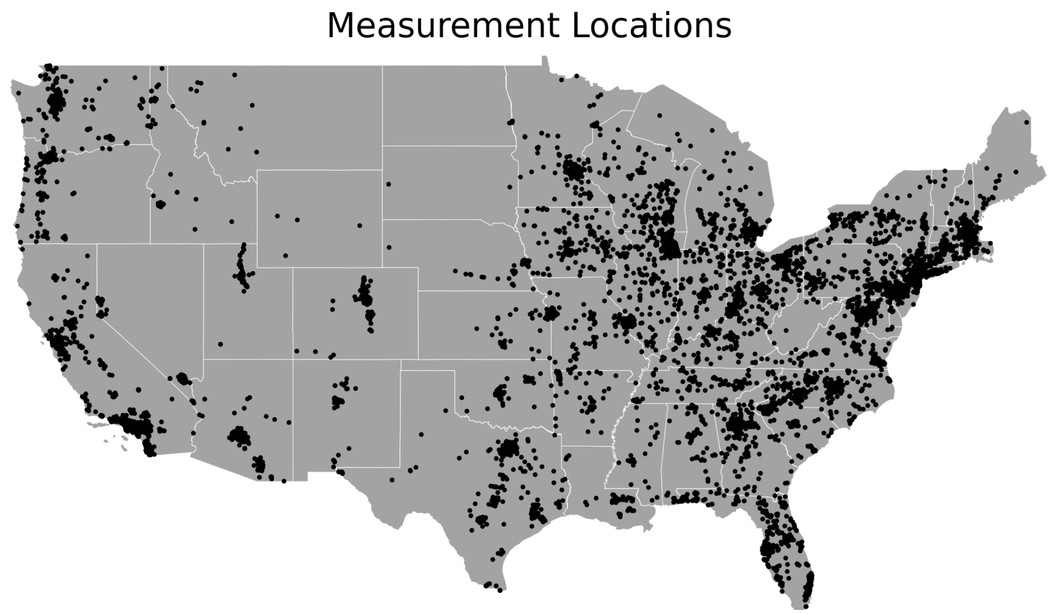


Figure 2.1. Map of all measurement locations used in FCC Broadband America study.

relevant to service performance such as geographic region, access technology and subscription level. Boxes were distributed to subscribers across 15 major ISPs, with the number of participants selected for each service provider proportional to its market share. Since the initial deployment to US broadband subscribers in late 2010, additional boxes have been distributed to new participants to replace boxes that were deactivated, ensuring that the number of measurement points remains stable despite the natural churn.

Since February 2011, the FCC has publicly released datasets using measurements executed by custom gateways in selected users' homes. The collected data includes a rich set of metrics, such as measurements of link capacity, latency, packet loss, and loading times of popular websites as well as hourly recordings of the number of bytes sent and received over the WAN link.[1] This data has been primarily used to create periodic reports on the state of broadband services in the US as part of the MBA initiative [**37**].

Throughout this work, we employ the full four years of measurements made available in order to quantify network performance in terms of latency, packet loss, and download/upload throughput and to measure user traffic. For this, we used four different measurement tables (out of eleven) from the dataset for our analysis: (1) UDP pings, (2) HTTP GETs, which measure download throughput, (3) HTTP POSTs, which measure upload throughput and (4) traffic byte counters.

The UDP pings run continuously, measuring round-trip time to two or three measurement servers. These servers are hosted by either M-Lab or within the user's provider's network. Over the course of each hour, the gateway will send up to 600 probes to each measurement server at regular intervals, less if the link is under heavy use for part of the hour. Each gateway reports hourly statistical summaries of the latency measurements (mean, min, max, and standard deviation) as well as the number of missing responses. We use the average latency to the nearest (in terms of latency) server to summarize the latency during that hour. We also use the number of missing responses to calculate the packet loss rate over the course of each hour.[2]

As mentioned above, the HTTP GET and POST tables record the measured download and upload throughput rate, respectively. Similar to the latency measurements, throughput measurements are typically done to two different target servers. However, throughput

[1] A full description of all the tests performed and data collected is available in the FCC's measuring broadband technical appendix [**36**].

[2] The FCC's hourly aggregation of the reported data prevents us from conducting an analysis of packet loss rates at a finer granularity.

measurements are run once every other hour, alternating between measuring upload and download throughput rates.

The traffic byte counters record the total number of bytes sent and received across the WAN interface over each previous hour. They also record counters for the amount of traffic due to the gateway's active measurements, which we subtract from the total volume of traffic.

We combined these performance measurements with user metadata, which includes the user's provider, service tier (i.e., subscription speed), service technology (e.g., cable or DSL), and geographic region. This allows us to group measurements by ISP, compare the achieved throughput as a fraction of the subscription capacity and differentiate between subscribers of the same ISP with different access technologies (e.g., Verizon customers with DSL or fiber).

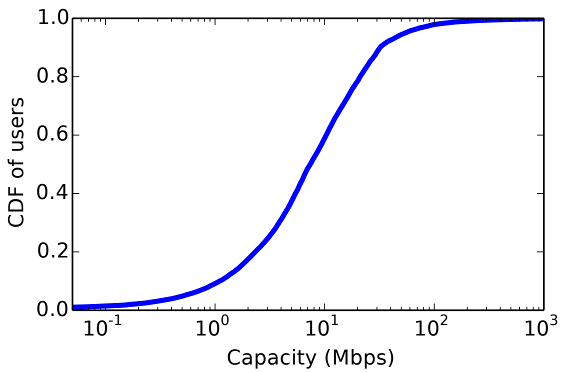
2.1.2. End host data

We used two different software tools to collect data from end hosts, Dasu [**84**] and Namehelp [**71**]. Both are described in the following paragraphs.

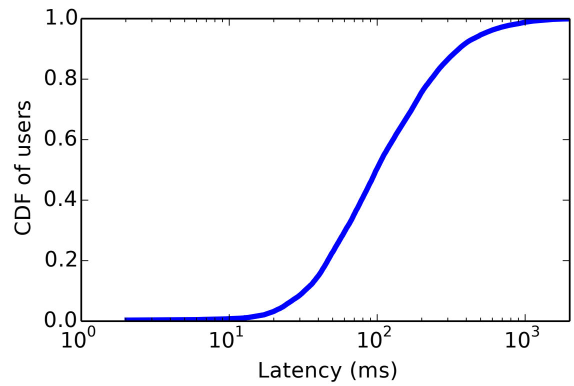
2.1.2.1. Dasu. Dasu [**84**], a previously released network experimentation and broadband measurement client, is available as both an extension to BitTorrent and as a standalone client. As an incentive for adoption, Dasu informs users of their ISP's performance, providing detailed information on their home network configuration, the volume of network traffic sent and received by the localhost, the volume of detected cross traffic in the home network, and the results of performance measurements on their ISP (e.g. a comparison of their ISP's web browsing and DNS performance). Dasu records network usage data from the localhost and home network to account for cross traffic during characterization or the execution of network experiments.

Since its release, Dasu has been installed by over 100,000 users in over 160 countries, with the majority of clients using the BitTorrent extension. From this dataset, we select users that either have UPnP enabled on their home gateway device or those that were directly connected to their modem (thus their machine is the only device generating traffic). UPnP-enabled gateways provide byte counters that we use to measure activity on the link, taking into account issues with UPnP counters raised in other works [**29**,**83**]. For users directly connected to their modem, we use byte counters available from `netstat` to monitor network usage (available by default on most popular operating systems). Traffic byte counters are collected at approximately 30 second intervals with some variations due to scheduling.

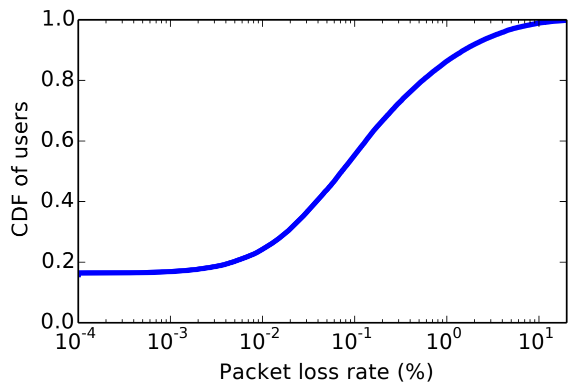
As it is the case with all observational studies, there is a concern about potential biases in our datasets, coming either from P2P or SamKnow's users (e.g., uniquely demanding users, early-adopters or "geek-effect") [**17**, **55**, **70**]. We account for some of these issues throughout our analysis by, for instance, focusing on measurements gathered when users are not actively downloading/uploading content on BitTorrent, restricting our users to those directly connected to a modem or wirelessly connected to a UPnP-enabled one, using neighbor matching with a caliper to ensure close matches. On the potential biases with our P2P users' data, we show in Sec. 3.1.1 that the average demand of Dasu users in the US – when not actively using BitTorrent – is comparable to that of participants in the FCC's study.



(a) Download capacity



(b) Latency



(c) Packet loss

Figure 2.2. CDFs of the maximum download capacities, average latency to nearest available measurement server, and average packet loss rates measured for every network connection in our Dasu dataset.

We now briefly describe the diversity of broadband connections in our global dataset, presenting distributions of their measured capacity, latency and packet loss. All the Dasu data were collected by running M-Lab's Network Diagnostic Tool (NDT) [**63**] within Dasu. NDT reports the upload and download capacity of a connection, as well as its end-to-end latency and packet loss rates.

Capacity. Figure 2.2a shows a CDF of the maximum download capacities, in Mbps, measured over each user's connection in our dataset. Our distribution has a median user download capacity of 7.4 Mbps and interquartile range of 14.3 Mbps (from 3.1 Mbps to 17.4 Mbps). Approximately 10% of users have download capacities below 1 Mbps, while the top 10% of users have capacities above 30 Mbps.

Latency. For latency, we measured the average latency to the closest NDT measurement server. Since measurement servers are hosted in a diverse set of networks of content providers (e.g. Google) and content distribution networks (e.g. Level 3), we believe such measurements provides a reasonable estimate of the latency to popular content. Figure 2.2b shows the distribution of measured latencies. We find that a "typical" user in our dataset has an average RTT of about 100 ms to the nearest NDT servers. The top 5% of users had an average latency above 500 ms. Based on the organization names that we found via `whois` lookups, the majority of connections with very high latencies appeared to be connecting over wireless modems or satellite providers.

Packet loss. Figure 2.2c shows the distribution of average packet loss rates reported by NDT tests. While the loss rate is relatively low for most users (less than 0.1%), approximately 14% of users saw an average loss rate above 1%. For the top 1% of users, average loss rates were above 10%. As was the case with high latency connections, the organization names of networks with very high packet loss rates indicated they were satellite or wireless (e.g. WiMAX, cellular) services.

2.1.2.2. Namehelp. For our analysis of the potential of broadband multihoming for improving reliability, we also collected measurements from approximately 6,000 endhosts running Northwestern University's Namehelp service [**71**]. Namehelp is a client DNS service that improves web performance by obtaining more accurate redirections to nearby content delivery network replica servers. For this work, we added a network experimentation platform to the Namehelp client, allowing us to perform diagnostic tests.

Despite it's smaller user base, one advantage of Namehelp is that as a background system service, so long as a device is operational, Namehelp is running and able to perform network measurements. The Dasu BitTorrent client, on the other hand, is only only able to run network measurements while the user is running Vuze. This is a preferred scenario for some of our experiments in our study, including those that test network availability.

Further details about the measurements collected from Namehelp clients are described in Chapter 5.

2.1.3. Connectivity plans

Our last dataset is a compilation of international retail broadband connectivity plans, made available by Google on their "Policy by the Numbers" blog [**78**]. This data was compiled by Communications Chambers, a consultant group, by visiting the websites of broadband service providers around the world. The dataset covers 1,523 service plans across 99 countries. It includes information on the upload and download speeds of each plan, the monthly traffic limits, and monthly cost in the local currency. We selected this dataset over those provided by the FCC, OECD, or ITU given the breadth of countries included and the depth of plans listed. The FCC and OECD datasets focus on the US and members of the OECD while the ITU dataset only includes a single service plan for each country. In a few cases, we expanded this dataset by manually visiting the websites of ISPs in countries where we had users but no broadband price data.

To directly compare the price of broadband plans across different economies, we convert the monthly cost to US dollars. We account for differences in relative purchasing power in each country by using the purchasing power parity (PPP) to market exchange ratio. In most cases, this is included in the broadband service survey provided by Google. When that is not the case, we use publicly available data from the International Monetary Fund's website[3]. All monetary figures throughout this work are normalized by purchasing power parity, including the gross domestic product (GDP) per capita data provided by the International Monetary Fund that we use later in our case study.

2.2. Experimental methodology

One goal of our study is to provide insight into the impact of broadband service market characteristics on network usage. Specifically, we study the impact of the following market features: connection capacity, the price of broadband access, the cost of increasing capacity, and connection quality. While there are many other variables that can affect user behavior, this set covers the key characteristics of broadband service markets. Given the rapid pace of development in broadband and the reported growth in network traffic, we also conduct a longitudinal analysis of user demands on broadband services.

Beyond gathering a sufficiently large and diverse perspective of broadband connections, a key challenge for a macroscopic study such as ours is the nature of experiments one is able to conduct. Classical controlled experiments – where subjects in the study are randomly assigned to "treated" and "untreated" groups for comparison – are clearly not feasible at a global scale. It is also unlikely that the features we explore are independent, e.g., one would assume that price or service diversity can impact capacity and service quality.

[3]International Monetary Fund. `http://www.imf.org/`

This has been a long, well understood problem in a range of fields, from epidemiology to sociology and economics. We address this challenge, as many studies do in those domains, by leveraging *natural experiments* in our analysis, a class of experimental design common in epidemiology, the social sciences, and economics [**33**].

By using natural experiments [**33**] and related study designs, we remedy the fact that we cannot control the application of a treatment, Matching users in our treated group with similar users in the untreated group we simulate random or as-good-as-random assignment, manually ensuring that differences are evenly distributed between the two groups. This allows us to infer whether or not the relationship observed are likely to be causal. For example, to test if bandwidth capacity affects user demand, we pair users that are similar in terms of connection quality and broadband market. We then check if the user with higher capacity generates more traffic. If so, our hypothesis holds true for that pair. After testing this for each pair of similar users, we calculate the percentage of times that our hypothesis is correct.

If neither of the two variables under study – in this example, capacity and demand – have an impact on the other, then their interaction would be random. In our example this would mean, for instance, that lower capacity will result in lower (or higher) demand about 50% of the time. Significant deviations from this would suggest that a causal relationship is likely to exist between the two.

We use the one-tailed binomial test to measure the statistical significance of deviations from the expected distribution. As is common in many study designs, we consider a p-value that is less than 0.05 to be a strong presumption against the null hypothesis (H_0). One potential issue with our application of the binomial test in this context is the known problem that given a large enough dataset, the test will consider even minor deviations to be significant. That is, with a large enough sample of throws, an unbiased coin could fail to pass a χ^2 test for fitting the predicted binomial distribution [**75**]. To account for this issue, we only consider deviations larger than 2% to be practically important. In other words the hypothesis holds at least 52% of the time.

We use network demand as a measurable metric that may reflect user experience. Recent work [**6**, **9**] suggest that this metric is a reasonable proxy for user experience. A significant change in network usage (e.g., bytes transferred or received) can be interpreted as a response to a change in the user's experience.

CHAPTER 3

Putting broadband in a broader context

The main goal of our study is to provide insight into the impact of broadband service market characteristics on network usage. Specifically, we study the impact of the following market features: connection capacity, the price of broadband access, the cost of increasing capacity, connection quality, and service reliability. While there are many other variables that can affect user behavior, this set covers the key characteristics of broadband service markets. Given the rapid pace of development in broadband and the reported growth in network traffic, we also conduct a longitudinal analysis of user demands on broadband services.

3.1. Service capacity and user behavior

The interplay between broadband service characteristics and user demand is complex [**99**]. For instance, while subscribers cannot directly affect the cost of their service, they have some freedom in choosing what package (capacity) they purchase and how much traffic they generate. On the other hand, although they come with needs and budgets when choosing a broadband plan, once acquired, their usage patterns are shaped by their selection. In addition, there is the potential impact of seemingly irrational and biased choices by subscribers [**42, 91**] that complicates any attempt at understanding and analytical modeling of the drivers of users' choices and demand. While we (or even most customers [47]) may not know the advertised service of a broadband connection, our study focuses on the impact of the actual maximum capacity provided to the user.

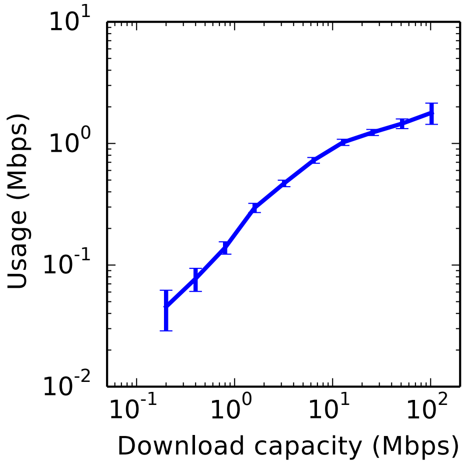
In this section, we begin to empirically explore the complex interactions between broadband service market features and user behavior by first studying the effects of capacity on user demand. When appropriate, we compare data collected from end hosts (via Dasu) and residential gateways (FCC/SamKnows).

3.1.1. Capacity vs. usage

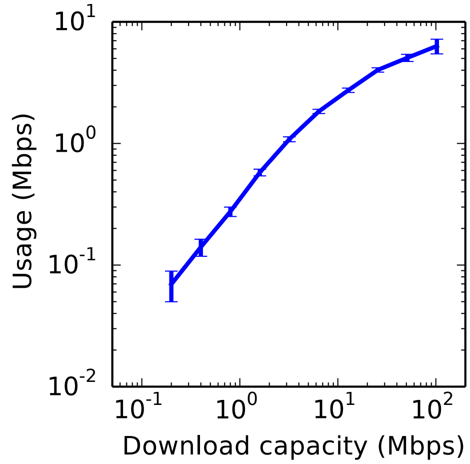
We first explore the relationship between access link capacity and the demand users generate on the access network. To describe user demand, we rely on two metrics of usage: the average and peak volume of traffic generated. We define peak as the 95th-percentile value of the time series (sampled every 30-secs) of downlink demand for each user.

Figure 3.1 presents both the mean and peak demand, for different classes of users based on their measured downlink capacity. Given the range of services across the different markets we analyze, we split services into ten classes where every user in class (k) has a download capacity in the range of $(100Kbps * 2^{k-1}, 100Kbps * 2^{k}]$. We analyze usage both throughout the entire measurement period (Fig. 3.1a and 3.1b) and during periods when users are not actively uploading or downloading content on BitTorrent (Fig. 3.1c and 3.1d).

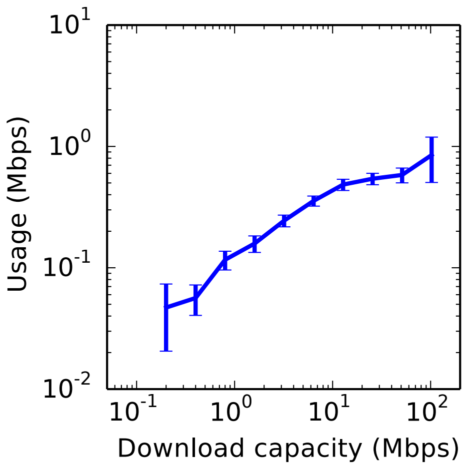
We also contrast Dasu' end-host collected data with that of users in the FCC study (gateway collected data), looking both at average and peak network usage. Figure 3.2 shows the mean and peak (95th percentile) demand of users, grouped by capacity, in the



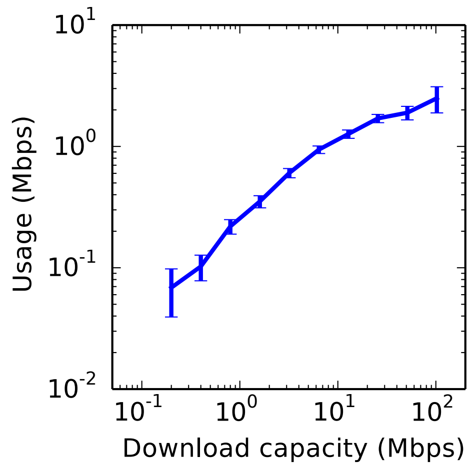
(a) Mean
w/ BT
($r = 0.870$)



(b) 95th %ile
w/ BT
($r = 0.913$)

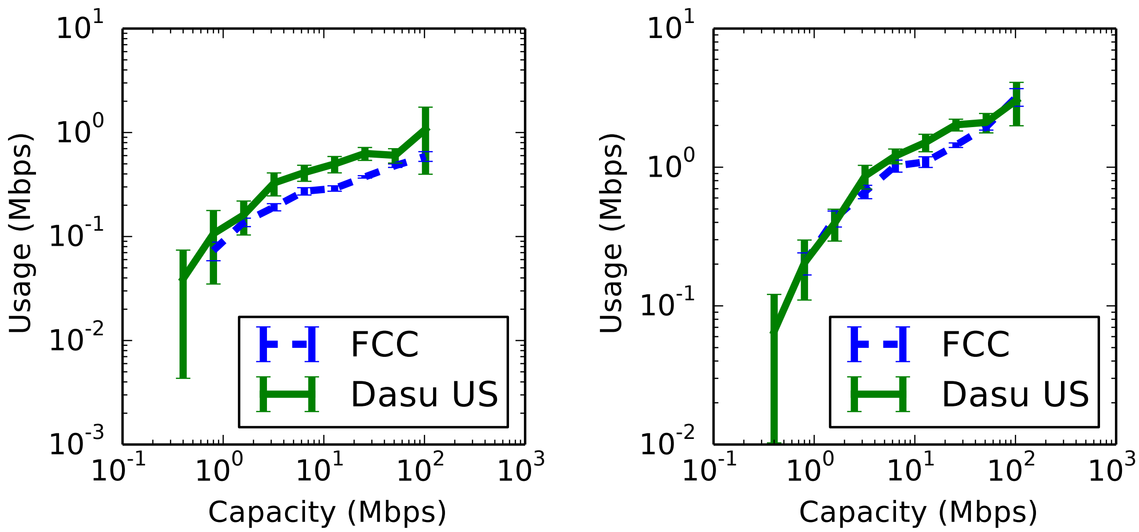


(c) Mean
no BT
($r = 0.885$)



(d) 95th %ile
no BT
($r = 0.890$)

Figure 3.1. Volume of download traffic generated by users versus their download link capacity. Users are grouped by their download capacity and each bin is averaged. The error bars represent the 95% confidence interval of the mean. In each case, usage is strongly correlated with link capacity.



(a) Mean
($r = 0.915$)

(b) 95th %ile
($r = 0.905$)

Figure 3.2. Mean and peak (95th percentile) download traffic generated for FCC gateway users and Dasu users within the US when not using BitTorrent. The error bars represent the 95% confidence interval of the mean.

FCC and US-based Dasu dataset (when not using BitTorrent). Although the average usage is slightly higher for Dasu users, the peak usage is nearly identical for both groups. The difference in average demand is likely due to the fact that the FCC data is collected evenly throughout the 24-hour period, while Dasu usage (and thus our data) is partially biased towards peak usage hours.

We find that usage grows with capacity, as the plots of Fig. 3.1 and 3.2 clearly show. This is *despite* the fact that users rarely utilize their link (even at the 95th percentile, average utilization ranges between 10 and 48%). For both mean and 95th percentile usage, with and without BitTorrent traffic, we find that usage is strongly correlated with the group's link capacity ($r \geq 0.87$ for each).

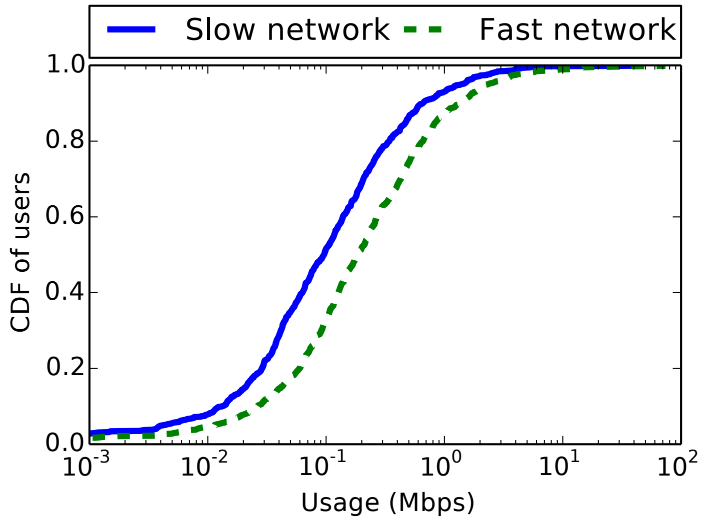
Figure 3.1 also show that as capacity increases, usage begins to level off (related to the findings in Sundaresan et al. [**94**]). This would suggest that the relationship follows a law of diminishing returns: the relative increase in demand is greater for lower capacity connections than for higher capacity connections.

3.1.2. Inferring causality

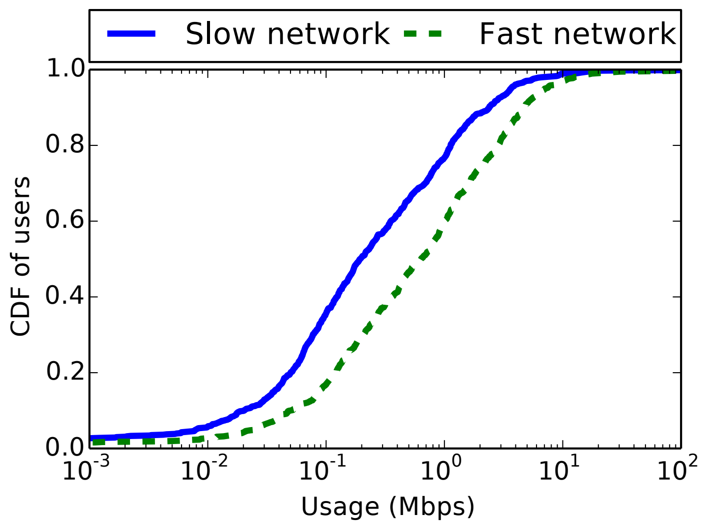
While the access capacity and the demand users generate are strongly correlated, inferring the causality between these variables is significantly more challenging. For instance, demand might drive capacity (i.e. users sign up for faster services because they have access to applications with higher bandwidth needs, such as HD video streaming) or be driven by it, with users changing their behavior when given a higher capacity and generating a higher demand. Additionally, there may be other factors that affect user demand such as the quality of the connection or the price of access.

To explore a causal relationship between access capacity and demand, we first design a natural experiment to see if the behavior of individual users changes when switching between networks of different capacities. This let us test the idea that when given a higher capacity link, users will increase their demand on the network. We then compare the demand of users that are similar in terms of price of broadband access, cost to upgrade, and link quality but differ in terms of service capacity.

User upgrades. To determine if their relationship between capacity and demand is causal, we need to account for differences in usage patterns between different users. We do this by looking at how individual users change their network demand when switching to



(a) Mean



(b) 95th %ile

Figure 3.3. CDFs of the mean and peak download link usage for individual users on "slow" and "fast" networks when not using BitTorrent.

faster services, allowing us to determine if the relationship between capacity and demand is likely causal.

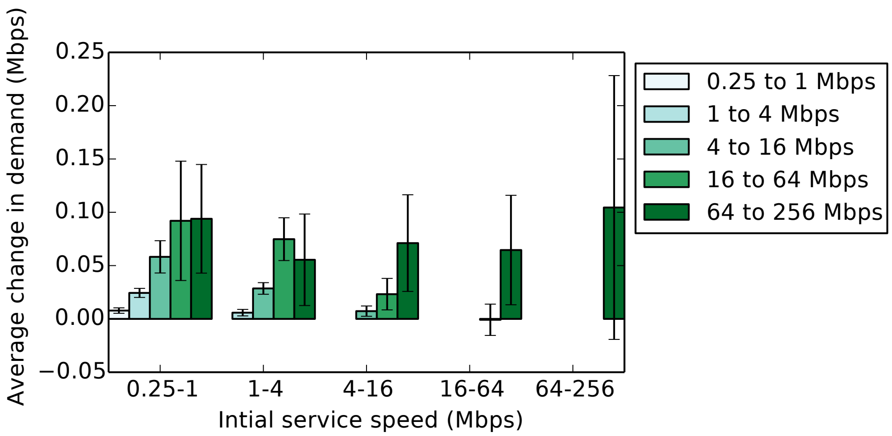
Metric	% H holds	p-value
Average usage	66.8%	1.94×10^{-25}
Peak usage	70.3%	1.13×10^{-36}

Table 3.1. Percentage of the time that an individual user's average and peak demand will increase when moving to a network with a higher capacity. In both cases, the control group is their behavior on the slower network and the treatment is their behavior on the faster network.

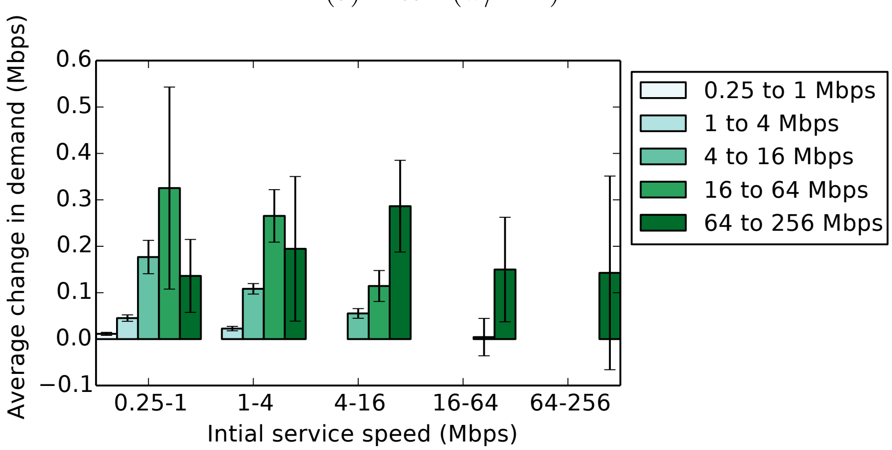
Figure 3.3 presents the CDFs of mean and peak download link usage for users switching between a "slow" and "fast" network. Both average and peak volume of traffic are when the client is not active on BitTorrent. Each network is identified by a tuple *(ISP name, network prefix, geolocated city)*. For both average and peak demand, we see that usage tends to be considerably higher on the faster network. For example, at the median, average usage doubles from 95 kbps to 189 kbps and peak usage (95th percentile) more than triples, from 192 kbps to 634 kbps.

To validate this assertion we use a natural experiment. Our hypothesis (H) is that when a user moves from a slower to a faster service, demand will increase. As such, our null hypothesis (H_0) is that demand will not be affected by a change in capacity. We test this assertion for both the mean and peak demand and present our results in Table 3.1. As the table shows, our original hypothesis (H), is true 66.8% of the time when comparing average demand and 70.3% of the time for peak demand. For both metrics, we find very small p-values, leading us to reject the null hypothesis (H_0) that capacity does not affect the demand of individual users.

For these analysis we limit our data to that collected while the users were not generating BitTorrent traffic. Including BitTorrent traffic, we find an even higher increase in usage, and so is the percentage of the time that our hypothesis holds true. This is likely



(a) Mean (w/ BT)

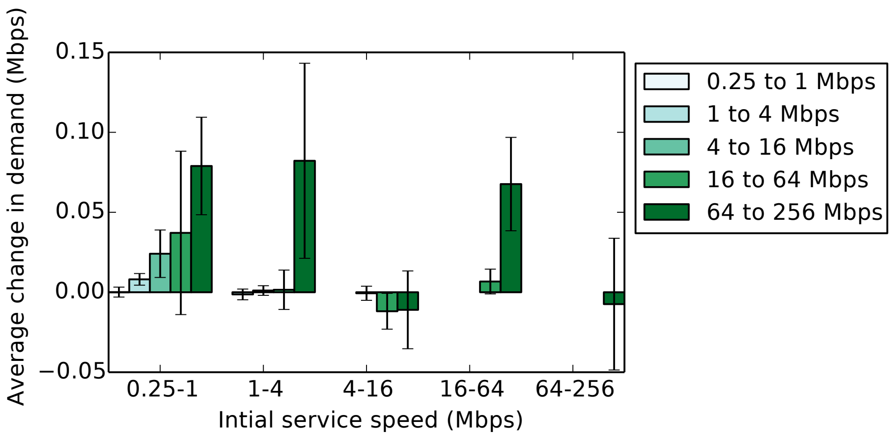


(b) 95th %ile (w/ BT)

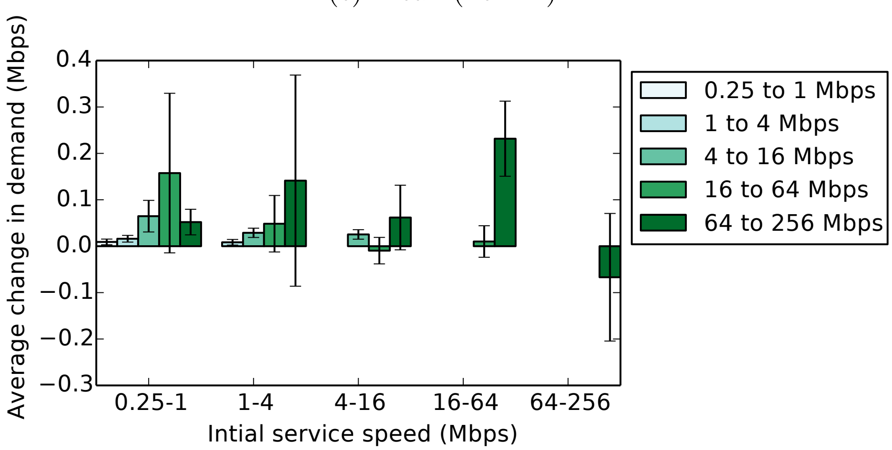
Figure 3.4. Change in volume of traffic generated when switching to a faster connection. The x-axis corresponds to the initial service speed while each bar represents the average change for users switching to a faster service within that group. The error bars represent the 95% confidence interval. (Continued on next page)

due to the fact that users are more likely to saturate their link for extended periods of time when using BitTorrent [**24**].

Impact of switching services. To further understand the interaction between capacity and demand, we explore the impact of service upgrades on demand, for different



(c) Mean (no BT)



(d) 95th %ile (no BT)

Figure 3.4. (cont.) Change in volume of traffic generated when switching to a faster connection. The x-axis corresponds to the initial service speed while each bar represents the average change for users switching to a faster service within that group. The error bars represent the 95% confidence interval.

initial capacities of connections. Figure 3.4 shows the average change in demand, grouping upgrades by the "before" and "after" download capacities. The labels on the x-axis represent the capacity range of the initial service and each bar is the average change in demand when switching to a faster service in the respective capacity tier.

Dasu data			
Control Group (in Mbps)	Treatment Group (in Mbps)	% H holds	p-value
$(0.1, 0.2]$	$(0.2, 0.4]$	75.2%	5.81×10^{-11}
$(0.2, 0.4]$	$(0.4, 0.8]$	63.4%	2.21×10^{-7}
$(0.4, 0.8]$	$(0.8, 1.6]$	59.9%	8.01×10^{-8}
$(0.8, 1.6]$	$(1.6, 3.2]$	59.3%	1.11×10^{-8}
$(1.6, 3.2]$	$(3.2, 6.4]$	53.3%	0.0166
$(3.2, 6.4]$	$(6.4, 12.8]$	57.5%	0.00707
$(6.4, 12.8]$	$(12.8, 25.6]$	56.8%*	0.0583
$(12.8, 25.6]$	$(25.6, 51.2]$	52.9%*	0.310
$(25.6, 51.2]$	$(51.2, 102.4]$	51.0%*	0.462
FCC data			
Control Group (in Mbps)	Treatment Group (in Mbps)	% H holds	p-value
$(0.4, 0.8]$	$(0.8, 1.6]$	66.4%	0.000223
$(0.8, 1.6]$	$(1.6, 3.2]$	58.1%	4.70×10^{-05}
$(1.6, 3.2]$	$(3.2, 6.4]$	56.2%	0.000487
$(3.2, 6.4$	$(6.4, 12.8]$	55.1%	0.00236
$(6.4, 12.8]$	$(12.8, 25.6]$	58.5%	2.54×10^{-7}
$(12.8, 25.6]$	$(25.6, 51.2]$	61.2%	6.76×10^{-17}
$(25.6, 51.2]$	$(51.2, 102.4]$	64.7%	0.00161

Table 3.2. Percentage of the time that increased capacity will increase demand when comparing similar users and each experiment's corresponding p-value. An asterisk denotes that a result was not statistically significant.

As the figure shows, for each metric, demand clearly increases when upgrading from slower services, particularly when looking at peak (95th percentile) usage. Increases in demand are less consistent when switching between already fast services, particularly above 16 Mbps, where there is a large variance on demand growth with capacity. In some cases, the large range in the 95% confidence interval shows that the upgrade likely had no significant impact on usage. These findings suggest that while capacities do drive demand, this is only true up to a certain point.

All users. We expand our comparison to all users in the datasets and use a matching study design to test the impact of increased capacity. As before, we place users into one of k bins, where $k = (100Kbps * 2^{k-1}, 100Kbps * 2^{k}]$. We then compare the usage of users in bins k and $k+1$. Our hypothesis (H) is that users in the "treated" group, $k+1$, will have a higher demand on the network due to their increased capacity. Our null hypothesis (H_0) is that the relationship is random and increased capacity will not result in higher demand.

To compare users from each group, however, we must ensure that each pair of users is similar in terms of connection quality (packet loss and latency), price of broadband access, and cost to upgrade capacity. For this and the remaining studies, we use nearest neighbor matching to pair similar users in "control" and "treatment" groups. We use a caliper to ensure that dissimilar users are not matched, requiring that users be within 25% of each other for each confounding factor. This means, for instance, that users with latencies of 50 and 62 ms and in regions where broadband Internet access costs $25 and $30 (USD) per month are considered to be sufficiently similar in terms of latency and cost of broadband access. Note the trade-off here, a tighter caliper will yield a potentially more accurate comparison, but will also reduced the number of comparisons we can perform.

Table 3.2 shows the experiment's results separated by the datasets used. For the Dasu data, increased capacity has the widest impact when comparing slower service groups. The increase in demand is statistically significant while comparing groups of users with capacities less or equal to 6.4 Mbps (though the achieved p-value when comparing groups $(6.4, 12.8]$ and $(12.8, 25.6]$ is very close to 0.05). When comparing users in bins above 12.8 Mbps, the difference tends to become random and our hypothesis holds about 50%

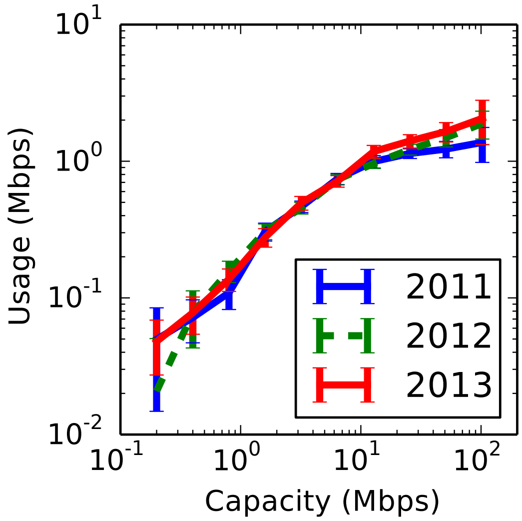
the time. These results suggest that increasing capacity beyond ≈10 Mbps is less likely to have a significant impact on peak user demand.

For the FCC data, increased capacity tends to result in increased demand across all bins. We believe that this is largely due to the fact that the FCC vantage point set is comprised solely of users in the US, where higher capacity broadband services are available, but at a moderately higher price (this does not apply in many of the countries in our study). We also observed a similar trend of increased usage when studying Dasu users in the US, as we will show in Sec. 3.3.

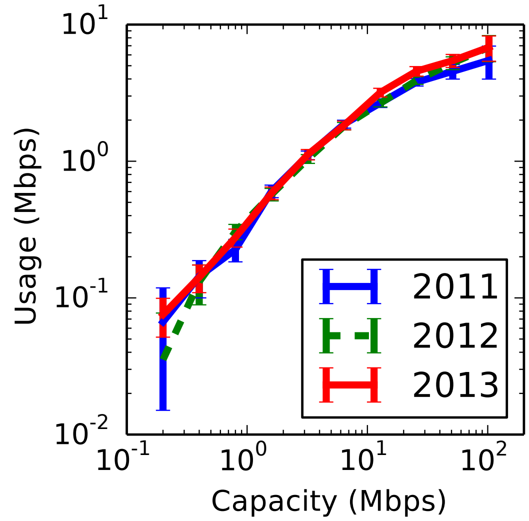
3.2. Longitudinal trends in usage

The last few years have witnessed a rapid growth on the capacity, coverage and affordability of broadband networks [**14**]. Concurrently, the volume of digital content and total IP traffic continue to grow at rapid pace. A recent Cisco report states that the total IP traffic has increased 18-fold since 2.4 exabytes in 2005 [**25**]. Meanwhile, the size of the "digital universe", the total amount of data created and replicated reported to be 2.8 zettabytes in 2012, doubles in size about every two years [**48**]. In this section, we look for changes in demand over time to see if these changes are reflected, and in what manner, in the network demand of broadband users.

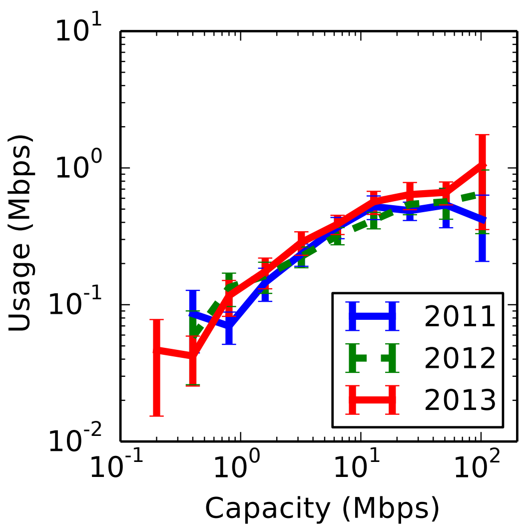
To this end, we carry a longitudinal analysis of broadband connections in our dataset. We compare changing trends in usage relative to capacity, both average and peak, between 2011 and 2013. Figure 3.5 shows average and demand over this period, with and excluding BitTorrent traffic.



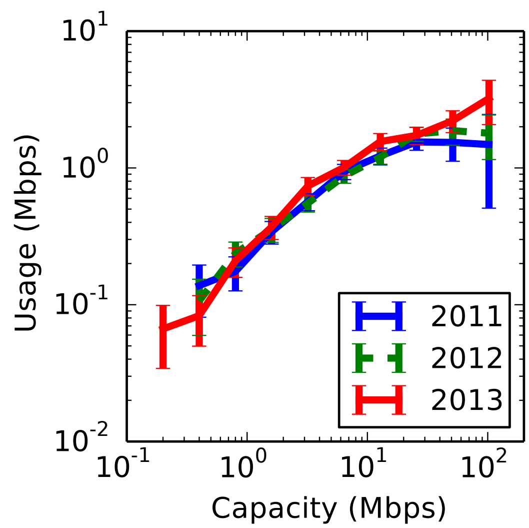
(a) Mean (w/ BT)



(b) 95th %ile (w/ BT)



(c) Mean (no BT)



(d) 95th %ile (no BT)

Figure 3.5. Peak and average usage versus capacity, grouped by year. The error bars represent the 95% confidence interval.

Trends in relative use are surprisingly different than what we expected. Despite the fourfold increase in global IP traffic, we find that subscribers' demand on the network remained constant at each speed tier. While we note a slight increase in demand for users with very fast connections (about 100 Mbps), overall the demand within the same broadband class has remained fairly consistent throughout the observation period.

Using a natural experiment, we are unable to find any significant change in demand at any given speed tier between 2011 and 2013. It would appear that growth in traffic is likely due to an increase in the number of broadband subscriptions and the increased capacity of available services.

3.3. Price of broadband access and user behavior

This section examines the impact that the price of broadband access has on user demand. In recent years, we have seen significant growth in the number of people accessing the Internet [**14**]. While increased affordability has played a critical role in this growth, the price of broadband Internet access remains unaffordable in many parts of the developing world. In countries like Iran and Botswana, a 1 Mbps plan could cost as much as $150 USD per month, after accounting for purchasing power parity (PPP). Contrast this with countries like Germany, Japan, and the US, where a 1 Mbps plan (or faster) are available for less than $25 per month.

We have seen how, up to a point, demand increases with capacity. If price is a factor that affects a customers' decision when selecting a broadband plan, then we would expect that higher prices will result in users signing up for lower capacity services despite their needs. Similarly, if two services with similar speeds are available at different prices in two markets, we would expect that the service in the more expensive market would experience higher network demand since subscribers are willing to pay more for it.

Control Group	Treatment Group	% H holds	p-value
$(\$0, \$25]$	$(\$25, \$60]$	63.4%	8.89×10^{-22}
$(\$0, \$25]$	$(\$60, \infty)$	72.2%	5.40×10^{-10}

Table 3.3. Percentage of the time that increased price results in increased usage for pairs of similar users and corresponding p-values.

We design the following study to test this idea. We define our hypothesis (H) such that users in markets where broadband Internet access is more expensive will have higher demands on the network than users in less expensive markets. Our null hypothesis (H_0) then, is that increased price does not have an affect on network demand.

For this experiment, we first need to group users based on price of broadband access in their region. We define the price of broadband access in a country as the monthly cost (USD PPP) of the cheapest service with a capacity of at least 1 Mbps. We grouped users by the cost of broadband access using the following bins: less than \$25 per month, between \$25 and \$60 per month, and over \$60 per month. Users in countries such as Germany, Japan, and the US fit in the first bin ($<$ \$25 per month). Countries such as Mexico, New Zealand, and the Philippines had prices between \$25 and \$60 per month, while prices in counties such as Botswana, Saudi Arabia, and Iran were above \$60 per month.

After placing users into groups based on the monthly price of broadband access, we compared the demand of otherwise similar pairs of users in each group. In these experiments, we use peak usage (when not active on BitTorrent) to measure demand.[1] For this experiment, users are "treated" with an increased cost, which our hypothesis says

[1]Results of experiments with and without BitTorrent for both average and peak demand were all comparable.

Country	Number of users in dataset	Median capacity (Mbps)	Nearest tier (Mbps)	Price in USD (PPP)	Annual GDP per capita (PPP)	Cost of Internet access as percent of monthly GDP per capita
Botswana	67	0.517	0.512	$100	$14,993	8.0%
Saudi Arabia	120	4.21	4	$79	$29,114	3.3%
US	3759	17.6	18	$53	$49,797	1.3%
Japan	73	29.0	26	$37	$34,532	1.3%

Table 3.4. The "typical" price of broadband in each country. The "Median capacity" column lists the maximum download capacity for the median user. We then matched the median capacity with the nearest speed tier in our set of Internet services available in a country. The "Price" column shows the price of that service (converted into US dollars using the purchasing power parity conversion factor). This price is used to calculate the monthly cost in each country as a percentage of monthly income.

will increase demand. The results are shown in Table 3.3. We find that indeed, as price increases, more users have a higher demand than those with a similar connection where access is cheaper.

Case study. We now illustrate the impact that price can have on usage with a concrete example using four markets: Botswana, Saudi Arabia, the US and Japan. We selected these four as examples of the diversity of markets in our dataset.

We chose Botswana and Saudi Arabia given that both countries were among those with the highest broadband access costs, but differed in terms of typical service capacities. Since its independence Botswana has enjoyed one of the highest GDP growth rates in the world.[2] In recent years, the country has seen rapid growth in the percentage of citizens with access to the Internet (from 3% in 2005 to 12% in 2013[3]). The cost of Internet access

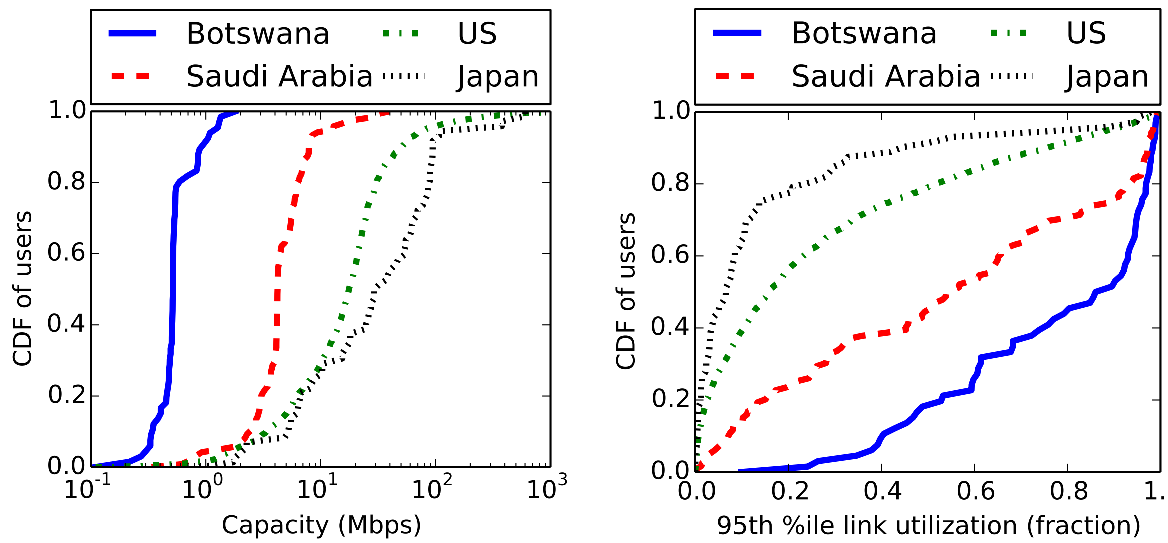
[2]CIA World Factbook. `https://www.cia.gov/library/publications/the-world-factbook/geos/bc.html`
[3]All statistics on Internet access and growth are is from ITU. `http://www.itu.int`

in Botswana, however, remains comparatively high. A 1 Mbps service, including a phone line, from Botswana Telecom costs about \$150 per month after accounting for purchasing power parity. In contrast, a 1 Mbps service in the US would cost about \$20 per month.

Over the past decade, Saudi Arabia has also experienced rapid growth in both GDP per capita (PPP) and the number of Internet subscribers. The percentage of the population using the Internet has tripled from just under 20% in 2007 to over 60% in 2013. However, according to the ITU only about 5% of the population with broadband subscriptions are on services faster than 10 Mbps (we see a similar percentage in our global dataset). A 1 Mbps connection is also relatively expensive in Saudi Arabia at about \$60 USD (PPP) per month, three times higher than a similar service in the US.

We include the US in our study as it presents another interesting case as one of the most diverse broadband service markets in terms of the available download capacities (from about 1 Mbps to over 100 Mbps). Japan, on the other hand, is one of the markets with widest availability of high-end broadband services. While the range of broadband service prices are similar to those in the US market, a larger fraction of users in Japan subscribes to high capacity services.

Table 3.4 summarizes the users and services seen in each market. We calculate the "typical" price of broadband in each of the country by matching the median capacity to the nearest service in our dataset. Compared to the US and Japan, customers in Botswana and Saudi Arabia are paying much more for slower services, especially as a fraction of monthly GDP per capita. Users in both Japan and the US appear to spend a similar fraction of monthly GDP per capita (1.3%). However, ISPs in Japan offered



(a) Download capacities

(b) 95th %ile utilization

Figure 3.6. CDF of the download capacities and peak utilization for users in each of the four markets.

higher capacities at the same fraction of monthly income. As a result, users in Japan were more likely to subscribe to faster services.

Figure 3.6a shows the maximum download throughput rates measured for connections in each country. The typical maximum download capacity increases across these markets (Botswana, Saudi Arabia, the US, Japan). We find a large number of Botswana users on a $\approx$ 512 kbps service while users in Saudi Arabia are heavily clustered around 4 Mbps. Both the US and Japan show a wider distribution across different service levels. In Japan, however, a higher fraction of users are on high-end services. The majority of users in Japan (60%) have download speeds of at least 25 Mbps, compared with over 71% of users in the US who are on services slower than 25 Mbps.

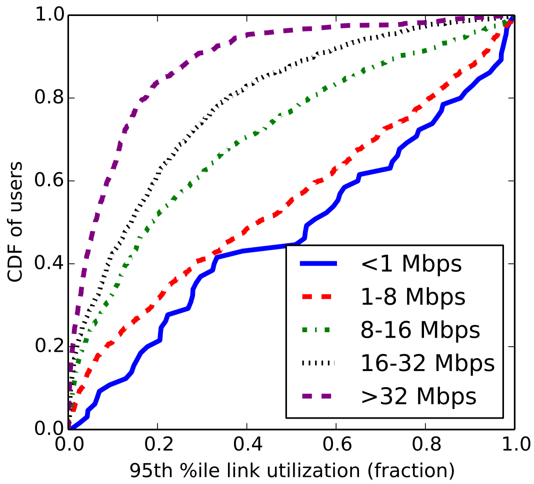
It is interesting to contrast maximum download throughout rates with the fraction of the link utilized during peak usage for each user in these four countries (Fig. 3.6b). The countries appears in exactly reverse order. Botswana shows the highest peak utilization while Japan shows the lowest. In Japan, and to some extent in the US, links tend to be very underutilized, even at the 95th percentile.

Based on our earlier findings, we expect that users in Botswana and Saudi Arabia will have higher network demands than users with similar services in the US, due to increased costs. On the other hand, users in Japan should have lower demand on the network than users with similar services in the US, due to lower service costs for the same capacity.

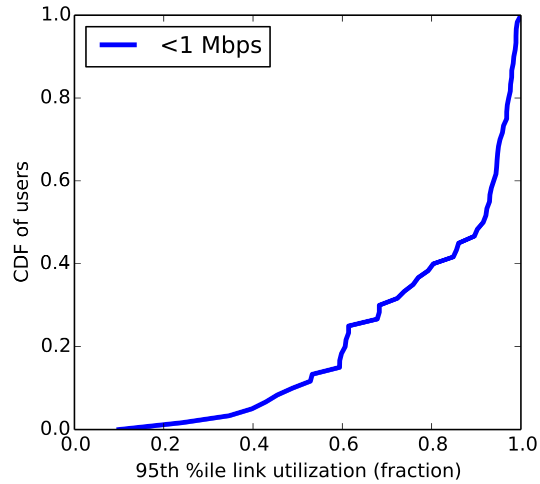
Unfortunately, at this point it is difficult to directly compare user demand in each market due to the large differences in service capacities. Therefore, we group users into different tiers of service based on their service capacity. We then compare usage within the same tier across markets. For this analysis, we selected the following tiers: below 1 Mbps, 1 to 8 Mbps, 8 to 16 Mbps, 16 to 32 Mbps, and above 32 Mbps. The selection of tiers was based on the speeds common among the broadband technologies in our dataset and the range of capacities in each country. In the following plots, we do not include data on a particular tier for a country with less than 30 users in our dataset.

Figure 3.7 shows the 95th percentile utilization of users, categorized by the aforementioned speed tiers. Figure 3.7a represents the utilization for users in the US. In this case, as customers sign up for faster services, they tend to be using less of the link during peak usage.

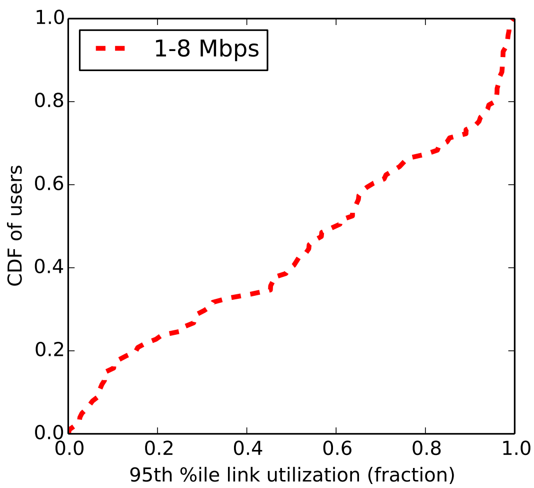
Note the higher link utilization in Botswana (Fig. 3.7b) compared to the utilization on the same tier in the US. In Botswana, the average 95th percentile link utilization was 80%; in the US, the average peak utilization was about 52%. Such a significant difference



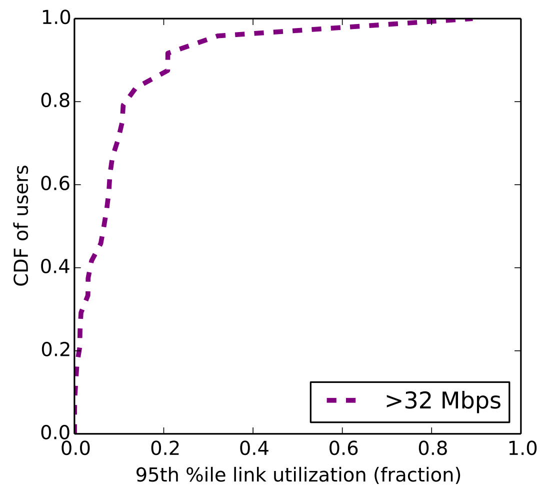
(a) US



(b) Botswana



(c) Saudi Arabia



(d) Japan

Figure 3.7. CDF of the 95th percentile link utilization for users in each country. Users are split into five different groups depending on their maximum download capacity.

could be explained by the much higher costs of faster service levels in Botswana where, for instance, a 2 Mbps plan costs about \$200 (PPP) per month!

Figure 3.7c shows a similar, but less pronounced trend in Saudi Arabia. The large majority of users in Saudi Arabia have capacities around 4 Mbps, in the 1 to 8 Mbps

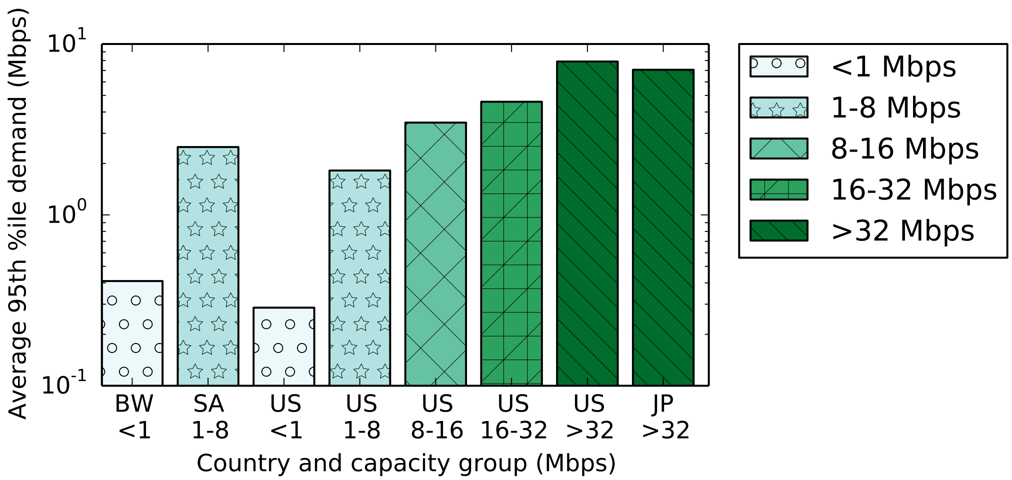


Figure 3.8. Average 95th percentile utilization for users in each country across each speed tier.

download throughput-rate range. Compared to broadband users in the US on the same tier, we also find higher utilization of the link in Saudi Arabia,. Specifically, for users in the 1 to 8 Mbps group, the median link utilization increases from about 43% in the US to 60% in Saudi Arabia.

At the other end of the spectrum is Japan, shown in Fig. 3.7d. Here we find that, for the majority of users, links tend to be very under-utilized, with an average link utilization of 10%. Overall, the fraction of the link utilized is similar to the same tier in the US, though it is slightly higher, on average, in the US.

Figure 3.8 shows the average peak demand for different tiers in each country. We note that in the US, demand increases on each tier, despite the fact that the fraction of the link utilized decreased (shown in Fig. 3.7a). We also find that when comparing across markets at the same capacity tier, in addition to having higher link utilization, users in more expensive markets also tended to have a higher total demand.

In Botswana, for example, users with less than 1 Mbps service used 410 kbps on average versus 286 kbps in the same tier in the US. Additionally, Fig. 3.8 shows that the demand on the network is 676 kbps (37%) higher in Saudi Arabia than on the same tier in the US. In fact, the average demand of the 8-16 Mbps tier in the US is only 39% higher than the 1-8 Mbps tier in Saudi Arabia, but is 90% higher than the 1-8 Mbps tier in the US. This difference supports our belief that the relatively high price of Internet access in the country, rather than user need, is preventing users in this market from signing up for faster services.

Similarly, users in the US with a service faster than 32 Mbps use 830 kbps more than users on the same tier in Japan. Despite the fact that the cost of broadband access is similar in both Japan and the US, the availability of faster services at a lower cost leads subscribers to sign up for services that will be less heavily used. We examine this trend in the next section.

3.4. Cost of upgrading and user behavior

Subscribers select broadband service based on their needs, the set of available plans and the plans' prices. Thus, given the diversity in service availability across markets, users with similar needs will end up choosing different broadband services, depending on what is available. In this section, we look at how the relative cost of alternative services impacts user demand.

Beyond price, broadband service markets differ in the relative cost of upgrading services. For example, according to our dataset of service plans, both Japan and the US have similar prices of broadband access with a connection of at least 1 Mbps costs less than \$25 per month. The two markets differ, however, in service availability and the cost of upgrading. In Japan, a 100 Mbps plan is considerably less expensive than in the US (\$40 per month instead of \$115 per month). Furthermore, in contrast to the US, the broadband service market in Japan has more options with capacities above 50 Mbps and fewer fixed-line services below 10 Mbps.

It is clear that the cost of upgrading capacity, similar to the cost of a particular service level, can have an impact on a demand users impose on their service. To explore this, we begin by generalizing the cost of increased capacity. To this end, we collect all service plans for each country, perform a linear regression analysis on each market, and measure the correlation between capacity and price. We find that, in the majority of these markets (66%) there is a strong correlation (> 0.8) between price and capacity and in 81% there is at least moderate correlation (> 0.4).

In markets where there is weak or no correlation, price is often affected by other factors. For example, in Afghanistan, it is possible to sign up for a dedicated (not shared) DSL connection that is slower and more expensive than alternatives, lowering the correlation coefficient between price and capacity. Whether or not a service is wireless or has a monthly traffic cap would also affect the relationship between price and capacity.

For markets where price and capacity are at least moderately correlated ($r > 0.4$), we use the slope of the linear regression line to estimate the cost of upgrades (the slope is measured in monthly price per Mbps increase in capacity). Figure 3.9 presents a CDF of the cost of increasing capacity by 1 Mbps for all markets in our dataset.

For illustration, we note in the figure where a few representative markets fall in the distribution. At the lower end of the curve (less than \$0.10 to upgrade), we find regions

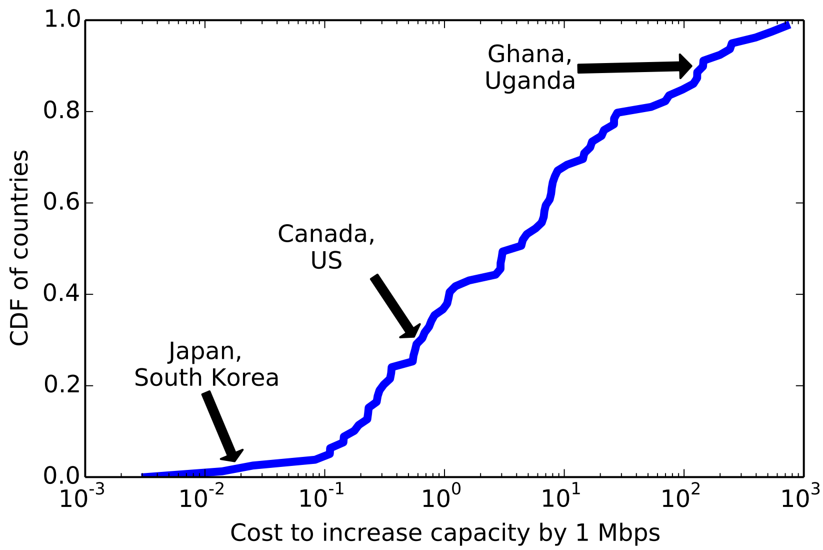


Figure 3.9. CDF of the monthly cost (after accounting for PPP) to increase broadband service capacity by 1 Mbps in a given country's broadband market. The arrows point out where the labeled countries were placed in the distribution.

such as Hong Kong, Japan, and South Korea. Countries such as Canada and the US are at slightly above \$0.50 per Mbps increase. The higher end of the distribution is largely comprised of countries in Africa and the Middle East, like Ghana and Uganda.

As shown in Fig. 3.9, increasing capacity by 1 Mbps tends to cost less than \$1 per month in developed countries but can be well above \$100 (PPP) in some developing countries (e.g., Paraguay and Ivory Coast).[5] Table 3.5 summarizes this distribution by aggregated region, presenting the percentage of countries, per region, where the cost of increasing capacity by 1Mbps is above \$1, \$5 and %10 (PPP) per month. The trends are strikingly clear – for 74% of the countries in Africa and 43% of those in the Middle East, for instance, the costs of an additional 1Mbps is above \$10 per month.

[4]As defined by the International Monetary Fund.

[5]Two exceptions in Asia are India and China, where upgrading capacity cost less than \$1 per Mbps per month.

Region	> \$1	> \$5	> \$10
Africa	100%	84%	74%
Asia (all)	67%	47%	33%
Asia (developed)	0%	0%	0%
Asia (developing)	83%	58%	42%
Central America/Caribbean	100%	86%	14%
Europe	10%	0%	0%
Middle East	86%	57%	43%
North America	0%	0%	0%
South America	78%	55%	33%

Table 3.5. The percentage of countries in each region where increasing capacity costs more than \$1, \$5, and \$10 per month for a 1 Mbps increase in capacity. We split Asia into two subgroups, developed and developing, given the diversity of economies within the area.[4]

To test for the impact of service upgrade on user demand we define a new study. For this experiment, our hypothesis (H) states that as the cost to upgrade increases, users are less likely to upgrade and will have higher network demand than users in markets where upgrading is cheaper. Our null hypothesis (H_0) is then that the price of upgrading will not affect demand. We use the cost of upgrade to split broadband markets into three classes: countries where the cost of increasing a service by 1Mbps is (i) below \$0.5, (ii) between \$0.5 and \$1 and (iii) above \$1.00 per Mbps.

We present the results of this experiment in Table 3.6, for average demand with and without including BitTorrent traffic. In general, increased upgrade prices do lead to higher demand. It is clear that users in developing countries tend to use more than similar users where faster service are more readily accessible. In cases where our results are statistically significant, we can reject the null hypothesis, and assert that the price of increasing capacity affects demand. Our results are inconclusive, i.e., p-value slightly

Control Group	Treatment Group	% H holds	p-value
($\$0, \$0.50$]	($\$0.50, \$1.00$]	53.8%	0.00717
($\$0.50, \$1.00$]	($\$1.00, \infty$]	58.7%	0.0110

(a) Average demand w/ BitTorrent

Control Group	Treatment Group	% H holds	p-value
($\$0, \$0.50$	($\$0.50, \$1.00$]	52.2%*	0.0947
($\$0.50, \$1.00$]	($\$1.00, \infty$]	56.3%	0.0265

(b) Average demand w/o BitTorrent

Table 3.6. Percentage of the time that a higher cost to increase capacity (price per 1 Mbps increase) will result in higher network usage. An asterisk denotes that a result was not statistically significant.

higher than 0.05 when comparing demand (without BitTorrent) between markets where the cost of upgrade are ($\$0, \$0.50$] and ($\$0.50, \$1.00$].

We have already visited an example of the impact that the cost of increasing capacity can have on (Figs. 3.7 and 3.8). While both Japan and the US have similar monthly cost of broadband access, the costs of increasing capacity is over five times higher in the US explaining the observed higher demand in the US.

3.5. Connection quality and user behavior

Previous works have shown that poor connection quality can have a negative impact on a user's quality of experience [**32**]. In this last section, we explore the potential impact that the quality of a connection, specifically latency and packet loss, has on user demand.

We hypothesize that a sufficiently poor quality of experience could lead to a decrease in demand on the broadband service. In the following paragraphs we test whether this is true by studying the impact of both long latencies and high packet loss rates. As we

Control Group	Treatment Group	% H holds	p-value
$(512, 2048]$	$(0, 64]$	63.5%	0.00825
$(512, 2048]$	$(64, 128]$	63.4%	0.00620
$(512, 2048]$	$(128, 256]$	59.4%	0.00766
$(512, 2048]$	$(256, 512]$	56.3%	0.0330

Table 3.7. Percentage of the time that decreasing latency will result in higher 95th percentile usage (without BitTorrent). Very high latency (over 512 ms) to the nearest NDT server appears to result in lower demand than comparable users with lower latencies.

have done in our previous comparisons, we study the effects of these factors by comparing users that are similar in terms of link capacity and location. When testing the effects of increased latency, we require that average packet loss rates are similar between matched users and vice versa.

3.5.1. Latency

We first look at the impact of latency on user behavior. In this case, our hypothesis (H) is that decreasing latency will result in higher demand. Therefore, our null hypothesis (H_0) is that decreasing latency does not affect demand and the interaction will be random.

We present the results of the study in Table 3.7. The table compares the peak demand (95th percentile usage when BitTorrent is not active) of users with problematically high latencies, above 512 ms in our dataset. Users are divided among exponentially increasing sized bins; our control and treatment groups in this case are the higher and lower latency groups, respectively. The results show that there is a significant increase in usage when switching from very high latency to any lower latency group, leading us to reject the null hypothesis.

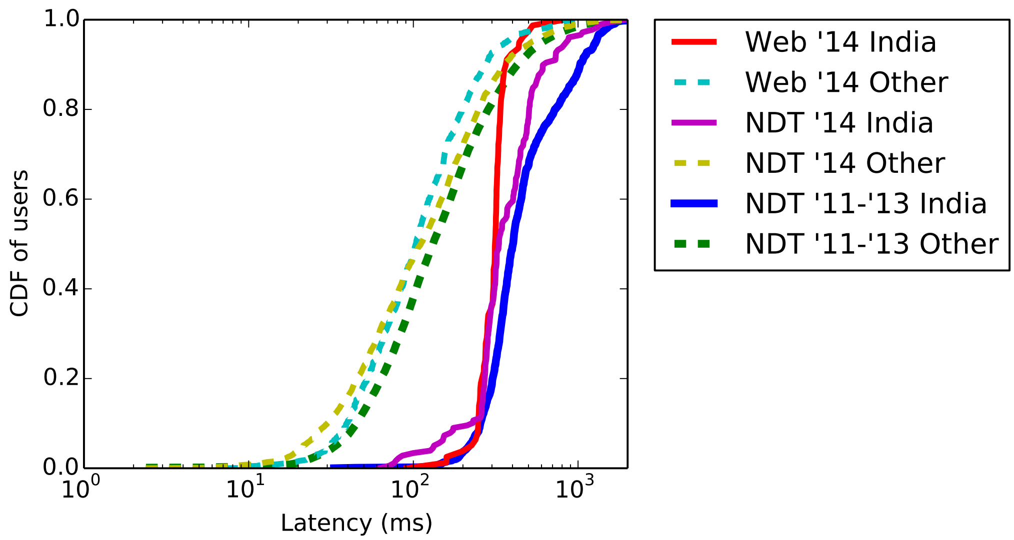


Figure 3.10. CDF of latency measurements for users in our dataset, grouped by location (India versus the rest of our sample population, labelled "Other"). "Web" represents each user's median latency to five of Alexa's Top Sites from our 2014 dataset. "NDT" represents the average latency to the nearest NDT server. We include NDT data from 2014 that was collected from the same set of users as the web 2014 data.

While the case of latency impacting demand is visible in multiple countries, the impact of high-latency is clear when focusing on users in India. In our previous analysis broadband service plans, we find that the cost to increase capacity is similar in both the US and India (both are within 25% of each other). The cost of broadband access, however, is much higher in India ($67 versus $20). Thus, we would expect usage to be higher in India. When comparing users in India to users with similar capacities in the US, we find, surprisingly, that users in India tend to impose lower demand 62% of the time (p-value < 0.001).

An analysis of NDT latencies shows that users in India report much higher latencies to NDT servers than users in other countries. The trends are not restricted to NDT servers but can also be seen when looking at latencies to the set of five globally popular websites:

Facebook, Google, Windows Live, Yahoo, and YouTube. Four of these websites (Google, Facebook, YouTube, and Yahoo) accounted for the top five most popular websites in India[6] while Windows Live was ranked 26th.

Figure 3.10 describes these latency measurements, and compares them by user location (India versus the rest of our sample population). The lines labelled "Web" correspond to the median latency to the five popular websites while "NDT" is the average latency to the nearest NDT server (measured by NDT). We include data from two time periods – 2011 through 2013 (labelled "'11-'13") and from May 2014 to August 2014 (labelled "'14") to compare NDT and website latencies. [7]

Figure 3.10 shows that the distribution of latency measurements is similar for both NDT traces and the typical latency to the top Alexa sites. For the majority of users in India, we find much higher latencies to both NDT and popular websites compared to the rest of our sample population; nearly every user has a latency longer than 100 ms. Since we rejected the null hypothesis that latency does does not affect demand, we believe that the higher latency for users in India contributes to the fact that we see a decrease in network usage in India.

3.5.2. Packet loss

Next we examine the impact of packet loss on user demand. Our hypothesis (H) is that decreased packet loss rates result in higher demand. The results of this experiment are shown in Table 3.8. We find that when comparing users with very low packet loss rates to comparable users with very high packet loss rates, usage tended to be higher

[6]Ranked by `http://www.alexa.com`
[7]We added the website latency experiment later in the study.

Control Group	Treatment Group	% H holds	p-value
$(0.1\%, 1\%]$	$(0, 0.01\%]$	55.4%	5.85×10^{-6}
$(0.1\%, 1\%]$	$(0.01\%, 0.1\%]$	53.4%	8.55×10^{-4}
$(1\%, 15\%]$	$(0, 0.01\%]]$	58.9%	2.16×10^{-5}
$(1\%, 15\%]$	$(0.01\%, 0.1\%]$	53.8%	0.0360

Table 3.8. Percentage of the time that decreasing packet loss will result in higher average usage (without BitTorrent). Very high packet loss (above 1%) appears to lead to lower demand than comparable users lower packet loss rates.

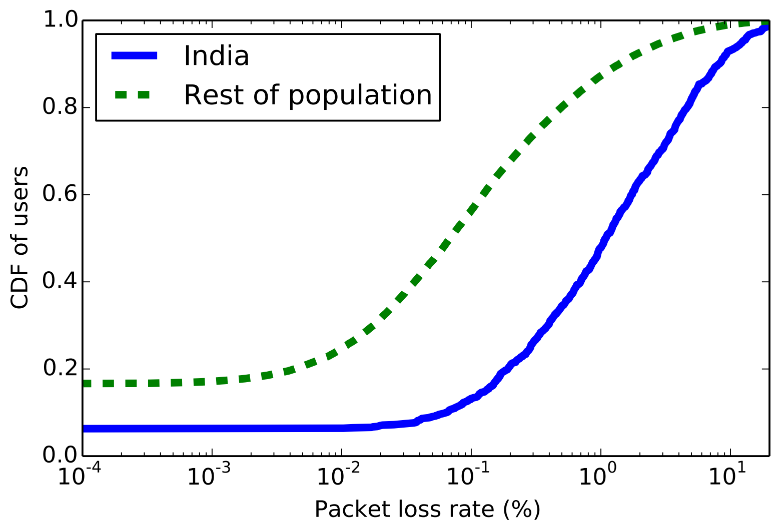


Figure 3.11. CDF of the average packet loss during measurements to the NDT servers for all users and users in India.

on connections with lower packet loss rates. This trend was most pronounced when comparing to connections with packet loss rates above 1%.

This impact of packet loss can be illustrated by looking at the behavior of users in India, as done in Sec. 4.2. We again found that users in India had much higher packet loss rates than the general population, as shown in Fig 3.11. As a result, we believe that

Treatment group	% H holds	p-value
(0.5%, 1%)	48.1	0.792
(1%, 2%)	57.7	0.0356
$> 2\%$	60.4	0.00862

Table 3.9. Percentage of the time that a higher average packet loss rates will result in lower usage. Users in the control group have similar download capacities with an average packet loss rate between 0% and 0.0625%.

the lower quality connections (both increased packet loss rates and latency) in India are the probable cause of lower demand on the network.

3.5.3. Experiment results

Several possible experiments can shed light on how service reliability affects user behavior. Although we expect that usage will drop around a single outage, we aim to understand how poor reliability over longer periods of time affects user behavior. Our experiments test the effects on user demand of connections that are consistently lossy and connections that have frequent periods of high loss.

High average loss. To look at how consistently lossy links affect user demand, we calculate the average packet loss rate over the entire period during which the user is reporting data. We then group users based on their average packet loss rate. We select users from each treatment group and match[8] them with users in the same region with similar download and upload link capacities (within 10% of each other) in the control group. Users in the control group have an average loss rate of less than 0.0625%. Our hypothesis, H, is that higher average packet loss rates will result in lower usage, due to

[8]In observational studies, matching tries to identify subsamples of the treated and control units that are "balanced" with respect to observed covariates.

Control group	Treatment group	% H holds	p-value
(0.5%, 1%)	(1%, 10%)	54.2	0.00143
(0.1%, 0.5%)	(1%, 10%)	53.2	0.0143
(0%, 0.1%)	(1%, 10%)	54.8	0.000421
(1%, 10%)	$> 10\%$	68.3	3.65×10^{-05}
(0.5%, 1%)	$> 10\%$	70.0	6.95×10^{-06}
(0.1%, 0.5%)	$> 10\%$	70.8	2.87×10^{-06}
(0%, 0.1%)	$> 10\%$	72.5	4.34×10^{-07}

Table 3.10. Percentage of the time that users with more frequent high-loss hours ($\geq 5\%$ packet loss) have lower network usage.

a consistently worse experience. Our null hypothesis is that average packet loss and user demand are not related. Table 3.9 shows the results of this experiment.

We see usage is significantly affected even for average packet losses above 1% – in 57.7% of our cases show a lower volume of traffic with a p-value of 0.0356. This leads us to reject the null hypothesis.

Frequent periods of high loss. To understand the effects of frequent periods of high loss on user behavior we calculate, for each user, the fraction of hours where the gateway measured more than 5% packet loss. We group users based on how frequently periods of high loss occurred. For example, users that recorded loss rates above 5% during 0% to 0.1% of measurements were placed in a group that we used as one of the controls. We then compared the network demands between each pair of user groups. In this case, our hypothesis, H, is that groups with a high frequency of high loss rates (treatment group) will have lower usage than groups with a low frequency of high loss rates (control group). Table 3.10 shows the results of this experiment.

We find that users with high packet loss rates during more than 1% of hours, tend to have lower demand on the network. As the difference between the frequency of high loss rates periods increases, the magnitude of this effect increases, with larger deviations form the expected random distribution.

Previous studies have discussed the importance of broadband service reliability [**58**], and surveys of broadband users have shown that reliability, rather than performance, has become the main source of user complaints [**68**]; our findings are the first to empirically demonstrate the relationship between service reliability and user traffic demand.

CHAPTER 4

Characterizing broadband services

In this chapter, we leverage longitudinal measurements from both gateways and end hosts to characterize the performance offered by broadband service providers. We characterize service in terms of service capacity, access latency, and service reliability.

One objective of broadband characterization efforts is to inform policymakers and businesses on the state of broadband Internet access in a country. Additionally, some works aim to inform consumers that are comparing broadband service providers. Typically, the findings of these works describe an ISP's performance at a country-wide granulartiy; they often provide a national average for each ISP's bandwidth capacity, last-mile latency, or packet loss rates.

In many cases, we find that performance for each metric can vary widely across regions, even within the same ISP and access technology. Such wide variations in performance demonstrate that basic descriptive statistics, such as mean and standard deviation, of low-level network metrics aggregated at a nation-wide granularity, are not sufficient for providing a meaningful characterization of broadband services.

4.1. Service capacity

We first evaluate ISPs in terms of download and upload throughput. While other works have looked at characterizing broadband access speeds, in many cases, their findings are based on "one-time" measurements being run on demand by users [**17**, **55**]. While such measurement techniques are able to capture large-scale trends in performance if run by a large enough group, they are, in general, unable to capture the variations in service that the individual sees from their Internet service. The datasets used in this study, both the FCC's Measuring Broadband America and Dasu, allow us to perform a longitudinal analysis of performance for individual users.

However, one challenge in comparing performance across providers and services, is that users do not have the same subscription speeds; individual ISPs typically offer a number of service capacities and the stated capacities of such offerings vary from one ISP to another. We address this issue in two different ways for the FCC and Dasu datasets to allow us to directly compare the consistency of performance.

For the FCC data, we first normalize throughput measurements by the speed that each user should be receiving. As mentioned in Chapter 2.1, in addition to including each participant's ISP and access technology, the FCC dataset also includes the download and upload subscription rate for the majority of users. We use this data to normalize throughput performance. For data collected from Dasu users, we normalize all measurements by maximum download throughput rate achieved for each user.

4.1.1. Gateway FCC measurements

Throughput distribution. Figure 4.1 shows a CDF of each normalized download throughput measurement from subscribers of four services: AT&T's DSL service, Clearwire, Comcast, and Frontier's fiber service. Of the services we studied, Frontier's fiber service had the most consistent throughput rates, both in terms of the fraction of measurements that measured at least 90% of the subscription speed and in showing the least

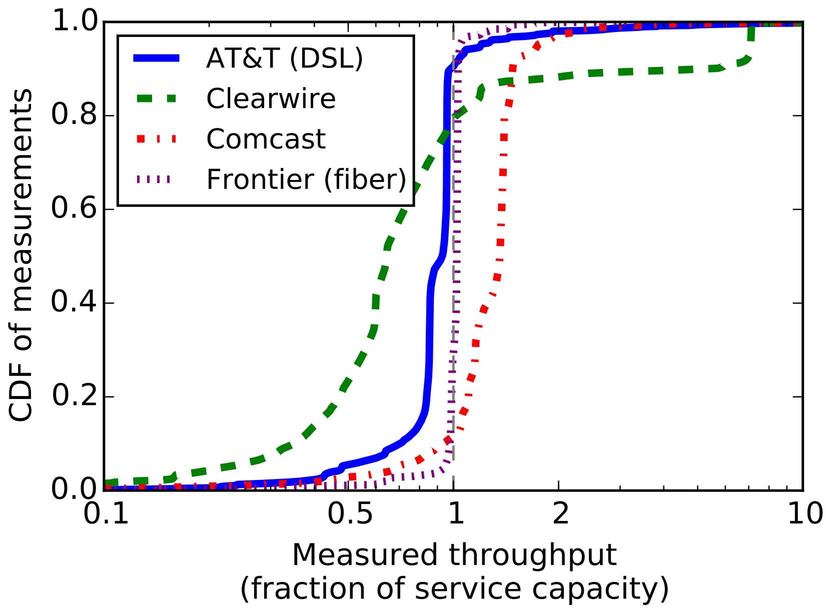


Figure 4.1. CDF of measured download throughput rates as a fraction of the subscriber's service capacity.

variations between measurements. Although measurements were unlikely to achieve download rates significantly higher than their subscription speed, 96% of measurements were above 90% of the subscription speed.

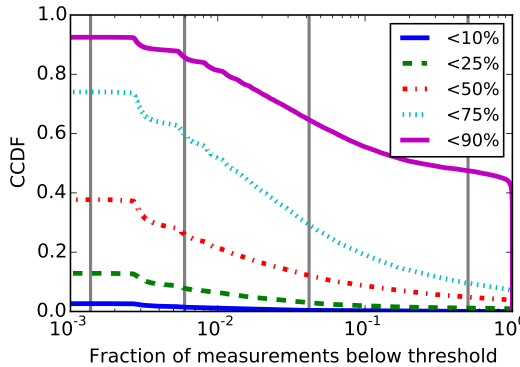
For Comcast (cable), measurements were slightly less likely to reach 90% of the subscription speed (about 91%). However, download throughput measurements were often much higher than the user's subscription speed; the median measurement on Comcast's network was 135% of the subscription speed. We observed a similar trend for most cable broadband providers, as well as Verizon's fiber service.

Download throughput measurements from subscribers of AT&T's DSL service were fairly consistent (i.e., showing little variation). However, in contrast to cable and fiber services, they rarely exceeded the subscription speed, with less than 10% of measurements at or above the subscription speed. Nearly half (48%) of measurements were below 90% of the subscription speed. Other DSL providers showed a similar trend. Of the ISPs in our study, Clearwire had the largest fraction of measurements (73%) below 90% of the subscription speed.

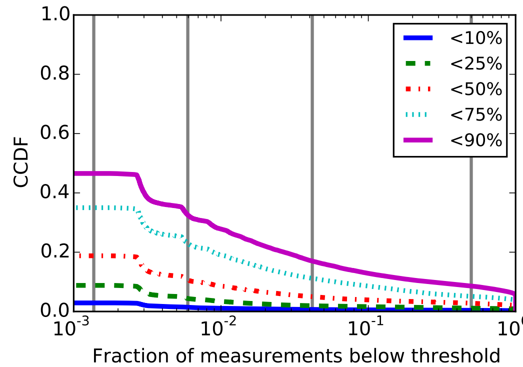
Variation over time. Looking only at Figure 4.1, it is still unclear how much performance can vary for an individual subscriber over the course of a month. To measure this, we aggregated all measurements that were conducted from the same vantage point and run during the same month, which we refer to as a "user-month". For each user-month, we calculate the fraction of measurements that were below a threshold of 10%, 25%, 50%, 75%, and 90% of the subscription speed.

Figure 4.2 shows, for AT&T, Comcast and Frontier fiber subscribers, how frequently measurements during the same month measured below a particular threshold. The vertical gray lines represent a particular frequency of throughput measurements being below a given threshold (from left to right: once a month, once a week, once a day, and once every other hour).

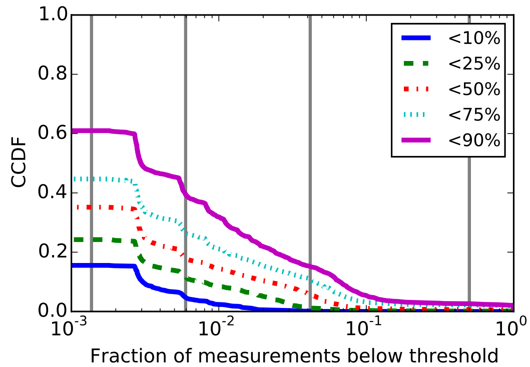
In the case of AT&T, shown in Figure 4.2a, during 47% of the user-months, subscribers got less than 90% of their service capacity at least once every other hour (the right-most vertical line). In contrast, for Comcast subscribers, shown in Figure 4.2b, only about 9% of user-months measured less than 90% of the subscription speed at the same frequency. Comcast users were also less likely to receive less than 50% of their subscription speed. Frontier's fiber service was even less likely to have degradations in download throughput every other hour; less than 3% of months of Frontier measurements saw throughput rates below 90% of the subscription speed.



(a) AT&T (DSL)



(b) Comcast



(c) Frontier (fiber)

Figure 4.2. CCDF of the fraction of a user's download throughput measurements per month that are below a percentage of the subscription capacity. Each gray vertical line represents a frequency of (from left to right) once a month, once a week, once a day, and once every other hour.

In general, the distribution of upload throughput measurements, shown in Figure 4.3, was similar to download throughput. The most obvious difference was that upload measurements from Clearwire subscribers were noticeably higher, more consistent, and much closer to the subscription speed. For each ISP in Figure 4.3, the median measurement was at least 90% of the subscription speed.

Figure 4.4 shows the frequency of measured throughput rates below a given threshold. Compared to the download throughput measurements for the same ISPs, a lower fraction

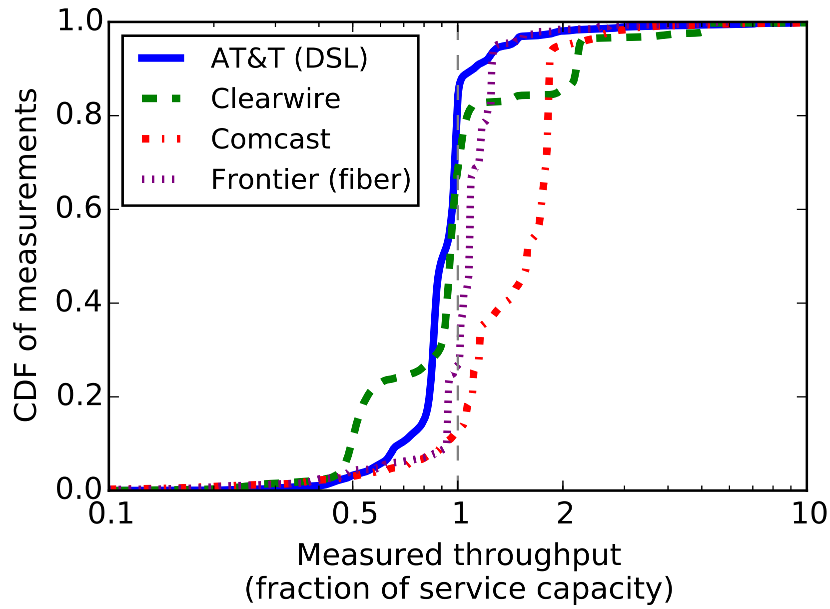


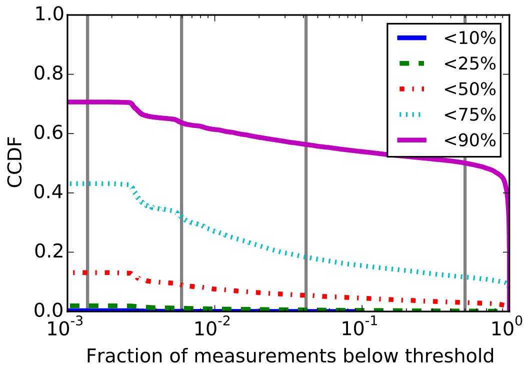
Figure 4.3. CDF of upload throughput rates as a fraction of the subscriber's service capacity.

of user-months had some sort of degradation in upload performance. However, the fraction of measurements below the threshold, for all thresholds – from once a month to once every other hour – remained fairly stable for these three providers. We observed the same trend for most ISPs in the FCC dataset.

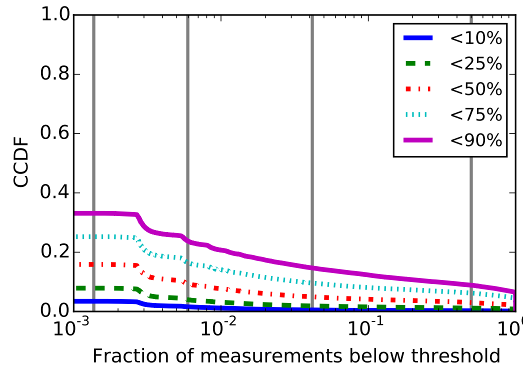
4.1.2. Dasu end host measurements

Variation over time. In this section, we compare the variance in performance seen by Dasu users subscribing to two ISPs in the UK.

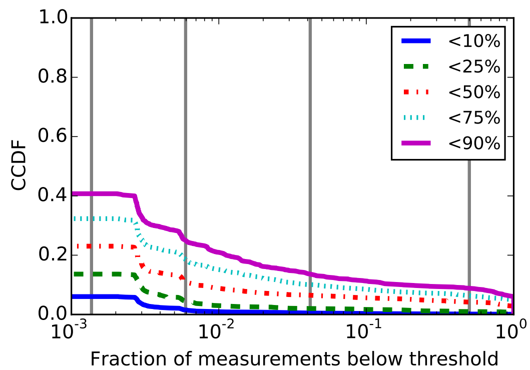
To measure the variations in performance seen by each user (identified by a UUID), we aggregated all NDT measurements run from the host over an nine month period. Each set of download throughput results was normalized by the maximum, then used to



(a) AT&T (DSL)



(b) Comcast



(c) Frontier (fiber)

Figure 4.4. CCDF of the fraction of a user's upload throughput measurements per month that are below a percentage of the subscription capacity. Each gray vertical line represents a frequency of (from left to right) once a month, once a week, once a day, and once every other hour.

calculate the standard deviation for each individual user. As such, the standard deviation is represented as a fraction of the maximum throughput rate.

Figure 4.5 compares the standard deviation in download throughput seen by users in BT and TalkTalk networks within the same city. The figure shows that a user in BT's network is more likely to see lower variance in their download throughput performance compared to TalkTalk's network. For example, the standard deviation of the median user in BT's network is about 16% of the maximum, 7% lower than the 23% of the median

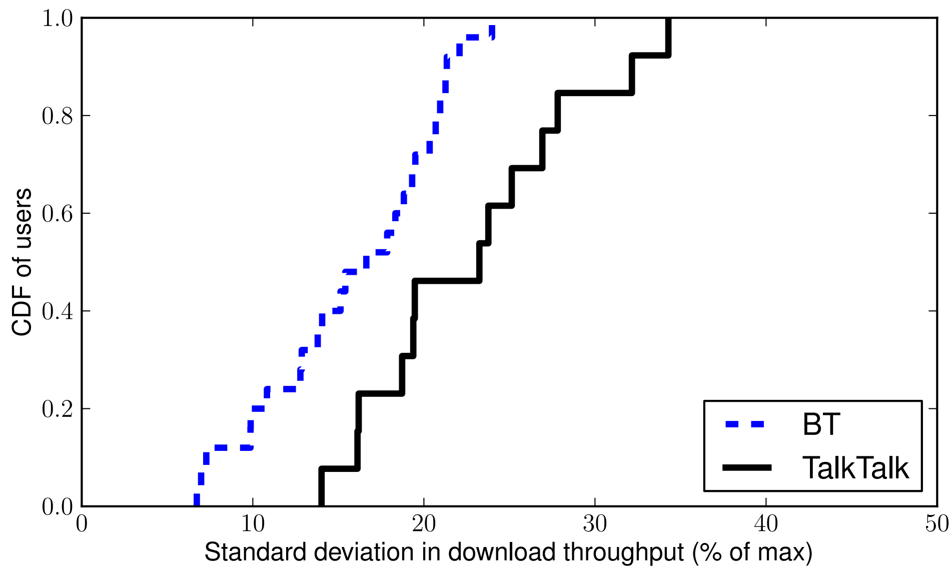


Figure 4.5. CDFs for all users in BT's and TalkTalk's network of the standard deviation of each user's NDT download rate measurements. Each user's measurements are normalized by that user's maximum measurement to account for differences in maximum throughput rates.

user in TalkTalk's network. Additionally, the maximum standard deviation in TalkTalk's network is 34%, 10% higher than the 24% in BT. On average, the distributions are separated by approximately 10% at each percentile.

To demonstrate what these differences in standard deviation look like, we selected three users that fell into lower half of the BT distribution and three users that fell into the upper half of the TalkTalk distribution. Figure 4.6 shows the normalized distributions of download rates for all six users. We see that the distributions of BT users with a lower standard deviation are pulled down and to the right, representing a more stable set of measurements. In fact, the slower measurements recorded by these hosts occurred during a similar time window. We believe this may have been caused by a temporary issue within BT's network. On the other hand, the distributions of TalkTalk users are pulled up and

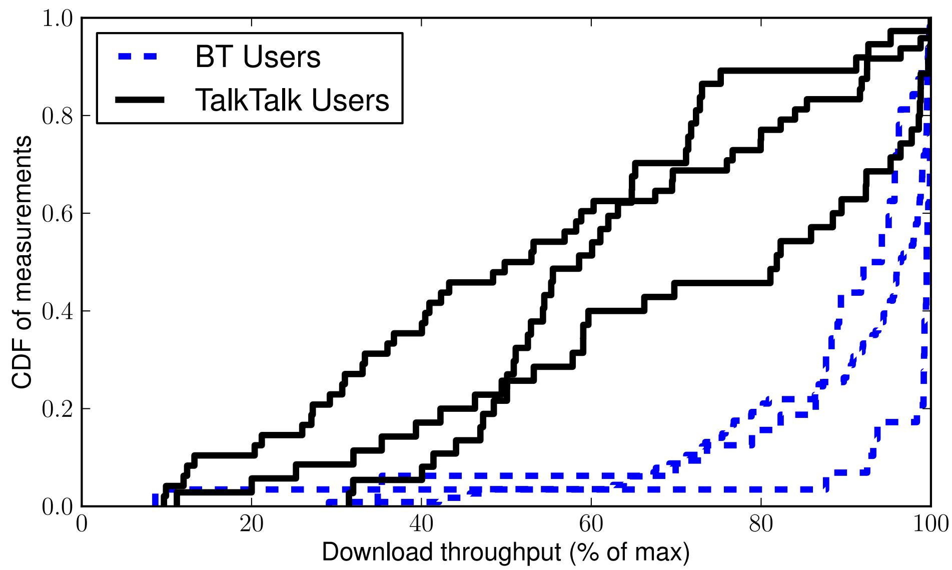


Figure 4.6. CDFs of the normalized download rate measurements for three individual users on BT's and TalkTalk's Internet services.

to the left. For two of the users, the 50th percentile measurement is less than 60% of their maximum.

Continuous monitoring from a single user is necessary to capture the variance in performance that an individual receives due to time-of-day effects or ISP policies and to capture capture transient issues, such as a service's downtime. The value of continuous monitoring is demonstrated further in the additional examples in the following paragraphs.

Time of day effects. Other studies have shown that the service users receive from their ISPs at any time of day partially depends on the particular access network technology (e.g. DSL or Cable) and the traffic management techniques their ISPs implement. Dischinger et al. [**31**] showed that while DSL ISPs have bandwidth rates that roughly correspond to those advertised, the performance of some Cable ISPs can vary significantly during the day.

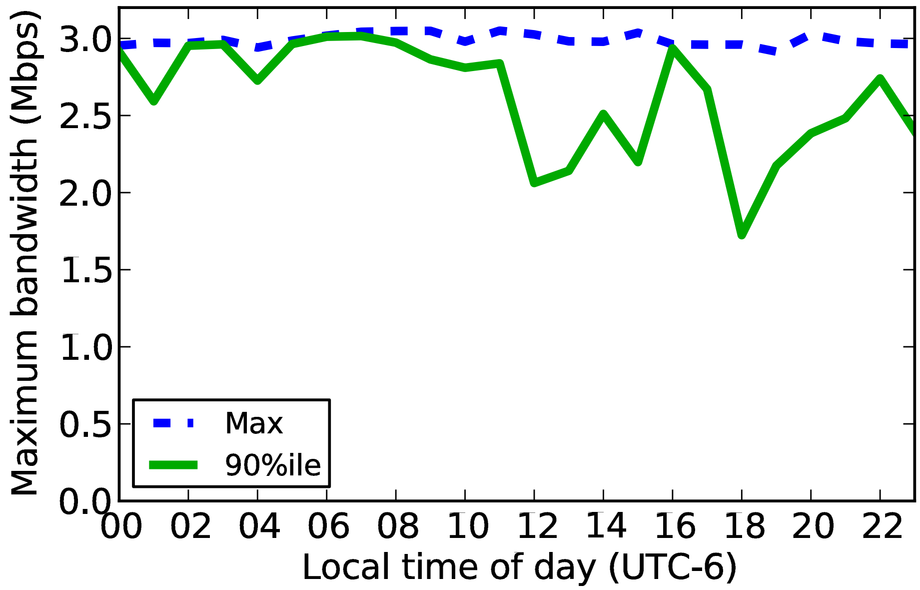


Figure 4.7. The maximum and 90th-percentile download speed of peers in Rogers' network for each hour-long period during Nov. 2009.

Accurately capturing ISP performance, thus, requires nearly continuous monitoring. By passively monitoring the performance of network-intensive applications, it is possible to efficiently capture such time-of-day effects.

We use Rogers, a Canadian ISP providing access via cable, as an interesting example, due to its diurnal patterns. Figure 4.7 shows the maximum and 90th-percentile download rate seen, over a 24-hour period, by Rogers customers subscribed to a level of service of 3 Mbps. Each data point is the average of all such rates from subscribers during a 1-hour interval (i.e., all statistics gathered between 2:00:00 PM and 2:59:59 PM on each day of the month are considered together) during November 2009. To ensure that all customers included in the analysis are subscribed to the same level of service, we only consider peers that had reported a maximum download rate within that level of service.

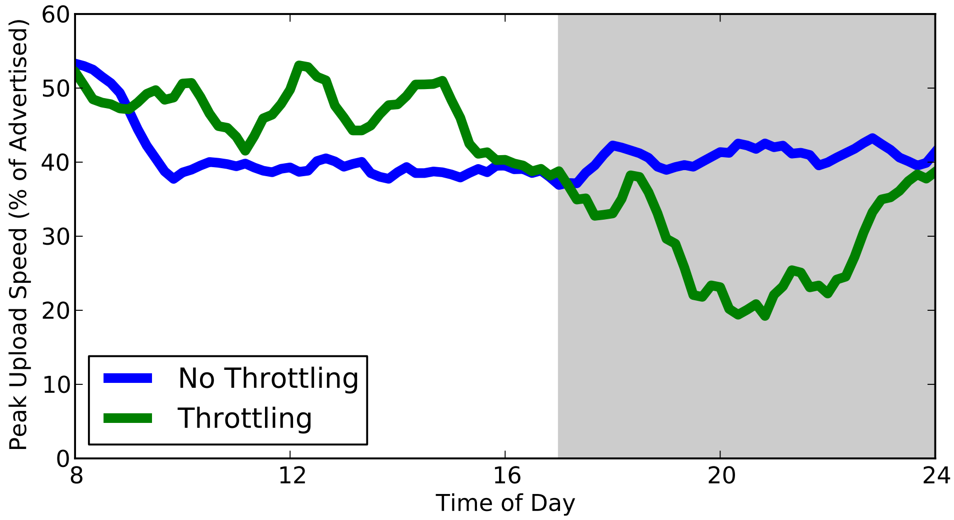


Figure 4.8. The weekday peak upload speed achieved by users on Virgin Media's service during the week before throttling was implemented ("No Throttling") and the first week during the trial ("Throttling"). The shaded region represents the time of day when traffic management was enabled, from 5 PM to midnight.

The graph shows that the reported maximum achieved transfer rate for each hour reaches the 3 Mbps limit fairly consistently. However, the 90th-percentile download rate drops from an average of 2.9 Mbps between 6 and 10AM to as low as approximately 1.7 Mbps during the peak residential hours of 5-10PM. This corresponds to approximately 96% of the advertised service during low activity and about 60% during peak residential usage hours. This drop is consistent with that reported by Dischinger et al. [**31**] for a different dataset.

Detecting changes in policy. Next we look at changes in an ISP's traffic management policy and how it can affect a user's performance. In March of 2011, Virgin Media, an ISP in the UK, completed a week-long test of its new upload bandwidth throttling policy that targets peer-to-peer traffic during peak usage times.

Figure 4.8 plots upload speeds of Virgin Media subscribers before and after the throttling policy was implemented. We find that this policy reduces users' upload bandwidth by as much as 50%, based on anonymized statistics contributed by P2P users inside the Virgin network. On weekdays in the week before the throttling trial, users consistently used about 40% of their upload capacity. However, during the test (when P2P upload rates were throttled from 5PM to midnight), the maximum upload speeds we observed were significantly lower, sometimes only 20% of their capacity. The most significant reduction in speed occurs between 8 PM and 10 PM.

Here we have demonstrated the importance of longitudinal measurements for broadband service comparison. In order to fully understand what type of service a potential customer will receive, an approach to characterization must support longitudinal measurements on a per user basis. Approaches that require a user initiated test will likely miss these variations in performance. As such, findings that demonstrate the degree to which performance can fluctuate on a particular ISP require multiple measurements over time.

4.2. Access latency

Next we look at how latency varies across different access technologies, service providers, and even across regions within the same ISP. In our analysis of the FCC's data, we use the latency to a nearby measurement server to estimate access latency. With our end host data from Dasu, we analyze traceroute data to measure last-mile latency by looking at the RTT to the first hop with a public IP address.

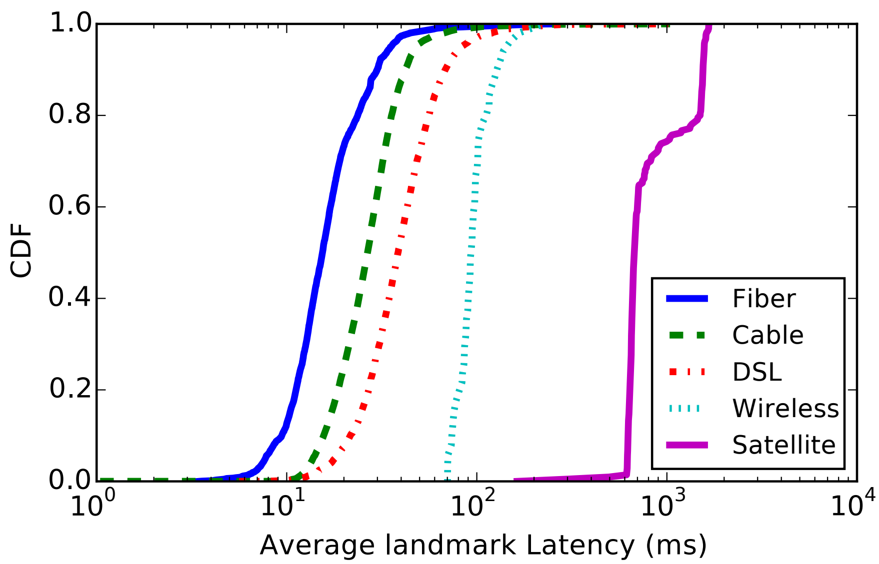


Figure 4.9. Landmark latency.

Figure 4.10. Latency performance of all users in the FCC study grouped by access technology.

4.2.1. Gateway FCC measurements

As mentioned in Chapter 2.1, in our analysis of the FCC data, we use the average latency (measured over an hour) to the nearest measurement server as an estimate of the subscriber's access latency. Unfortunately, except for the most recently published data, the measurement gateways distributed by the FCC do not directly measure last-mile latency.

Access technology. Figure 4.9 shows the average RTT to the measurement server for all users in the FCC study, grouped by access technology. It shows the distinct performance characteristics between different technologies, as each technology operates under different latency performance due to the nature of their access technology. For instance, satellite and wireless operators see much higher latency due to their last mile restrictions.

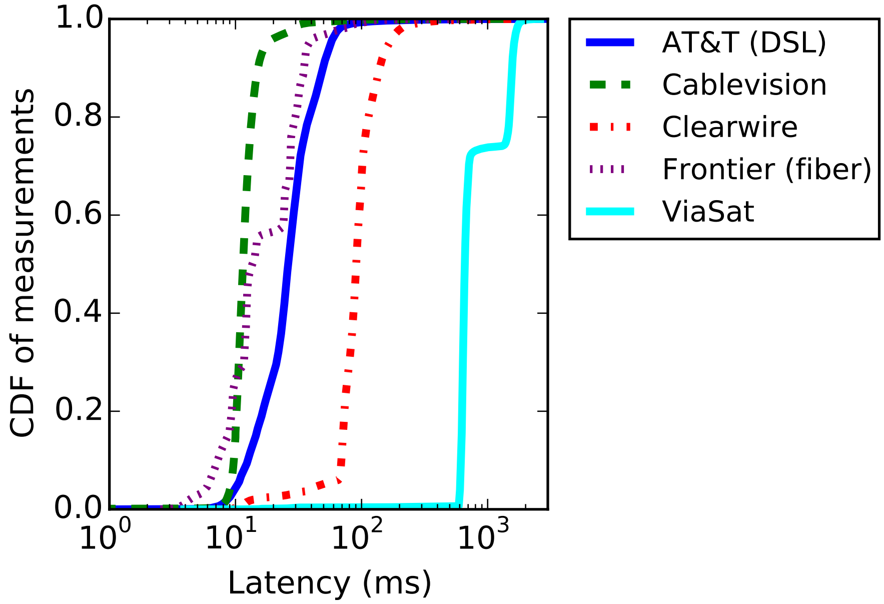


Figure 4.11. CDF of latency measurements to servers.

Unsurprisingly, fiber providers, overall, had the lowest latencies to the measurement servers. However, some cables were able to do as well or better than some fiber providers. Figure 4.11 shows a CDF of the hourly average latency for five ISPs, one for each technology represented in Figure 4.9. Cablevision, with 96% of hourly averages below 20 ms, showed the lowest latencies of all ISPs in our dataset and appeared to consistently have the lowest latencies.

Other fiber, DSL, and cable ISPs had slightly higher latencies, but were fairly consistent in terms of ordering, with at least 90% of average latency reports for each provider being less than 70 ms. AT&T, with 95% of measurements below 57 ms had the lowest latencies of all DSL providers, but the overall distribution was higher than all fiber and most cable providers.

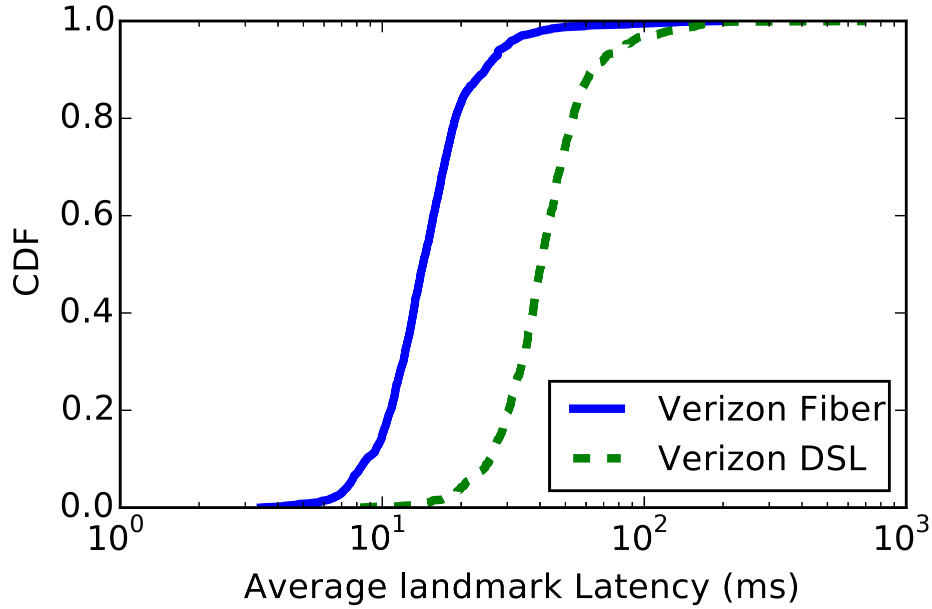


Figure 4.12. Average landmark latency for Verizon subscribers using two different access technologies.

Latency measurements from Clearwire subscribers were noticeably higher, with a median of approximately 90 ms. Satellite providers had the highest latency measurements, consistently measuring above 600 ms as a result of the fundamental limitations of the technology.

The previous figures show that there is a clear relationship between access technology and latency. However, using causal inference to determine whether differences are due to differences in technology is difficult, due to the fact that access technology largely depends on the service provider – the majority of the providers in our dataset only use a single access technology. In other words, differences in performance between cable and fiber subscribers could be caused by a difference in provider, not access technology.

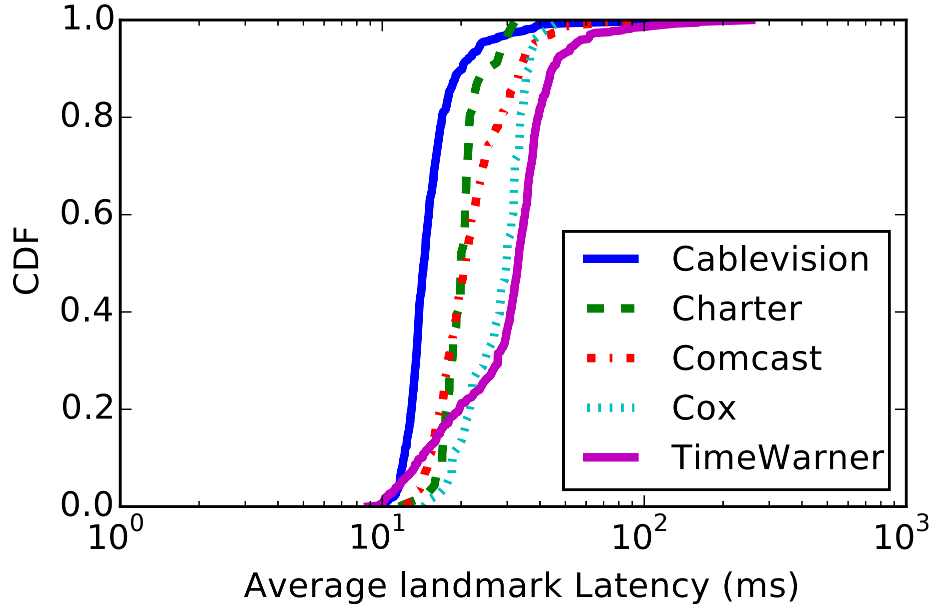


Figure 4.13. Average landmark latency for cable service subscribers in the Northeast with a capacity of at least 10 Mbps.

Fortunately, a small number of ISPs in the US provide services over multiple access technologies. Frontier and Verizon subscribers both had a large number of DSL and fiber to the premises (FTTP) subscribers. Figure 4.12 shows the latency for both sets of subscribers In the case of Verizon, the median latency for Verizon DSL subscribers (40 ms) was nearly three times higher than the latency for Verizon Fiber subsribers (15 ms). This suggests that access technology does play a significant role in determining service latency.

Service Provider. Overall, we found that a subscriber's provider was the most informative feature for predicting a user's network latency. However, it is important to remember that this feature is strongly correlated with access technology. Still, even providers using the same access technology showed significant variation in latency.

We compared the distribution of average network latency for subscribers of different ISPs, keeping constant the region, service capacity and access technology of each ISP compared. Figure 4.13 shows the average landmark latency for cable broadband subscribers in the Northeast region with a capacity of at least 10 Mbps. We observed the largest difference in latency between Cablevision and TimeWarner, where the overall average doubled from 16 to 33 ms. In most cases, we found that the differences between each provider's distributions of measurements were statistically significant.

Geographic Location Broadband access networks offer various levels of service to customers. These performance differences are the result of infrastructure heterogeneity, which differ in the access technologies used and in the differential quality of each provider's underlying physical infrastructure. Today's broadband service providers exist as large conglomerates, built from acquisitions of various small, regional companies offering telephone and cable services. For example, Comcast Communications, the largest cable and high-speed internet provider in the United States, began as a regional cable company in Mississippi with only 15,000 subscribers in 1963 and has grown to its current base of over 22 million subscribers, nearly through acquisition alone.

The effect of this growth pattern is a large amount of diversity within the performance of a large ISP's network. Each smaller ISP acquired differed in the quality of their design and construction of the underlying infrastructure.

As a result, location can also play a large role in an ISP's performance, particularly for nationwide ISPs, due to the fact that many of today's broadband service providers grew through a series of acquisitions. For example, today's AT&T is a result of a multiple breakups and acquisitions; this history is reflected in the diversity of their network.

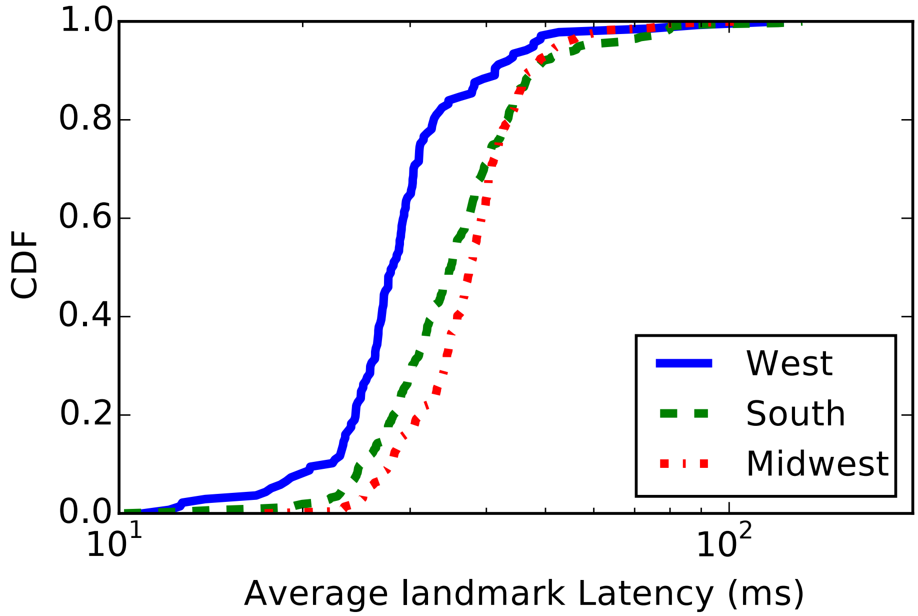
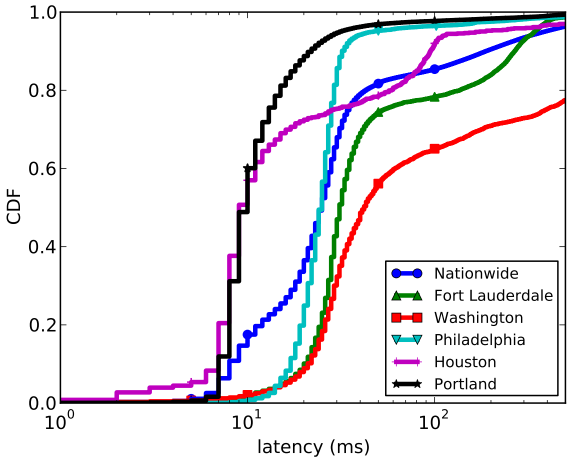


Figure 4.14. Average landmark latency for AT&T subscribers with a capacity of at least 10 Mbps.

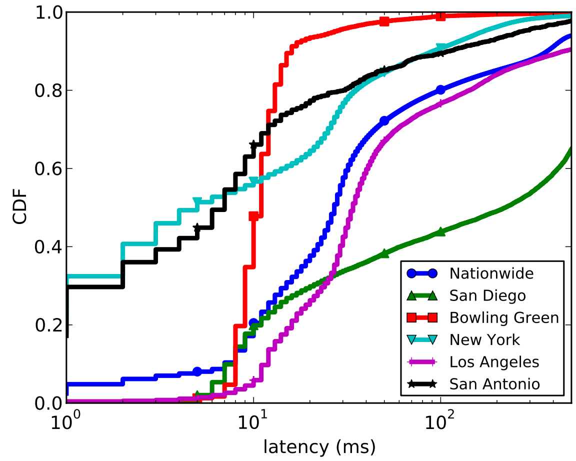
Figure 4.14 shows the average landmark latency for subscribers of AT&T's services with capacities above 10 Mbps across three geographic regions. While the difference between the South and Midwest distributions is relatively small, they both differ significantly from the West region.

4.2.2. Dasu end host measurements

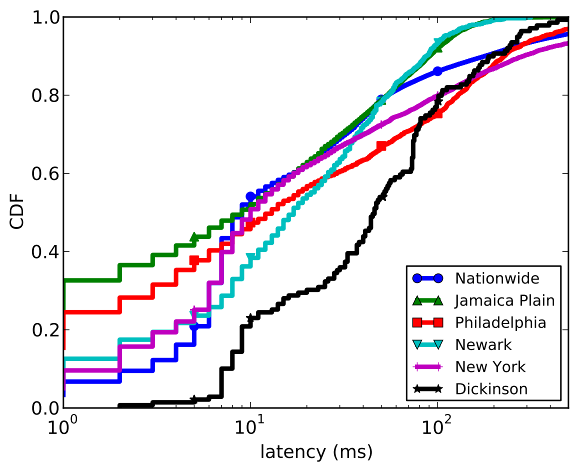
Last mile latency is frequently a significant contributor to the end-to-end latency of Internet paths [**93**]. A high last-mile latency effectively increases the delay for all network communication, which can significantly reduce application performance. As such, this is an important metric for benchmarking ISPs.



(a) Comcast



(b) Time Warner Cable



(c) Verizon

Figure 4.15. CDFs of last-mile latency measurements from users in three large ISPs, aggregated across all locations ("Nationwide") or separated out for 5 sample cities.

To evaluate variations in last-mile latency, we select measurements from users in 3 large American ISPs and either aggregate them by all users ("Nationwide") or separate out those from users in 5 different cities. Figure 4.15 plots CDFs for each of these sets of measurements. For instance, the nationwide median latency for Comcast (Fig. 4.15a) is 25 ms, but the median for individual cities ranges between 10 (Portland and Houston) and 40 ms (Washington); this means that typical performance can vary from 33% to 160% of the nationwide value. In addition, the bimodal distributions of Houston and Fort Lauderdale (with smaller modes at the 90% mark at 90 ms and 250 ms, respectively) indicate the presence of additional variability at a finer granularity that is not captured in our city-level aggregation.

In the case of Time Warner Cable (Fig. 4.15b), the variability between cities compared to the nationwide distribution is even wider than in Comcast, ranging from 10% (New York) to 1000% (San Diego) of the nationwide median latency. Though most of Time Warner's locations are characterized by significant variation within the city, Bowling Green is an exception, with 90% of measurements in a very tight range, between 8 and 20 ms, indicating that Time Warner Cable is consistent in that city.

In contrast with Comcast and Time Warner Cable, each of the latency distributions for Verizon (Fig. 4.15c) is spread across a wide range of values. The middle 50% of the nationwide distribution ranges from 6 ms to 45 ms, nearly an order of magnitude difference. The lack of obvious modes in the distributions for Jamaica Plain and Philadelphia, for instance, mean that any estimate of last mile latency in these cities would have a very wide confidence interval given the observed variations.

While we have focused only on last mile latency here, these results indicate diversity in the underlying access networks that serve subscribers from the same ISP across geographic regions. The impact of this diversity is not restricted to last mile latency; previous work has shown that throughput rates may also vary across geographic regions – even when controlling for the same ISP service level [**11**].

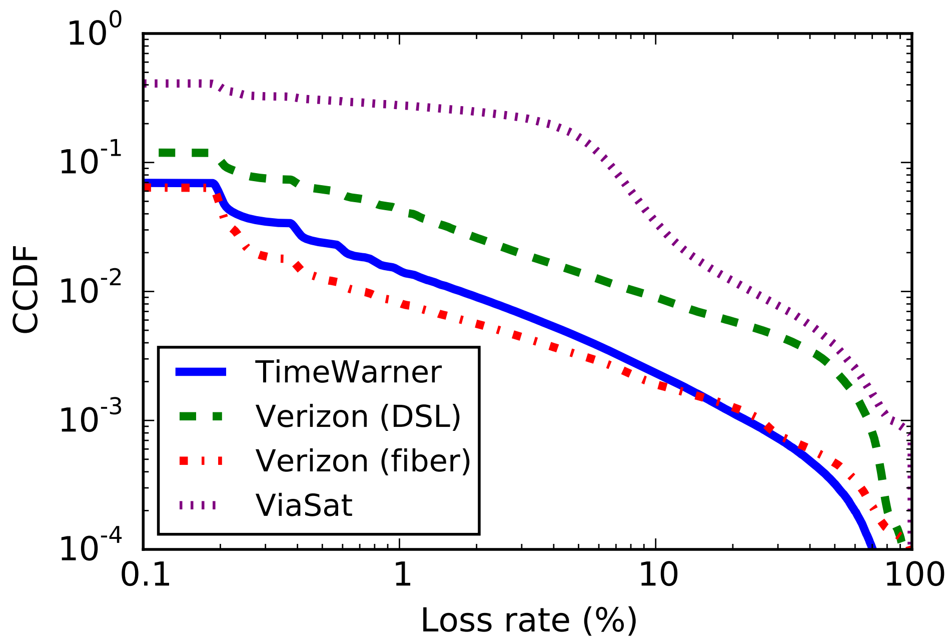


Figure 4.16. CCDF of hourly loss rates to servers.

4.3. Service reliability

Using the number of UDP pings that succeeded and failed to the target measurement server, we calculated the percentage of packets lost over each hour. Figure 4.16 shows the CCDF of the hourly packet loss rates for four ISPs. On average, fiber providers tended to have lower loss rates and had the lowest frequency of high loss. More specifically, Verizon had the lowest frequency of loss rates above 1%, occurring during only 0.82% of hours. Comcast (not in the figure) and TimeWarner had the lowest frequency for cable providers, with loss rates above 1% occurring in approximately 1.5% of hours. Satellite providers had the highest frequency of loss rates above 1%, occurring during over 26% of hours.

In this section we present an approach to characterizing broadband service reliability designed to be applicable to the datasets being collected by many ongoing national broadband measurement studies. This design decision imposes several constraints on our analysis method, such as the type and granularity of metrics and the placement of vantage points. At the same time, this makes our approach applicable to the various dataset being generated and, we hope, can spur future designs to better capture all aspects of broadband service reliability.

In the following paragraph we describe ongoing broadband measurement efforts before presenting our methods and metrics for characterizing service reliability. We then discuss our findings concerning the reliability of broadband services in the US.

4.3.1. Approach

Available data. Over the last decade, the number of governments with national broadband plans has increased rapidly [**52**] and several of these governments are funding studies to characterize the broadband services available to their citizens. Two well-known examples are the efforts being carried by the UK Ofcom and the US FCC in collaboration with the UK company SamKnows. In the few years since their initial work with Ofcom, SamKnows has started working with at least six additional governments including the US, Canada, Brazil, the European Union and Singapore. Data for these efforts is typically collected from modified residential gateways distributed to participants in a range of service providers.

Metrics. To analyze the data from these efforts, we use a number of traditional metrics and measure the reliability of a service. These metrics are defined based on an understanding of what constitutes a failure which we discuss in the following paragraphs.

We define the *reliability* of a broadband service as the average length of time that the service is operational without interruption and *availability* as the fraction of time the service is in functioning condition. Note that a system can be highly available, but have low reliability.

We adopt several classic metrics from reliability engineering, including *Mean Time Between Failure* (MTBF) and *Mean Down Time* (MDT). MTBF is the average time that a service works without failure; it is the multiplicative inverse of Failure Rate, formally defined as

$$MTBF = \sum \frac{\text{Total uptime}}{\text{\# of failures}}$$

To characterize the length of time a service is unavailable during each failure, we use MDT, which is defined as

$$MDT = \sum \frac{\text{Total downtime}}{\text{\# of failures}}$$

We can now define availability (A) as the probability that at any given point in time, the service is functioning/operational. Unavailability is the complement of availability. More formally

$$A = \frac{\text{MTBF}}{\text{MTBF} + \text{MDT}}$$

$$U = (1 - A)$$

Definition of a failure. What constitutes a failure or outage in the context of broadband services is a critical issue tightly coupled to the collected metrics. Although the definition of failure is obvious in many systems, it is less clear in the context of "best-effort" networks.

We choose to identify connections failures by detecting significant changes in lost packets. It is unclear what packet loss rate (or rates) should be used as thresholds for

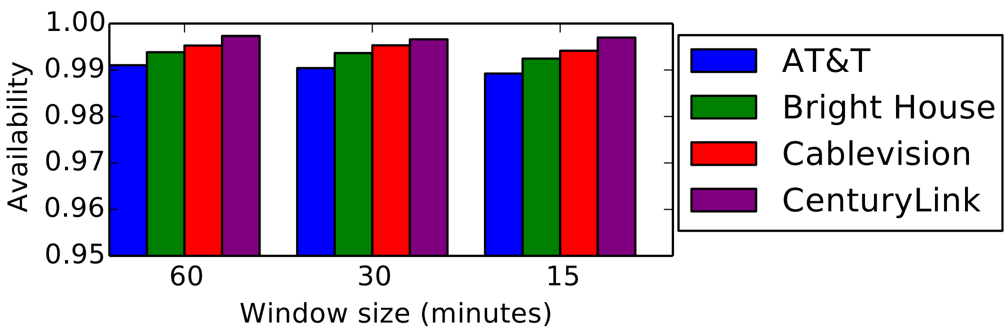


Figure 4.17. Service availability for four ISPs across multiple observation window sizes.

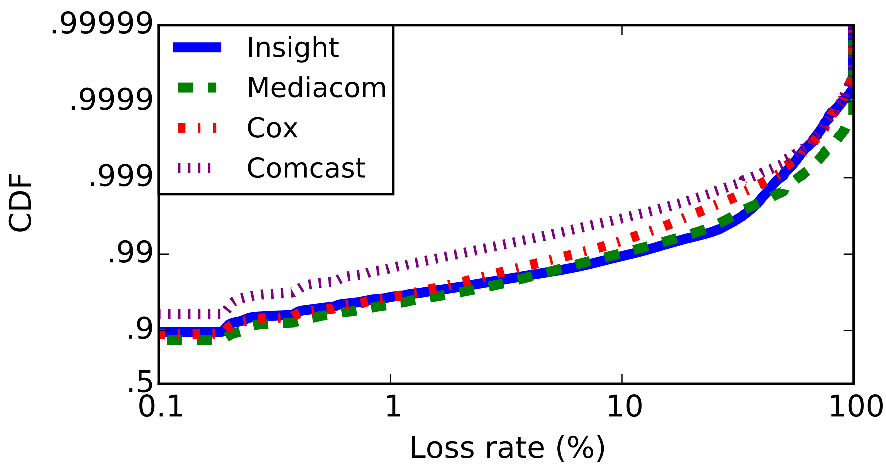


Figure 4.18. CDF of the hourly loss rates measured from gateways of four cable providers. Lower curves indicate a less available service; curves crossing over each other implies that different loss-rate thresholds would yield different rankings.

labeling failures. Achievable TCP throughput varies inversely with the square root of loss rate [**61**, **73**] and even modest loss rates can significantly degrade performance. Xu et al. showed that video telephony applications can become unstable at a 2% bursty loss [**98**], with significant quality degradation occurring around 4% in some cases. In our analysis, we use three thresholds for classifying network failures – 1%, 5%, and 10%.

While the FCC MBA dataset is currently the largest publicly available dataset on broadband service performance, relying on it for our analysis means we are only able to measure loss rates at a one-hour granularity. To evaluate the impact of monitoring granularity, we rely on a platform installed in 6,000 end-hosts to measure loss by sending packets approximately every five seconds and use this data to calculate loss rate using different sizes and loss rate thresholds. Figure 4.17 shows the availability of four ISPs in our dataset using the 10% loss threshold. We found that changing the window size has little impact on our calculation of availability and the relative ranking of ISPs.

The distribution of loss rates are quite different for different broadband technologies, and can vary even across providers with the same technology at different loss rate thresholds. Figure 4.18 shows the CDF of loss rate of four cable providers, with the y-axis showing the cumulative fraction of all hourly time intervals. Although two providers may offer the same MTBF for a particular loss rate threshold, considering the difference in loss rate distributions, a different definition of "failure" could result in a different ranking. For instance, defining a failure as "an hour with $>$ 1% packet loss" yields a similar MTBF for both Cox and Insight Cable ($\approx$ 27.5 hours), using a 10% loss rate threshold, but results in a MTBF over 50% higher for Cox ($\approx$ 150 hours) than for Insight ($\approx$ 94 hours).

The assessment of broadband reliability could focus on different aspects, ranging from the reliability of the network connection, the consistency of performance, and the availability of services offered by the ISP, such as DNS servers and email [**58**]. The primary focus of this work is on broadband service reliability, under which we include both the

Technology	% of participants
Cable	55%
Cable (business)	1%
DSL	35%
Fiber	7%
Satellite	1%
Wireless	1%

Table 4.1. Percentage of the sample population in the FCC's dataset using each access link technology.

availability of connection itself as well as that of the ISP's DNS service. From the perspective of most users, failures in either are indistinguishable. We plan to study other aspects of service reliability, such as performance consistency, in future work.

4.3.2. Characterization of service reliability

We apply the approach presented in the previous section to characterize the reliability of broadband services in the US using the FCC MBA dataset. We first provide a short summary of the population of participants in the SamKnows/FCC study. In our study we seek to understand the role that a set of key attributes of a subscriber's connection play in determining its reliability: (1) How does reliability vary across different providers? (2) What is the impact of using different access technologies or (3) subscribing to different tiers of service? (4) Does geography have an effect? (5) How reliable is the provider DNS service?

Sample population description. As part of the MBA dataset, the FCC also provides metadata about each participant including the user's service tier (i.e., subscription speed), service technology (e.g., cable or DSL), and geographic location. Combining this information with the loss rate data described in Section 4.3.1, we compare the reliability of broadband services across different axis.

The list of ISPs covered in the sample population includes both large, nationwide ISPs and smaller, regional ISPs. Since the number of devices per ISP is weighted by the number of subscribers, most devices (71%) are located in larger ISPs (AT&T, Comcast, and Verizon).

The FCC's dataset includes a diverse set of technologies, including satellite and fixed wireless providers. Table 4.1 shows a summary of the distribution of participants by access technology. "Wireless" access refers to fixed wireless (not mobile) from providers such as Clearwire, where users connected there FCC-provided device to a wireless modem. Additional information, such as the process used for selecting participants, can be found in the technical appendix of the FCC's report [**36**].

To understand the relative importance of the different attributes collected, we calculated the information gain—the degree to which a feature is able to reduce the entropy of a target variable—of each attribute of a subscriber's connection (ISP, download/upload capacity, region, and access technology). We found the subscriber's ISP to be the most informative feature, with access link technology as a close second, for predicting service availability. In the rest of this section we analyze the impact of these attributes on service reliability. We close with an analysis of DNS and ISP reliability.

4.3.2.1. Effect of ISP. We first characterize service *availability*— the probability that a service is operational at any given point in time—for each provider in our dataset. Table 4.2 lists the average availability per ISP, as well as the provider's unavailability,

ISP	Average availability			Average annual downtime (hours)		
	1%	5%	10%	1%	5%	10%
Fiber						
Frontier (Fiber)	98.58	99.47	99.77	124	46.8	20.3
Verizon (Fiber)	99.18	99.67	99.80	72	29.2	17.8
Cable						
Bright House	98.21	99.28	99.58	156	62.8	36.7
Cablevision	98.33	99.53	99.70	146	41.4	25.9
Charter	97.84	99.29	99.59	189	62.5	36.1
Comcast	98.48	99.45	99.66	134	48.0	29.7
Cox	96.35	98.82	99.33	320	103.0	58.4
Insight	96.38	98.31	98.94	318	148.0	93.0
Mediacom	95.48	98.34	99.03	396	146.0	85.3
TimeWarner	98.47	99.48	99.69	134	45.9	26.9
DSL						
AT&T	96.87	99.05	99.42	274	83.3	51.1
CenturyLink	96.33	98.96	99.39	322	90.9	53.7
Frontier (DSL)	93.69	98.18	98.87	553	160.0	98.7
Qwest	98.24	99.24	99.51	154	66.7	42.8
Verizon (DSL)	95.56	98.43	99.00	389	137.0	88.0
Wireless						
Clearwire	88.95	96.96	98.13	968	266.0	164.0
Satellite						
Hughes	73.16	90.15	94.84	2350	863.0	453
Windblue/Viasat	72.27	84.20	96.37	2430	1380.0	318.0
Windstream	94.35	98.72	99.42	495	112.0	50.6

Table 4.2. Average availability and annual downtime for subscribers, per service, for three different loss-rate thresholds.

described as the average annual downtime (in hours). We evaluate both metrics in the context of the three loss rate thresholds for network failures measured over an hour.

We find that, at best, some providers are able to offer two nines of availability. Verizon's fiber service is the only one with two nines of availability at the 1% threshold. At 5%, about half of the providers offer just over two nines. The satellite and wireless services (Clearwire, Hughes, and Viasat) provide only one nine of availability, even at the 10%

ISP	A	% change in U	A	% change in U
	1%		10%	
Satellite				
Hughes	60.97	+45.4	91.38	+66.9
Wildblue/ViaSat	69.44	+10.2	94.14	+61.2
Windstream	89.17	+91.8	99.13	+50.4
Wireless				
Clearwire	86.35	+23.6	97.57	+29.9
DSL				
Frontier (DSL)	87.98	+90.4	98.42	+39.9
Verizon (DSL)	93.95	+36.2	98.90	+9.9
CenturyLink	94.19	+58.2	99.35	+6.9
AT&T	95.85	+32.4	99.38	+5.4
Qwest	97.92	+18.5	99.51	+1.2
Cable				
Cablevision	97.76	+34.2	99.64	+22.6
TimeWarner	98.03	+28.5	99.69	+1.3
Insight	95.31	+29.4	98.98	-3.9
Charter	97.75	+4.2	99.61	-6.4
Mediacom	94.52	+21.1	99.09	-7.0
Comcast	98.39	+5.3	99.70	-11.7
Brighthouse	98.15	+3.5	99.63	-11.8
Cox	96.30	+1.3	99.42	-13.3
Fiber				
Frontier (Fiber)	98.56	+1.4	99.78	-4.6
Verizon (Fiber)	99.11	+8.7	99.83	-14.7

Table 4.3. Average availability (A) and percent change in unavailability (U) for subscribers of each ISP during peak hours.

loss rate threshold. For comparison, five nines is often the target availability in telephone services [**65**].

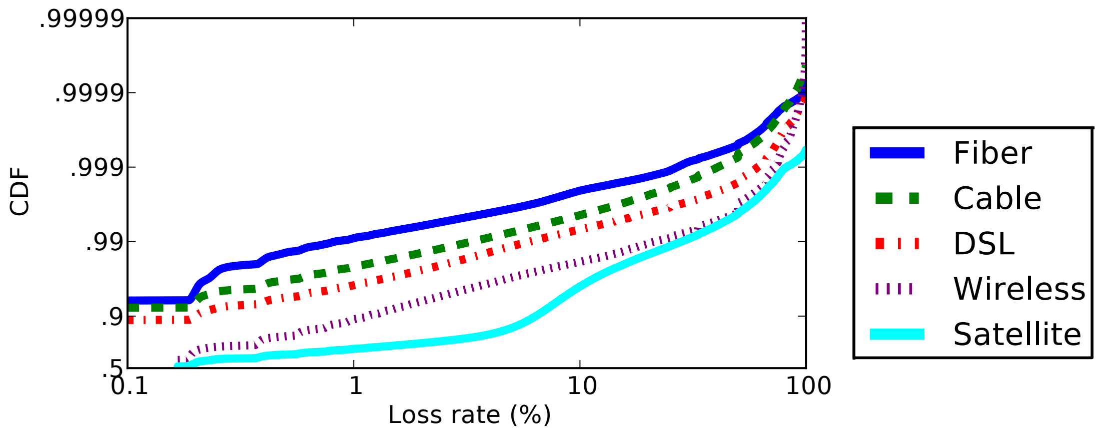
Because broadband users are more likely to be affected by outages in the evening, we also measured availability during peak hours (from 7PM to 11PM, local time), as shown in Table 4.3. Although all providers show a lower availability at the 1% loss rate threshold compared to their full-day average, most cable providers actually performed better at a 10% loss rate threshold. We expect that some of these providers may perform planned maintenance, which would introduce extremely high periods of loss (> 10%), during the early morning or midday. On the other hand, DSL, wireless, and satellite providers continued to have lower availability during peak hours, as compared to their average availability over all time periods.

We also analyzed the MTBF for each provider, which represents the average time between periods with high packet loss. Most ISPs appear to maintain a MTBF of over 200 hours (≈8 days), but a few experience failures every 100 hours, on average. ClearWire, Hughes, and Viasat again have notably low MTBF: 73.8, 26.0, and 4.78 hours, respectively. CenturyLink and Mediacom offer the two lowest MTBFs for DSL and cable providers, respectively.

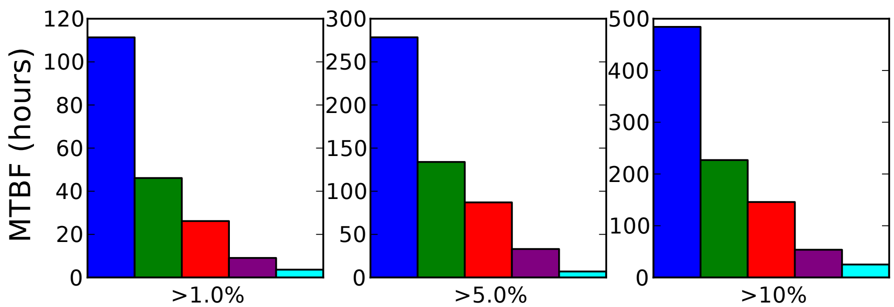
These network issues are repaired, on average, within one to two hours for most ISPs. The main exception is satellite providers—more specifically Viasat— with a MDT (mean downtime), close to 5.5 hours.

In general, most ISPs ranked similarly across both MTBF and MDT, with a few exceptions. For instance, Verizon's fiber service had the highest MTBF, but its periods of downtime were often over 2.5 hours. Frontier's DSL service, on the other hand, had frequent failures, but these periods of failure were relatively short.

4.3.2.2. Effect of access technology. Next, we study the impact of a subscriber's access technology. Figure 4.19a shows a CDF of packet loss rates for each access technology. As expected, we find that fiber services provide the lowest loss rates of all technologies in our dataset with only 0.21% of hours having packet loss rates above 10%. Stated differently, fiber users could expect an hour with 10% packet loss to occur approximately



(a) CDF of the hourly loss rates

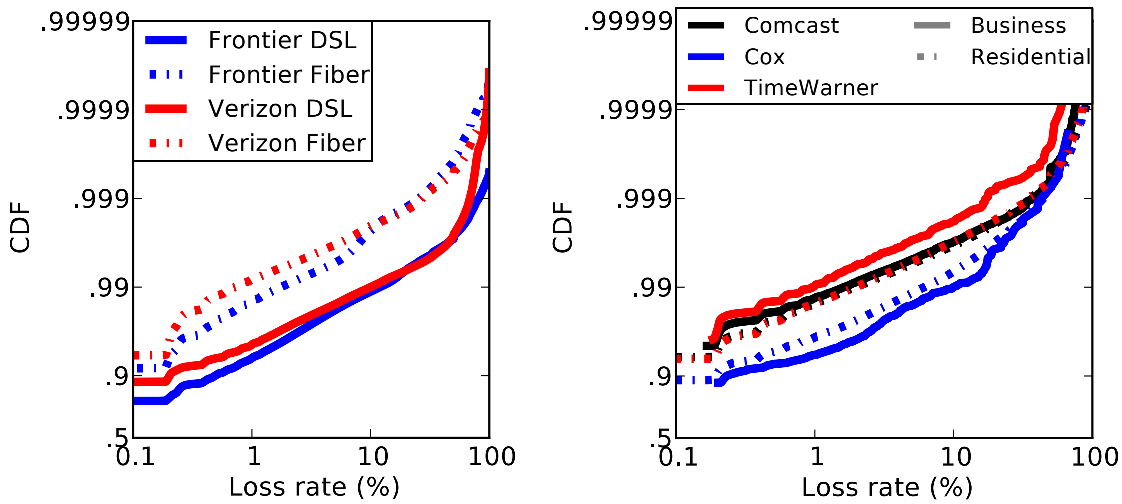


(b) MTBF

Figure 4.19. CDF of hourly loss rates and MTBF for each type of access technology. There is a clear separation between technology for both metrics.

once ever 20 days. Cable and DSL services are next in terms of reliability, with periods of 10% packet loss only appearing 0.44% and 0.68% of the time, respectively. Periods with packet loss rates above 10% were almost a full order of magnitude more frequent for wireless (1.9%) and satellite (4.0%) services.

We compare the average interval between hours with loss above the different loss-rate thresholds, shown in Figure 4.19b. For each threshold, fiber performs significantly better, with cable and DSL again showing relatively similar performance.



(a) Fiber vs. DSL.

(b) Business vs. residential.

Figure 4.20. CDF of the hourly loss rates for subscribers of each service. Technology, rather than provider, is the main determinant of availability and service tiers has little effect.

Other factors that affect the reliability may in fact be related to access technology; for example, network management policies of a particular ISP might be correlated with the ISP's access technology and could hence play a role in determining network reliability. To isolate such effects, we compare the difference in service reliability within the same provider, in the same regions, but for different technologies. Only two providers offered broadband services over more than one access technology: Frontier and Verizon, both of which have DSL and fiber broadband services. Figure 4.20a shows a CDF of the loss rates measured by users of both services. Although there are differences across the two providers, in general, subscribers using same access technology tend to experience similar packet loss rates. Verizon and Frontier DSL customers measured high loss rates (above 10%) during 1.56% and 1.82% of hours, while Verizon and Frontier fiber customers saw high loss rates during 0.33% and 0.53% of hours.

4.3.2.3. Effect of service tier. In addition to offering broadband services over multiple access technologies, a number of ISPs offer different service tiers on the same access technology. For example, Comcast, Cox, and Time Warner all have a "business class" service in addition to their standard residential service. We explored how reliability varies across different service tier offerings within the same provider.

Figure 4.20b shows a CDF of the loss rates reported by users of each provider's residential and business class Internet service. In general, the service class appeared to have little effect on the reliability of a service. The differences in packet loss rates are small compared to the difference between access technologies in the same provider. Comcast business subscribers see about the same loss rates as the residential subscribers, while Time Warner's business subscribers report slightly lower packet loss rates. On the other hand, Cox business subscribers actually report a slightly higher frequency of packet loss when compared to residential subscribers. In particular, there are occasionally anecdotes that providers might be encouraging subscribers to upgrade their service tier by offering degraded service for lower service tiers in a region where they were offering higher service tiers; we did not find evidence of this behavior.

4.3.2.4. Effect of demographics. We also explored the relationship between population demographics and the reliability of Internet service. For this, we combined publicly available data from the 2010 census with the FCC dataset to see how factors such as the fraction of the population living in an urban setting, population density and gross state product per capita relate to network reliability. We found a weak correlation between demographics and reliability, as we describe in more detail below.

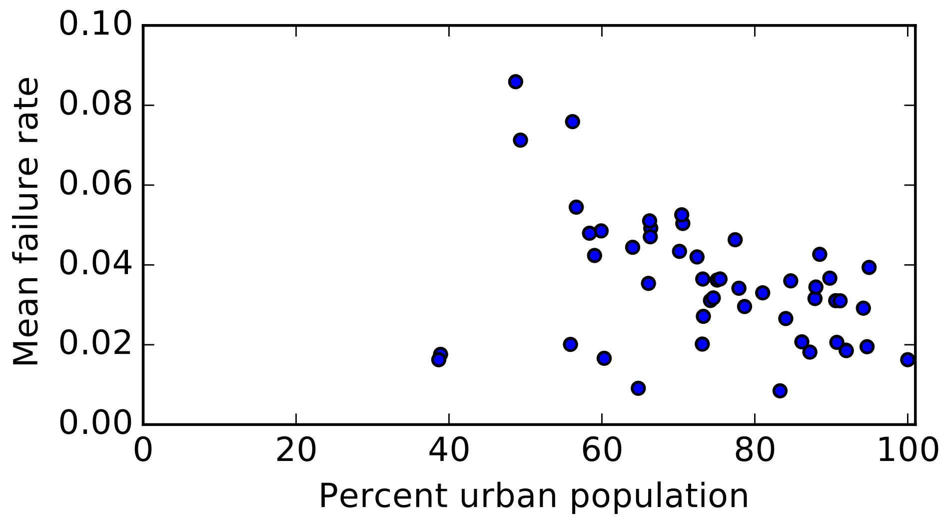


Figure 4.21. Mean failure rate in each state versus the percent of the population living in urban areas ($r = -0.397$). There is a weak negative correlation between failure rate and percentage of urban population.

We first look at service reliability and urban/rural population distributions per state. The US Census Bureau classifies residents into three different categories depending on population. According to the US Census Bureau, "urbanized areas" are defined as areas with more than 50,000 people, "urban clusters" are areas with a population between 2,500 people and 50,000, and "rural" areas encompasses all remaining, non-urban areas [**16**].

Figure 4.21 shows the average failure rate in each state versus the fraction of the population living in an urban setting. Although the two are not strongly correlated, we see a weak to moderate negative correlation between failure rates and percent of urban population ($r = -0.397$). In other words, states with larger rural populations see higher failure rates than more urban states. This correlation results from the fact that subscribers in urban areas have wider access to more reliable access technologies, like fiber and cable.

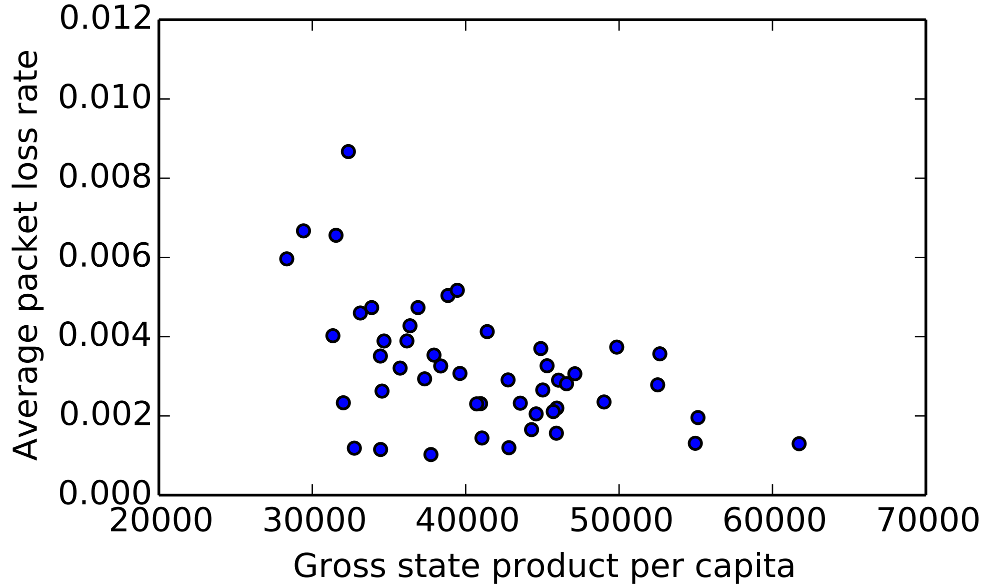


Figure 4.22. Mean loss rate in each state versus the gross state product per capita ($r = -0.358$). There is a weak negative correlation between failure rate and GSP per capita.

Higher urban population densities are sometimes correlated with better economic performance, which in turn may be correlated with better broadband (and other) services. Figure 4.22 plots the average loss in rate in each state versus the gross state product (GSP) per capita. GSP is a measurement of the economic output of a state (or province) and is the sum of all value added by industries within the state, serving as a counterpart to the gross domestic product (GDP).

The relation between population and economic output has been debated for a few hundred years [**60**], Figure 4.22 shows a weak to moderate correlation between failure rates and GSP per capita ($r = -0.358$), As the GSP per capita increases, we are less likely to see areas with higher packet loss rates, suggesting that reliability of broadband services is correlated with local economics.

These findings highlight the importance of considering context when comparing the reliability of service providers, including urbanization and economic levels and also suggest a relationship between population demographics and broadband service reliability (the actual direction of the causal relationship is an interesting area for further study).

4.3.2.5. ISP and DNS reliability. We close our analysis of broadband reliability with a study ISPs' DNS services. Previous work has shown that DNS plays a significant role in determining application-level performance [**71, 97**] and thus users' experience. Additionally, for most broadband users, the effect of a DNS outage is identical to that of a faulty connection.

For DNS measurements, the gateway issues an hourly A record query to both the primary and secondary ISP-configured DNS servers for ten popular websites. For each hostname queried, the router reports whether the DNS query succeeded or failed, the response time and the actual response. Every hour, the FCC/SamKnows gateway performs at least ten queries to the ISP-configured DNS servers. For this analysis, we calculate the fraction of DNS queries that fail during each hour. To ensure that we are measuring DNS availability, we discard measurements during hours when the gateway measures a loss rate above 1%. These are fewer than 3% of hours in our dataset. We classify hours when 50% of DNS queries fail or timeout as periods of DNS unavailability.

Figure 4.23 shows the probability of each provider experiencing one and two DNS server failures during a given hour. We sort providers in ascending order based on the probability that two servers will fail during the same hour.

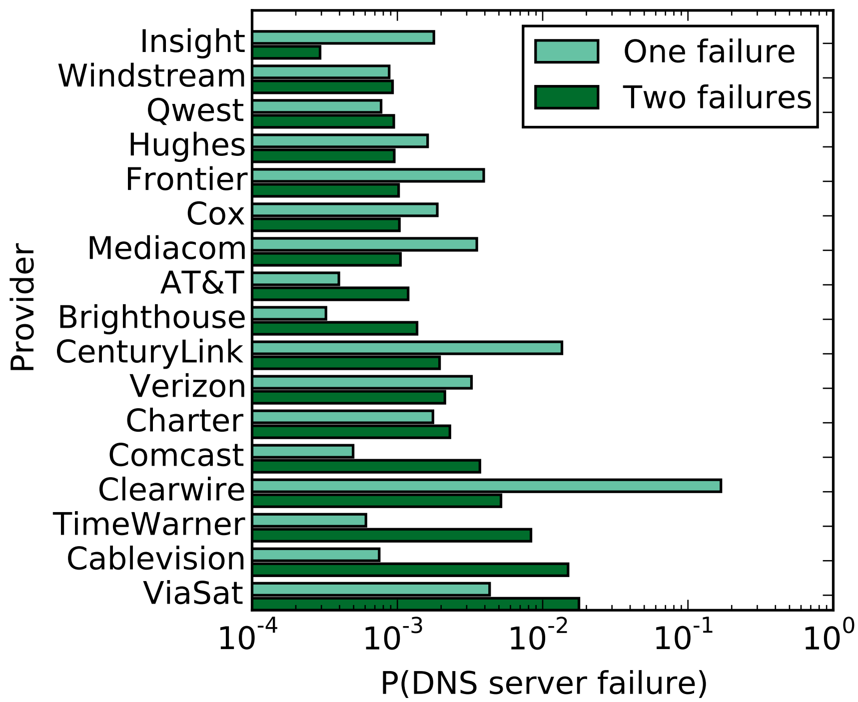


Figure 4.23. Probability that one (or two) DNS servers will be unavailable for each ISP's configured DNS servers. We consider the two cases independently (i.e., "one failure" reflects the event that exactly one server fails to respond to queries).

Surprisingly, we find that many ISPs have a higher probability of two concurrent failures than a single server failing. For example, Comcast's primary and secondary servers are almost an order of magnitude more likely to fail simultaneously than individually.[1]

As one might expect, a reliable access link does not necessarily imply a highly available DNS service. For example, in our analysis of the reliability of access link itself, Insight was in the middle of the pack in terms of availability, offering only one nine of availability (Table 4.2), yet the results in Figure 4.23 show Insight having the lowest probabilities that queries to both DNS servers would fail simultaneously.

[1]One possible explanation is the reliance on anycast DNS. We are exploring this in ongoing work.

CHAPTER 5

Improving service reliability

Our characterization of broadband reliability has shown that even with a conservative definition of failure based on 10% packet loss, current broadband services achieve an average availability no higher than two nines, with an average downtime of 17.8 hours per year. Defining availability to be less than 1% packet loss leaves only a single provider of the 19 ISPs in the FCC dataset with two nines of availability.

Motivated by these findings, as well as the results of our analysis demonstrating that a consistently unreliable service can affect user behavior, we seek to develop an approach to improve service reliability by two orders of magnitude, to the minimum four nines required by the FCC for the public switched telephone service. Our solution should be *easy to deploy*, *transparent to the end user* and should *improve resilience at the network level*.

- **Easy to deploy:** The solution must be low-cost, requiring no significant new infrastructure and the ability to work despite the diversity of devices and home network configurations. It should, ideally, be plug-and-play, requiring little to no manual configuration.
- **Transparent to the end user:** The solution should transparently improve reliability, "stepping in" only during service interruption. This transition should be seamless and not require any action from the user.

- **Improve resilience at the network level:** There have been proposals for improving the access reliability within particular applications, such as Web and DNS (e.g., [**3**, **74**]). Although these solutions are easy to deploy, they are less transparent than a network-level solution could be.

In this section, we present a multihoming-based approach for improving broadband reliability that meets these requirements. Multihoming has become a viable option for many subscribers. The ubiquity of broadband and wireless access points and the increased performance of cell networks means that many subscribers have multiple alternatives with comparable performance for multihoming. In addition, several off-the-shelf residential gateways offer the ability to use a USB-connected modem as a backup link.[1] While the idea of multihoming is not new [**2**, **87**], *we are not aware of previous work showing its potential for improving the reliability of residential broadband.*

We use measurements from Namehelp clients and the FCC's dataset to motivate our approach, showing that the majority of issues occur between the home gateway and the service provider (§5.1), that multihoming can provide the additional two nines of availability we seek (§5.2) and that there are current opportunities for multihoming with neighboring APs (§5.3). We then describe and evaluation our prototype system that allows users to multihome their access by using nearby wireless networks.

5.1. Where network failures occur

The first question to improving service reliability is where in the network the majority of broadband connectivity issues appear. We deployed a network experiment to approximately 6,000 endhosts running Namehelp [**71**] in November and December 2014. For each

[1]E.g., a wireless 3G/4G connection or a second fixed-line modem as in the case of the Asus' RT-AC68U [**4**].

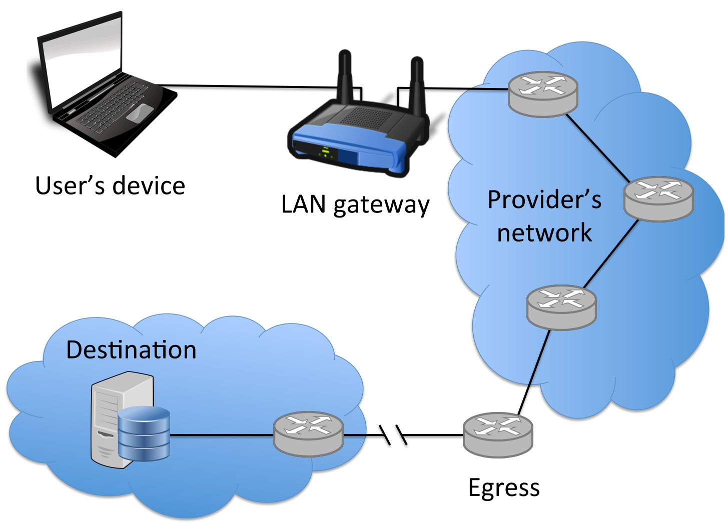


Figure 5.1. Categories used for classifying network reachability.

Farthest reachable point in network	Percent of failures
(1) Reached LAN gateway	68%
(2) Reached provider's network	8%
(3) Left provider's network	24%

Table 5.1. Farthest reachable point in network during a connectivity issue, according to traceroute measurements.

end-host, our experiment ran two network measurements, a ping and a DNS query, at 30-second intervals. We chose to target our measurements to Google's public DNS service (i.e., 8.8.4.4 and 8.8.8.8). For this experiment, we considered this to be a sufficient test of Internet connectivity.

If both ping and DNS query failed to get a response, we immediately launched a traceroute to the target. If the traceroute did not receive a response from the destination, our experiment recorded the loss of connectivity and reported the traceroute results once Internet access had been restored. We used this traceroute data to categorize the issue according to how far into the network the traceroute's probes reached. Figure 5.1

illustrates the categories used in our analysis, while Table 5.1 lists the farthest reachable point in the network during a connectivity issue.

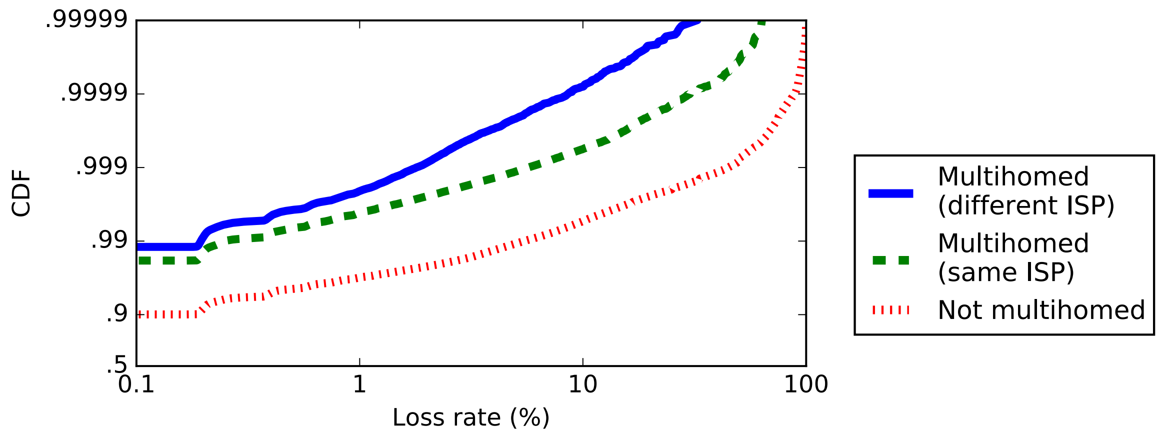
We find that most reliability problems occur between the home gateway and the service provider. During 68% of issues, our probes were able to reach the gateway, but not the provider's network. We cannot determine whether there was a problem with the access link, the subscriber's modem, or the gateway configuration, but in each case, we ensure that nothing had changed with the client's local network configuration (e.g., connected to the same access point and has the same local IP address) and that the probes from the client reached the target server during the previous test. Another 8% of traces were able to reach the provider's network, but were unable to reach a network beyond the provider's. The remaining 24% left the provider's network, but could not reach the destination server.

5.2. Potential benefits of broadband multihoming

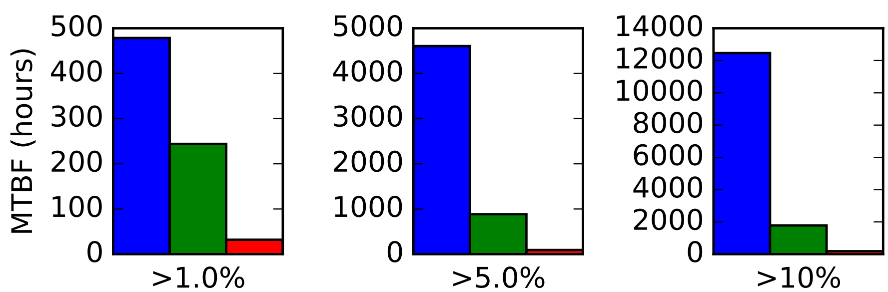
Because the majority of issues occur between the home gateway and the service provider, we posit that a second, backup connection—multihoming—could improve service availability.

To estimate the potential benefits of broadband multihoming for improving service reliability, we use the FCC dataset and group study participants by geographic region based on their Census Bureau block group. A Census block is the smallest geographical unit for which the Census Bureau publishes data, such as socioeconomic and housing details. Blocks are relatively small, typically containing between 600 and 3,000 people.[2] We identify blocks with at least two users online during the same time period. For each pair of users concurrently online in a region, we simulate a multihomed connection by

[2]`https://www.census.gov/geo/reference/gtc/gtc_bg.html`



(a) CDF of the hourly loss rates



(b) MTBF

Figure 5.2. CDF of hourly loss rates and MTBF measured from each gateway and simulated multihomed connection.

identifying the minimum loss rate between the two connections during all overlapping time windows. We distinguished between simulated multihomed connections depending on whether both users subscribed to the same ISP.

Figure 5.2a shows the results of this experiment as a CDF of the loss rates reported for each simulated multihomed connection. As a baseline for comparison, we include the original reported loss rates for the same population of users, labeled "Not multihomed". For both types of simulated multihomed connections (same and different ISP), high packet loss rates are at least an order of magnitude less frequent. Furthermore, the benefits of multihoming with different ISPs as opposed to using the same ISP increase as the loss rate threshold increases. For example, using a 1% threshold as a failure, both scenarios provide two nines of reliability (99.59% when using the same ISP, 99.79% when using different ISPs). However, at 10% loss, multihoming on the same ISP provides only three nines (99.94%), while multihoming on different ISPs provides four nines (99.992%).

Figure 5.2b shows the average interval between periods of high packet loss rates, with thresholds of 1%, 5%, and 10%. Though both types of multihomed connections have a large advantage over a non-multihomed connection, we find that as the loss rate threshold increases, the difference between connections multihomed on the same ISP and connections multihomed on different ISPs increases rapidly; at the 10% packet loss rate threshold, a multihomed connection using different ISPs provides four nines of availability, versus three nines for a connection multihomed on the same provider and about two nines on a non-multihomed connection.

5.3. Using neighboring networks to multihome

There are multiple ways that broadband subscribers could multihome their Internet connection. One possibility would be for users to subscribe to a cellular network service, adding to their existing wireless plan. This approach would be straightforward to implement, as users would only need to add a 4G dongle to their device. However, the relatively high cost per GB of traffic would likely be too expensive for most users, preventing them from using network-intensive services, such as video streaming.

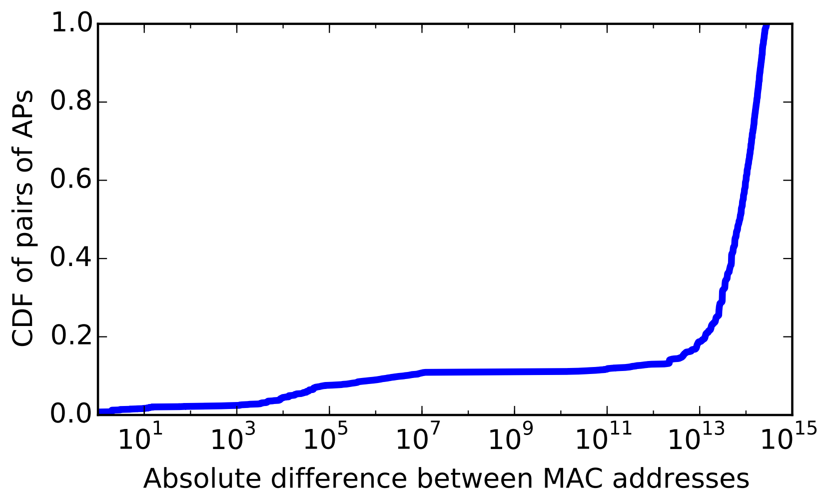


Figure 5.3. Absolute difference between pairs of MAC addresses seen during the same scan of wireless networks.

An alternative, and cheaper, realization of our approach could adopt a cooperative model for multihoming between neighbors either through a volunteer model [**34**, **41**] or a provider's supported community WiFi [**54**].[3]

To show the feasibility of this model, we used end-hosts to measure wireless networks between December 7, 2015 and January 7, 2016. For each user, every hour we recorded their current wireless configuration and scanned for additional wireless networks in the area using OS X's `airport` and Windows' `netsh.exe` commands.

One challenge to estimating the number of available APs is that, in many cases, an individual AP device will host multiple networks (e.g., 2.4 Ghz, 5 Ghz, and/or guest networks) using similar MAC addresses. To avoid overestimating the number of available APs, we used multiple techniques to group common SSIDs that appeared in our wireless scans. We first grouped MAC addresses that were similar to each other (i.e., string

[3]Providers offering such services include AT&T, Comcast, Time Warner, British Telecom (UK) and Orange (France).

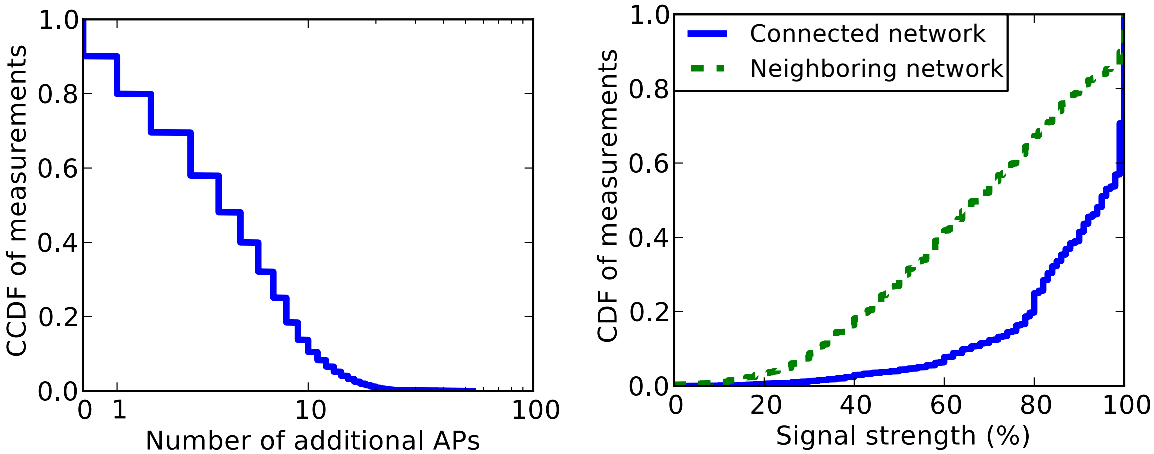
comparisons showed they differed in four or fewer hexadecimal digits or only differed in the 24 least significant bits).

To quantify this, we calculated the absolute difference between each pair of MAC addresses seen in each individual scan, shown in Figure 5.3. Approximately 10.8% of MAC address pairs differ by less than 10^7. However, very few pairs ($< 1\%$) have a difference between 10^7 and 10^{11}, leading us to believe that bottom 10.8% of MAC address pairs are hosted by the same AP. As a conservative estimate, we use 10^9 as a threshold for differentiating between APs. In other words, we consider networks with MAC addresses that differ by less than 10^9 to be hosted by the same AP.

We did notice a few cases where two networks would have nearly identical MAC addresses but differed slightly in the first two bytes of the MAC address. For example, some of Comcast's Xftinity hot spots followed this pattern. Since these were likely hosted by the same AP, we also performed string comparisons between each pair of MAC addresses, grouping together those that differed in four or fewer hexadecimal digits.

We then manually inspected these groups and removed any with an SSID that clearly did not correspond to a gateway, such as network devices and WiFi range extenders (e.g., SSIDs that contained "HP-Print", "Chromecast", or "EXT"). We consider the AP groups remaining as gateway devices.

Figure 5.4a shows the CCDF of the number of additional unique groups seen across all measurements. Since we combine our findings at the end of this section with those of the previous section (§5.2), we only include measurements collected from clients within the US in our analysis. In 90.2% of cases, one or more *additional* wireless APs are available to the client. In approximately 80% of cases, two or more additional APs are available.



(a) Unique additional APs

(b) Signal strength of neighbor's AP

Figure 5.4. Number of additional APs available (a) and signal strengths for the current and strongest alternative AP (b).

These results highlight the potential for using nearby APs to improve service availability via multihoming.

The availability of neighboring AP is a necessary but not sufficient condition; a remaining concern is whether clients would actually be able to connect to these APs. Figure 5.4b shows a CDF of the signal strength percentage of both the AP to which the client is currently connected as well as the signal of the strongest available alternative network ("Neighboring network"). While the signal strengths of the neighboring networks are typically lower than that of the home network, it is still sufficiently strong in most cases, with a signal strength of 40% or higher for 82.7% of measurements.

Last, to estimate the potential improvement in service availability of using a neighboring AP as a backup connection, we infer the ISP of an AP by analyzing its SSIDs. For example, we found a large number of APs advertising SSIDs that clearly identify the provider, such as those starting with "ATT" and "CenturyLink". Similarly, we classified

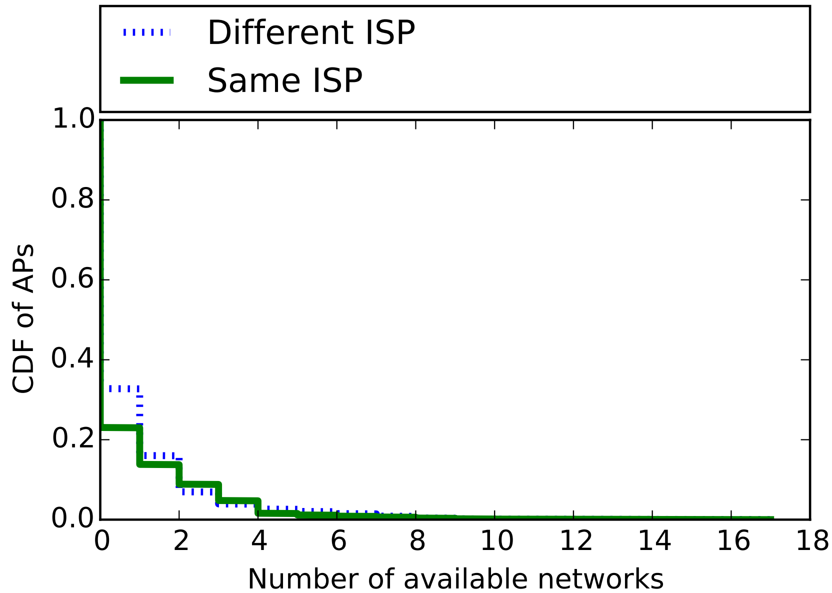


Figure 5.5. CCDF of the number of additional APs that were identified as belonging to a subscriber of the same or a different ISP from the measurement client.

APs that hosted an "xfinitywifi" network in addition to other SSIDs as neighboring networks that belonged to Comcast subscribers. We were able to infer the ISP of at least one neighboring AP in 45% of all scans. Of these, 71% of APs appeared to belong to subscribers of an ISP different from that of the client. Figure 5.5 shows a CCDF of the measurements when we were able to identify the ISP of the additional AP.

These results, combined with our findings in Section 5.2, suggest that if clients used these additional APs as a backup connection, service availability would improve by two nines in at least 32% of cases and by one nine in at least an additional 13% of cases. Since the majority of APs advertised user-defined or manufacturer default SSIDs (e.g., "NETGEAR" or "linksys"), preventing us from inferring a neighbor's access provider,

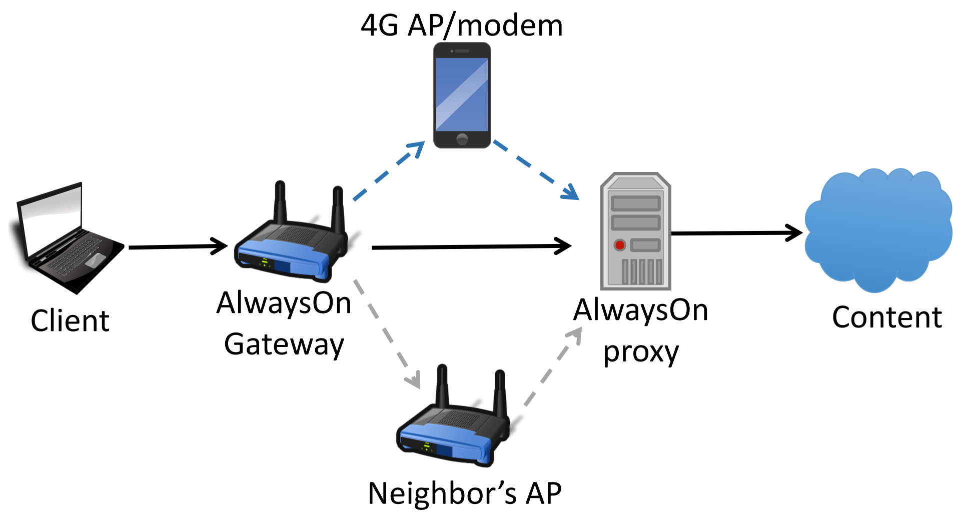


Figure 5.6. Two *AlwaysOn* configurations using a neighboring AP or a 4G hotspot. The black solid line represents a client's normal path while the gray lines represent possible backup routes.

this is a lower bound estimate of the potential of improving service availability through multihoming.

5.4. *AlwaysOn*: a gateway for broadband multihoming

Multihoming a residential broadband connection presents different challenges than traditional multihoming. Independent of the model used, failing over to a neighbor's AP or to a 4G connection, a naïve implementation may interrupt the clients' current open connections and require a restart since the switch will result in a different source IP address for outgoing traffic. A broadband multihoming solution would ideally be able to seamlessly switch between the primary and secondary interface without interrupting the user's open connections.

There are also a number of additional concerns related to usage policies and user privacy that must be addressed. For example, some backup connections (e.g., 4G) may have data caps. A common broadband use like streaming ultra HD videos may have to be restricted over those connections considering their cost. Similarly, neighbors sharing connections with each other may also prefer to impose limits on how they share their connection, for instance, in terms of available bandwidth, time or total traffic volume per month. In locations where there is more than one alternative backup connection, users may want to state their preference in the order of which networks to use based on factors such as the structure of their sharing agreement or the amount of wireless interference on a network's frequency. Finally, there are privacy concerns for both parties when multihoming using a neighbor's network. Users that "borrow" their neighbor's network would, by default, be allowing their neighbor to capture their unencrypted traffic. Conversely, neighbors "loaning" access to their network want their traffic to be private.

The following paragraphs describe how we address these concerns in our prototype service called *AlwaysOn*, followed by a evaluation of the prototype's performance.

5.4.1. AlwaysOn Prototype

AlwaysOn consists of two components – a client running in the gateway and a proxy server deployed as a cloud service. A diagram of this deployment is shown in Figure 5.6. The gray dashed line in this figure represents a backup path via a neighboring AP and another through a 4G hotspot as a backup link, represented by the dotted blue line.

AlwaysOn leverages Multipath TCP (MPTCP) [**72**] to make the transition from primary to secondary links without interrupting clients' open connections. The AlwaysOn Gateway creates an encrypted tunnel to the MPTCP-enabled AlwaysOn Proxy. All traffic from the private LAN is routed via this tunnel. Our current implementation uses an encrypted SOCKS proxy via SSH.[4] Using an encrypted tunnel ensures that a user's traffic is private when routed via the secondary AP. The AlwaysOn gateway also uses a "guest" network that is isolated from the private LAN when sharing its connection. This feature is available on many commodity residential gateways and protects users from exposing their own traffic when sharing network access.

AlwaysOn has multiple settings that can be configured by the user. Some configuration settings, such as traffic policies must be synchronized between the gateway and proxy. For example, while the gateway can throttle outing traffic, traffic shaping policies for incoming traffic must be implemented on the AlwaysOn Proxy. Our current prototype uses `tc` and `iptables` to enforce traffic management policies.

For outgoing traffic, the AlwaysOn gateway can throttle traffic traversing the neighbor's AP as well as traffic on it's guest network. At the proxy, users are given a unique port number to use for their tunnel. Traffic to or from secondary links is then identified by IP address. Using `iptables`, we are able to identify and mark traffic according to the user and whether it is a primary or secondary connection. We then use `tc` to apply the appropriate traffic shaping policy.

In our current implementation these policies must be manually configured both at the AlwaysOn Gateway and proxy. While this is not an issue when users rely on their own backup connection, our current prototype does not have a secure way to enforce policies at the proxy, particularly those that a collaborating neighbor may want. We are currently exploring some alternatives to realize this through a third-party service that accept, encode and enforces such policies on outgoing and incoming traffic.

[4] Other solutions, such as a VPN, are possible.

Deploying AlwaysOn requires an MPTCP-enabled kernel, which current home network field deployments such as BISmark do not currently have; despite these challenges, we are investigating the possibility of deployment on this and other platforms.

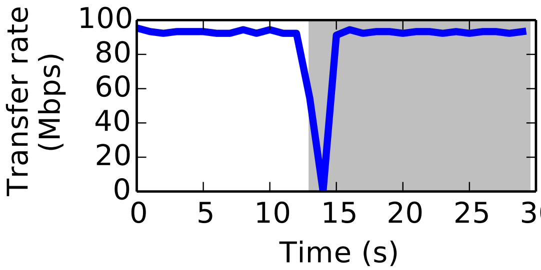
5.4.2. AlwaysOn Evaluation

To evaluate the AlwaysOn prototype, we instantiated the proxy server on a university network. There are two operating modes for AlwaysOn that we want to evaluate: during the failure of the primary link and during normal operation. Considering that the least reliable service offers $\approx$95% availability (or 36 hours of downtime per month), routing traffic via the AlwaysOn proxy should have minimal impact on performance during normal operation. In addition, to limit the impact of outages on user quality of experience, AlwaysOn should respond quickly and route traffic via the backup link.

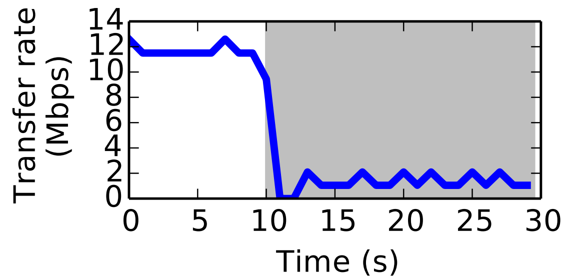
In our experiments, we ran `iperf` for 30 seconds from a client behind the AlwaysOn Gateway, recording the `iperf`'s measured throughput rate each second. We emulated different outages, represented in the plots by time periods highlighted in gray.

We set up three different scenarios for our evaluation, shown in Figure 5.7. In the first scenario (Fig. 5.7a), we used a university network connection of 100 Mbps as the primary connection and another identical university connection as the secondary. In the second scenario (Fig. 5.7b), we used a Comcast 75 Mbps service as the primary connection and a 3 Mbps AT&T service as the secondary connection

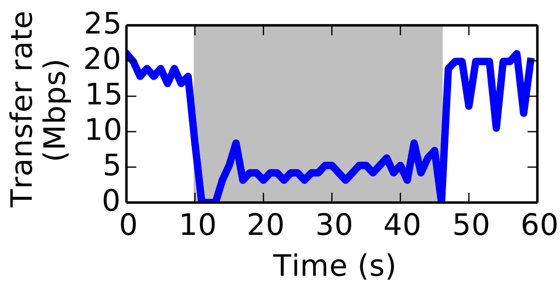
In the third scenario, shown in Fig. 5.7c, we used an RCN 150 Mbps service as the primary, and a Verizon Wireless 4G LTE hotspot as the secondary connection. In this case, AlwaysOn re-enabled the primary connection after approximately 30 seconds. Once



(a) Univ. / Univ.
(100 Mbps / 100 Mbps)



(b) Comcast / ATT
(75 Mbps / 3 Mbps)

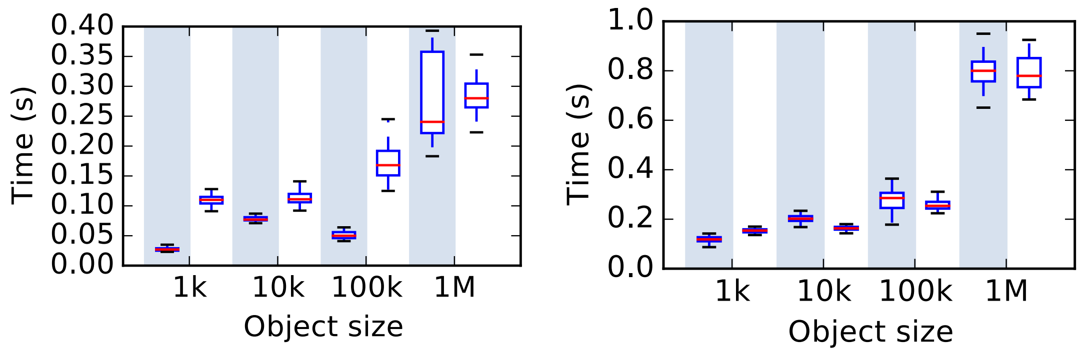


(c) RCN / VZW
(150 Mbps / 4G LTE)

Figure 5.7. Throughput rates measured using `iperf` using *AlwaysOn* in two different settings. Each figure lists the service providers used and the speed of the primary and secondary connections (respectively). The gray highlighted section represents the time during which the we simulated an outage on the primary link.

the change primary connection was reestablished, AlwaysOn switches traffic back to the RCN connection.

In each case, AlwaysOn can recover relatively quickly once it realizes the primary link is no longer working, and does not require the connection to be reestablished. We also ran `iperf` in each case directly between the same client and the AlwaysOn sever (bypassing the AlwaysOn Gateway) and consistently measured similar throughput rates. The relatively slow performance compared to the access link speed in Fig. 5.7b and Fig. 5.7c is likely



(a) RCN
(150 Mbps)

(b) VZW
(4G LTE)

Figure 5.8. Box and whisker plots showing the time to fetch objects hosted by Akamai while using (highlighted) and not using the AlwaysOn proxy. for RCN and Verizon Wireless (VZW). Each box and whisker represents the median, interquartile range (IQR), and 1.5 IQR for each experiment configuration.

due to other limiting factors such as end-to-end latency, congestion on the path, and only using a single TCP connection.

To see how our AlwaysOn proxy affected network performance, we also measured the time to fetch objects hosted on Akamai's CDN, both when using and not using the proxy. For this test, we downloaded files of varying sizes (1 kB, 10 kB, 100 kB, and 1 MB) 100 times. The box plots shown in Figure 5.8 summarize the distribution of download times for objects of each size while using the RCN (Fig. 5.8a) and Verizon Wireless (Fig. 5.8b) connections. The highlighted box plots show the fetch times while using the proxy for each respective file size.

Overall, performance was similar across the two settings. In many cases, we found that download times actually improved while using the AlwaysOn proxy.

CHAPTER 6

Adopting broadband SLAs

In today's broadband markets, service plans are typically described in terms of their maximum download throughput rate, often advertised as "up to X Mbps". This advertised capacity, along with the associated monthly cost, are the two primary, and many times only, pieces of information available to consumers when comparing service providers. Such "constrained" service agreements place services using technologies as diverse as fiber, DSL, WiMAX or satellite on nearly equal grounds, and leave consumers without clear expectations given that, strictly speaking, any speed slower than X meets guarantee.

We argue that as Internet users and the devices they use become more dependent on connectivity and consistent performance, broadband will move from a loosely regulated luxury to a key utility,[1] ushering in a growing demand for more encompassing SLAs. Most common utilities, such as electricity and water, have very well defined SLAs.

The adoption of broadband SLAs will benefit all players – service providers, customers, and regulators. From the ISP's perspective, contracts with SLAs could allow them to better differentiate their retail services and fine-tune their contracts to the needs of particular classes of customers (e.g., a service for gamers).[2] For customers, SLAs could significantly simplify the process of comparing services offered by different providers, allowing customers to make more informed decisions. This would in turn provide more transparent

[1]In Feb. 2015, the US FCC voted to regulate broadband services as a public utility.
[2]Some ISPs already try this if in coarser terms; e.g., Comcast's "What type of Internet connection is right for you?" `http://www.xfinity.com/resources/internet-connections.html`

competition in the market and potentially lower prices. Similarly, for policymakers and regulators, SLAs would improve their ability to gauge the broadband infrastructure across communities.

Despite these potential benefits, there are several challenges in defining SLAs for broadband services, from identifying metrics and defining the appropriate SLA structures to engineering compliance monitoring. SLAs must be designed so that they can be accurately and efficiently monitored and add value to consumers and providers, while limiting the risk of non-compliance. In this paper, we make a case for broadband SLAs and follow a data-driven approach to explore some of these key challenges.

We expect broadband SLAs to be specified, as other network SLAs, in terms of transport-level performance assurances using Quality of Service metrics such as bandwidth, packet loss, delay and availability. While the relationship between such QoS metrics and users' experience with different applications is a topic of ongoing research (e.g., [**19**, **66**]), existing approaches rely on such QoS metrics as input to application specific models of QoE estimation.

An SLA could be seen as an insurance policy against the risk of not receiving the contracted level of service. Consequently, SLA-enhanced services would come with a price-tag that depends on the structure of the SLA and degree of risk involved in the delivering the desired levels of service. Using data from the FCC dataset, we study the design space of broadband SLAs and demonstrate that certain SLAs could be offered almost immediately with minimal impact on retail prices and network investment.

In this chapter, we present the following contributions:

- We analyze different QoS metrics for use in SLA and define a set of broadband SLAs. We find that, across all ISPs and access technologies, bandwidth is the most consistent of the three studied performance metrics (bandwidth, latency and loss rate).
- We evaluate the relationship between access technology, the SLA structure and the cost of having SLAs. We show that many of the studied ISPs could offer moderate SLAs with minimal impact on their existing business, but that SLAs with stringent constraints are much harder to deliver across the whole user-base.
- We examine if SLAs could be tailored for each end-user individually and show that ISPs (or third parties) could accurately (with accuracy comparable to that in the car or credit insurance industry) infer the risk of offering SLA to individual customers and price the SLA service accordingly.

6.1. Metrics for broadband SLAs

An SLA is a contract between a service provider and its customers that specifies what services the provider will support and what penalties it will pay upon violations. A meaningful SLA should (*i*) capture the needs of consumers, (*ii*) be feasible to support by most service providers today and (*iii*) be expressed in measurable terms that can be validated by both consumers and services providers.

To understand the need of broadband consumers, we must consider the requirements of commonly used network applications. Clearly, one would not expect that "broadband consumers" would be a homogeneous class in either the type of applications they value most or their expectations. For some consumers, being able to browse the web or read

SLA	Throughput (% of service)	Latency	Packet loss
A	> 90%	< 50 ms	< 1%
B	> 50%	< 150 ms	< 5%
C	> 10%	< 250 ms	< 10%

Table 6.1. Three examples of possible broadband SLAs.

their email may be sufficient, and paying for a higher guaranteed throughput would not be a priority. Others may have higher performance requirements, wanting a connection that reliably allows them to stream HD video content or play online games with strict timing requirements.

Driven by these observations and the existing literature on the needs of different application classes (e.g., [**20**, **21**, **98**]), we drafted three potential broadband SLAs that cover a wide range of user requirements. Note that these are mere examples of possible SLAs, focused on the points relevant to our argument, and ignoring the specifics of a practical SLA, such as the form of reporting quality of service violations, the procedure to be invoked in case of violations or the exact cost model of violations.

Our basic SLAs (see Table 6.1) are stated in terms of throughput, latency and packet loss. Considering that the subscription capacity varies across users, we structure SLAs in terms of the percentage of subscription speed that is available. For latency and packet loss, we adopt a simple "below-threshold" model. *SLA A* represents a service that could be able to fit the demands of users with very strict performance requirements, while *SLA C* represents one that should be able to support simple applications, such as browsing the web. *SLA B* matches the middle-of-the-road services, capable of supporting most applications but with less than perfect performance for network-intensive applications.

Although they are somewhat arbitrary, the thresholds we use for our sample SLAs – from fractions of throughput to latency and loss rate – are based on existing literature and earlier studies of broadband services. We selected 10% of the service capacity as a bottom threshold, given that the vast majority of users had a connection much faster than 1 Mbps and that 100 kbps can support basic browsing and email requirements. We chose 50% as a threshold because of a study published by Ofcom in the UK found that in 2010, users were often only getting, on average, about half (46%) of the their advertised speed [**67**]. For our highest SLA we opted for 90% as a threshold to highlight providers that consistently provide capacities close to their subscription speeds.

In terms of packet loss, previous works have shown that rates above 1% can have a negative impact on users' QoE while using gaming applications [**21**]. High loss rates can also affect other common services such as audio and video calls [**20**]. Xu et al. [**98**] showed that loss rates above 4% can significantly degrade iChat video calls and rates larger than 10% result in a sharp increase in packet retransmissions.

Latencies above 250 ms result in a large increase in page loading time [**10**] and would likely have a significant negative impact on QoE while using telephony applications. End-to-end latencies below approximately 150 ms should be sufficient for Skype calls [**86**]. Our low threshold for *SLA A* was based on previous work showing that an increase of just 10 ms can yield an increase in page loading delays by hundreds of milliseconds [**94**].

6.2. Supporting SLAs today

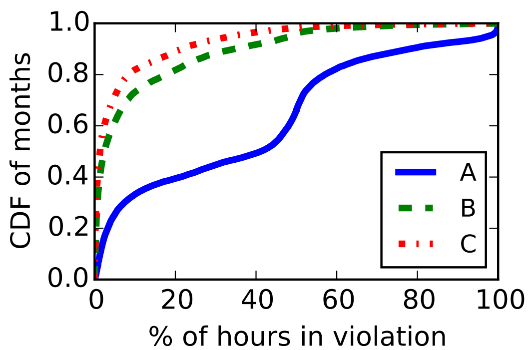
In Section 6.1, we defined SLAs in measurable terms with thresholds that would be meaningful to users' Quality of Experience. We now explore how effectively today's ISPs could meet our proposed set of SLAs.

There are a number of ways that a broadband SLA could be structured in terms of how users are compensated for periods of poor performance. As an example, we looked at how some broadband ISPs structure the agreements that they offered to businesses. In the case of Comcast [**26**], business class subscribes are compensated once the network become unavailable for more than four hours in a single month. For each hour of downtime after the first four, customers are reimbursed 1/30 of the monthly subscription price[3]. We believe that general broadband service plans could have a similar structure. For example, the SLA could state the network may be unavailable for up to two hours per day (or about 8.33% of hours in a month). This would allow ISPs to schedule downtime for maintenance and provide a guarantee for subscribers that their service will not be down for days at a time (or that they will be compensated if it is).

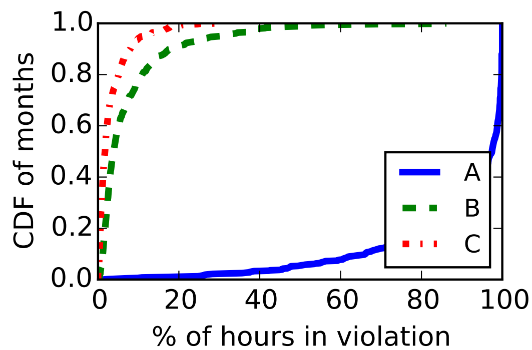
However, our focus in this paper is not on the structure of compensation for SLA violations. Instead, we look at how well the ISPs in the FCC's dataset are able to meet the SLAs defined in Section 6.1, and whether it would be at all feasible to provide guarantees of service.

Figure 6.1 summarizes the total number of SLA violations per month for four sample ISPs. AT&T, shown in Figure 6.1a struggles to meet the requirements of *SLA A* but is

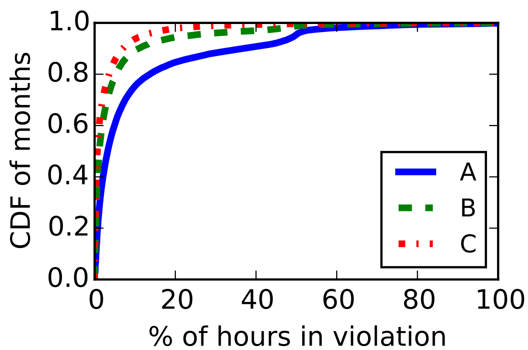
[3]This effectively means that if the service was 'unavailable' for 34 hours in a month (approximately 5% of the month) the user gets the monthly subscription for free.



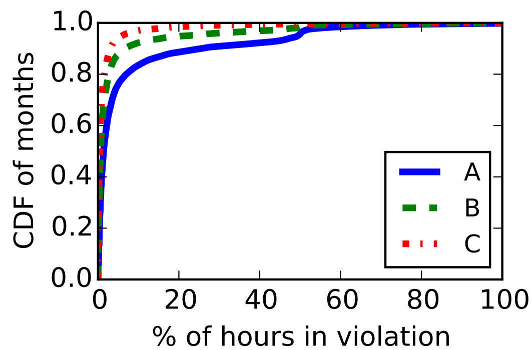
(a) AT&T (DSL)



(b) Clearwire



(c) Comcast



(d) Verizon (fiber)

Figure 6.1. CDF of the percent of hours per month in violation of a given SLA for four broadband services.

able to meet *SLA B* during 90% of hours per month for 73% of users and meets *SLA C* during 90% of hours for 82% of users.

The wireless provider in our dataset, Clearwire (Fig. 6.1b), face difficulties in meeting *SLA A*, as the average latencies were almost always higher than 50 ms. Interestingly, Clearwire actually did a better job of meeting *SLA C* than AT&T, with 94% of users meeting *SLA C* performance during at least 90% of hours in a month.

Both Comcast and Verizon's fiber service did a relatively good job of meeting the requirements of all three SLAs. Comcast was able to meet *SLA A* during 90% of hours in a month for 75% of users while Verizon was able to do the same for 83% of fiber subscribers. Both were able to provide both *SLA B* and *C* during 90% of hours for at least 90% of users.

To summarize our findings in this section, moderate SLAs (those which require SLA compliance up to 90% of time) are very feasible nowadays and could be offered by many ISPs with minimal effect on their current business. However, stricter SLAs (those which require SLA compliance 99% of the time or more) would be much more challenging to offer across the whole user base. In the following section, we examine how difficult would it be to assess the individual risk of breaking SLA which a central challenge in offering personalized SLA offerings.

6.3. Personalized SLAs

As we noted in the previous sections, SLA can be seen as an insurance policy against poor broadband experience (which may in turn also have financial consequences such as cash payback in case of broken SLA). In this section we study if SLAs could be tailored for each end-user individually. The key question is whether the SLA provider would be able to infer the likelihood of delivering the SLA. Certain user characteristics may be correlated with the quality of service the user receives and hence the SLA provider may choose to price the service (premium in insurance terms) according to the risk of not delivering promised SLA. With a good understanding of how likely it is to break the SLA the insurer (either a third party or the broadband provider itself) can fine tune the SLA parameters and the premium (in $ per month) in order to improve the user satisfaction of the service and ensure the profitability of the SLA service.

We train a simple model to examine the predictability of the service of individual subscribers complying to an SLA based on several simple user features available to us: (1) access technology, (2) base latency (to the nearest measurement server), (3) aggregate usage (in bytes per month) and (4) city population (a proxy of urban/rural residence). More advanced models, using a range of additional demographic and technological features, would likely improve the prediction accuracy, yet such analysis is out of scope of this short study and is left for future work.

We use supervised learning for estimating the likelihood of breaking the SLA, for three SLA types described in Figure 6.1 with 95% time thresholdḟootnoteMeaning the users' performance complies with the SLA 95% of the time during the month. It is basically a binary classification task, where we use four user features described above to predict whether the user complies with SLA or not. The features are extracted on 4038 active users in October and November 2014. The categorical feature describing access technology is projected to a binary vector of length 4 encoding the access technology of every user. We experimented with several classification methods. Each showed comparable performance so we report the results from only one of them: random forests. The hyper-parameters were optimized using a grid-search over a validation set extracted from the training set. We use fourfold cross validation to predict the chance of breaking SLA. The features are extracted in October 2014 and the (binary) SLA compliance is extracted for November 2014.

We use a standard metric for measuring the performance of the binary classifiers: Area Under Curve Receiver Operating Characteristic (AUCROC) [**43**]. The ROC curve as well as the AUCROC are reported in Figure 6.2 for the three SLAs from Table 6.1. The

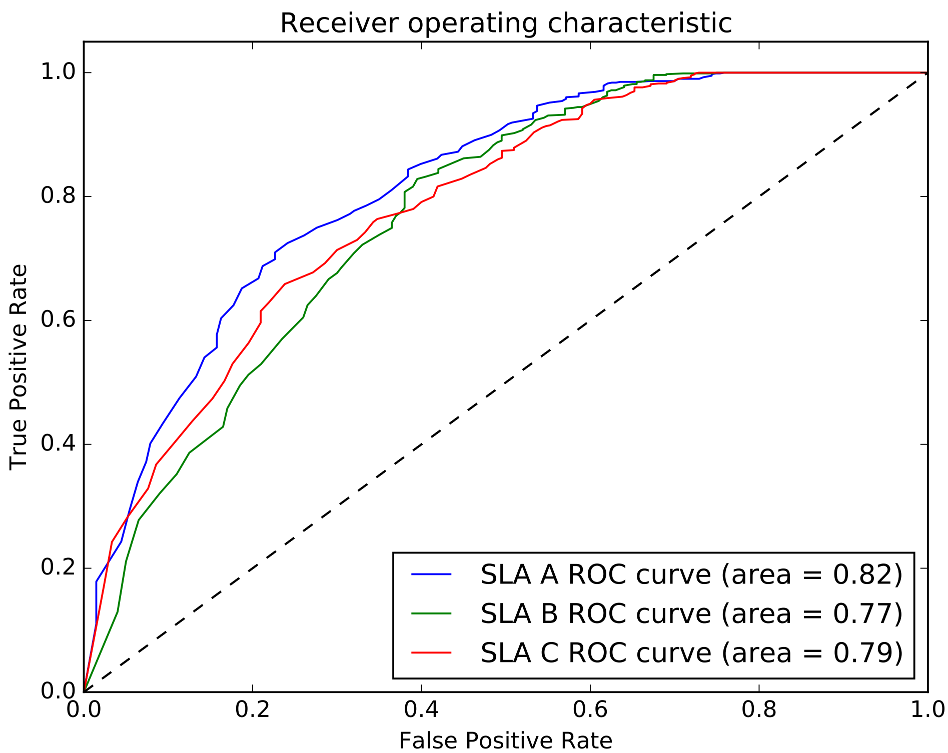


Figure 6.2. Area under ROC for Random Forest classifiers.

AUCROC for all three SLAs: A, B and C, is similar and is around 0.8. Such AUCROC is comparable to the precision of classifiers build on demographic user info in other insurance products for car insurance and credit ratings [**76**]. This accuracy of prediction for SLA compliance suggests that it would be possible to offer personalized SLAs with a price which accurately matches the likelihood of breaking the SLA.

6.4. Discussion of SLAs and contributions

Recent efforts [**11**, **82**, **84**, **93**] have attempted to address the lack of detailed evaluations of ISPs. Annual reports published by the FCC in the US and Ofcom in the UK have studied whether or not ISPs are providing the capacities promised to users. The recent

Net Neutrality ruling from the FCC [**38**, **39**] discussed the issue of how service plans are described to subscribers; one part of the ruling states that ISPs must disclose reasonable estimates of performance metrics, including both latency and packet loss. Unfortunately, what exactly is a "reasonable" estimate of these metrics is somewhat unclear. Additionally, providing the estimates alone does not offer any protection for consumers that may experience seriously degraded performance.

Contrast the state of broadband service agreements with those of utilities such as electricity or water. In the developed world, most common utilities have very well defined set of performance guarantees and, being very mature services, such guarantees are strongly embedded into the service and rarely, if ever, broken. In residential electricity delivery, voltage must be within 10% of the nominal voltage, and the frequency within 5% of the nominal value [**49**]. This is important as many home appliances rely on the strict voltage/frequency for normal operation and could even break when supplied with voltage/frequency that is far from expected values.

This work points to a number of interesting research directions that are crucial for implementing broadband SLAs. For example, perhaps the largest roadblock to adoption of broadband SLAs is the lack of infrastructure for monitoring performance and reporting SLA violations. One potential avenue to explore would be the deployment of a system, such as SLAM [**89**] on home gateways or modems, that could monitor SLA compliance as those presented in this work. Furthermore, reliable processes for the automatic generation and filing of SLA violation reports and reporting, to both the subscriber and the ISP, could be another interesting research direction.

Recent peering disputes between content providers and broadband access providers [**46**, **88**] highlight the importance of measuring congestion on a provider's peering links and its potential impact on performance. Poor quality of experience while streaming via Netflix or making Skype calls would not be captured by the measurements used in this paper if caused by congestion at the edge of the provider's network.

Another aspect we have not explored is the design of SLAs that both fit what a user's needs and what they can afford, an area that has been explored in other works [**9**]. For example, an SLA that promises to provide lower latency, from 25 ms to 15 ms, could be extremely expensive for the ISP, but provide little value to the subscriber. Additionally, the availability of other services that are typically hosted by the ISP, such as DNS or email, may be more important to some users than a guaranteed throughput rate. Other works suggest that consumers could benefit from improvements in how service offerings are described to customers [**95**]. In our case, the benefits of SLAs would need to be described in terms that non-technical consumers could understand.

In this chapter, we explored the possibility of implementing broadband SLAs and demonstrated that certain SLAs could be offered almost immediately with small impact on the retail prices and network investment. This work is partially motivated by the FCC's recent classification of broadband as a utility. We believe that this is a natural course for broadband Internet, as it progresses from being a luxury to a key utility and, in some countries, considered a basic human right. We believe that the adoption of broadband SLAs will eventually help improve the efficiency of broadband markets, as ISPs will be able to differentiate their service from competitors, and will also help customers understand the costs of providing more consistent and reliable broadband access.

CHAPTER 7

Conclusion

7.1. Related Work

Broadband analysis has recently attracted much attention from the research community and the general public given its important business and policy implications. A number of efforts have focused on characterizing the availability and performance of broadband services around the world [**10**, **11**, **17**, **31**, **55**, **81**, **90**, **93**]. The focus of our work is on exploring broadband services in their broader context, evaluating the complex interplay between broadband service characteristics, their market features and user demand.

Different aspects of the complex interplay between user behavior, network services and operation has been explore in previous work. Some recent studies have examined the relationship between user behavior, network services and the providers. In Dobrian et al. [**32**] the authors show that poor connection quality can have a negative impact on a user's quality of experience. Blackburn et al. [**12**] study how user behavior affects the economics of cellular operators. Chetty et al. [**22**] perform a user study to understand the effects of usage caps on broadband use. Other efforts have explored additional factors that may influence service demand, including the weather [**18**], service capacity [**99**] and the type of region [**23**].

The difficulty or outright impossibility of conducting controlled, randomized experiments of user behavior at Internet scale has been pointed out before. In his SIGCOMM

2011 Award presentation, Vern Paxson pointed to this issue and suggested the use of natural experiments to explore potential causal relationships with observational data. In a recent paper, Krishnan and Sitaraman [**56**] explore the use of related quasi-experimental design (QED) to evaluate the impact of video stream quality on viewer behavior and Oktay et al. [**69**] relies on it for causal analysis of user behavior in social media.

Previous work on broadband networks has focused on characterizing services in terms of performance (e.g., link capacity and latency) from a range of platforms and vantage points, including customized home gateways [**81**,**93**], applications in end-user devices [**11**, **30**,**71**,**84**], Web-based tests [**17**,**55**,**79**], and well-provisioned measurement nodes outside the access networks [**31**]. We leverage longitudinal data collected by two of these efforts, the FCC's MBA initiative and Namehelp [**71**], to study broadband reliability.

Reliability of phone networks. A number of efforts have tackled reliability characterization in other contexts, such as telephone and cellular networks. Thanks in part to the pioneer work at Bell Labs [**28**] by the end of the 20th century, public switched telephone networks (PSTN) had become so reliable that AT&T expected no more than two hours of failure over a 40-year period [**57**]. Today the FCC requires that PTSN providers document and report outages affect more than 30,000 users or last longer than 30 minutes [**35**] which corresponds with at least four nines of availability.

Effect of network factors on user behavior. Recent work has explored the effect of network factors on user experience with applications, including VoIP [**20**], Web [**5**] and Internet video [**6**]. Rather than deriving a model for user experience based on multiple factors, our work focuses on the effect of reliability on user demand. Others have started to explore the use of alternative experimental designs to evaluate user experience. Krishnan et al. [**56**] apply quasi-experimental design to evaluate the effect of video stream quality on viewer behavior, Oktay et al. [**69**] relies on it for causal analysis of user behavior in social media. Bischof et al. [**9**] explores the effect of contextual factors such as price and competition on user demand. We apply similar methods to understand the effect of service reliability on user behavior.

Reliability of broadband providers. Baltrūnas et al. [**7**] presented a study of the reliability of four mobile broadband providers in Norway using the Nornet Edge dedicated infrastructure. This work illustrates the value of end-to-end measurements to identify failures and performance events not always captured by the operators' monitoring systems. Broadband reliability has received little attention until recently. Lehr et al. [**58**] discuss some of the challenges of characterizing reliability and their economic and policy implications and identify three different ways in which the "reliability" of broadband services can be measured: (1) the reliability of the service itself; (2) the reliability of network services offered by the ISP (e.g., DNS); and (3) the consistency of the service's performance. We focus on characterizing reliability in terms of the former two categories and leave the latter as future work.

Multihomed access networks. Beyond improvements in access link technology, one way to enhance the reliability of access networks is through redundancy. Gummadi et al. [**44**] propose a detouring approach to recover from Internet path failures. Andersen et al. improved web availability with their system, MONET, an overlay network of multihomed proxy servers [**3**]. We have seen the recent introduction of consumer-grade residential gateways that support a second WAN connection (such as a 3G or 4G modem) [**4**] and some work exploring the performance benefits of the on-loading of broadband traffic using a 3G connection [**80**]. Other works have looked at using MPTCP to improve performance by bonding DSL and LTE connections [**13**, **77**].

7.2. Summary

The findings reported in this dissertation represent an import step towards understanding how user behavior, and the market features that shape it, affect broadband networks and the Internet at large. These findings should provide valuable insight to the research community, network operators and policy makers. For policy makers, there is a growing consensus that broadband access should be treated as a fundamental right and several efforts around the world are aimed at closing the digital gap. We believe that understanding digital inequality requires us to place broadband access in a broader context [**62**]. This work, to the best of our knowledge, is the first attempt to that end.

In our longitudinal study of usage between 2011 and 2013, we found that subscribers' demand remained relatively constant in a particular service class, despite the fourfold increase in global IP traffic over the past five years. Thus, we believe the growth in broadband traffic comes from a combination of increased service capacities and a rapidly increasing number of broadband subscribers, rather than higher demand at users' existing service levels.

We find a strong correlation between service capacity and user demand, despite the fact that users rarely fully utilize their links. We used two study designs to infer causality between capacity and usage by studying how individual users change behavior when switching to faster services and by comparing demand between users with different capacities that are otherwise similar. Their relationship also follows a law of diminishing returns; in both cases, we observed relatively lower increases in demand at higher capacities. This trend is particularly noticeable at approximately 10 Mbps, where usage begins to plateau for many users.

This suggest that as service capacities continue to increase, network operators can plan on higher over-provisioning rates. We did observe, however, larger increases in demand when including BitTorrent traffic in our analysis. Beyond capacity, we also showed the impact that the quality of a connection, in terms of latency and packet loss rates, has on user demand. For instance, we note that very long latencies (above 500 ms) and high packet loss rates (starting at 0.1%) clearly result in lower network usage. We speculate that the relationship between capacity, quality and demand will evolve with technology improvements and new applications with greater bandwidth requirements become widely available (e.g. 4k video streaming or telemedicine).

We examined how the price of broadband access affect user demand by comparing the behavior of users with similar broadband services located in different markets. We found that users in markets where broadband connections or additional capacity was more expensive, were more likely to impose higher bandwidth demands on their service than subscribers of comparable services in less expensive markets. For policy makers, this would imply that a focus on wider access to a medium, high-quality capacity service (around 10 Mbps) may have a more significant impact than a focus on increased service capacity. For operators, these trends may suggest a possible role for service pricing in network planning.

We empirically demonstrated the importance of broadband reliability on users' quality of experience. We presented an approach for broadband reliability characterization using data collected by common national efforts to study broadband and discussed key findings from applying it to four-year dataset collected through the FCC's MBA program. Motivated by our findings on both the importance of reliability and the current reliability of broadband services, we presented the design and evaluation of a practical approach to improving broadband reliability through multihoming.

We explored the possibility of implementing broadband SLAs and demonstrated that certain SLAs could be offered almost immediately with small impact on the retail prices and network investment. This work is partially motivated by the FCC's recent classification of broadband as a utility. We believe that this is a natural course for broadband Internet, as it progresses from being a luxury to a key utility and, in some countries, considered a basic human right. We believe that the adoption of broadband SLAs will eventually help improve the efficiency of broadband markets, as ISPs will be able to differentiate their service from competitors, and will also help customers understand the costs of providing more consistent and reliable broadband access.

References

[1] Finland makes broadband a 'legal right'. http://www.bbc.com/news/10461048, July 2010.

[2] Aditya Akella, Bruce Maggs, Srinivasan Seshan, Anees Shaikh, and Ramesh Sitaraman. A measurement-based analysis of multihoming. In *Proc. of ACM SIGCOMM*, 2003.

[3] David G. Andersen, Hari Balakrishnan, M. Frans Kaashoe, and Rohit Rao. Improving web availability for clients with MONET. In *Proc. of USENIX NSDI*, 2005.

[4] Asus. RT-AC68U. http://www.asus.com/Networking/RTAC68U/.

[5] Athula Balachandran, Vannet Aggarwal, Emir Halepovic, Jeffrey Pang, Srinivasan Seshan, Shobha Venkataraman, and He Yan. Modeling web quality-of-experience on cellular networks. 2014.

[6] Athula Balachandran, Vyas Sekar, Aditya Akella, Srinivasan Seshan, Ion Stoica, and Hui Zhang. Developing a predictive model of quality of experience for Internet video. In *Proc. of ACM SIGCOMM*, 2013.

[7] Džiugas Baltrūnas, Ahmed Emokashfi, and Amud Kvalbein. Measuring the reliability of mobile broadband networks. In *Proc. of IMC*, 2014.

[8] Zachary S. Bischof and Fabián E. Bustamante. A time for reliability - the growing importance of being always on. In *Proc. of ACM SIGCOMM*, 2014. Poster.

[9] Zachary S. Bischof, Fabián E. Bustamante, and Rade Stanojevic. Need, want, can afford – broadband markets and the behavior of users. In *Proc. of IMC*, November 2014.

[10] Zachary S. Bischof, John S. Otto, and Fabian E. Bustamante. Up, down and around the stack: ISP characterization from network intensive applications. In *Proc. of W-MUST*, 2012.

[11] Zachary S. Bischof, John S. Otto, Mario A. Sánchez, John P. Rula, David R. Choffnes, and Fabián E. Bustamante. Crowdsourcing ISP characterization to the network edge. In *Proc. of W-MUST*, 2011.

[12] Jeremy Blackburn, Rade Stanojevic, Vijay Erramilli, Adriana Iamnitchi, and Konstantina Papagiannaki. Last call for the buffet: economics of cellular networks. In *Proc. of MobiCom*, pages 111–122. ACM, 2013.

[13] M. Boucadair and et al. An MPTCP Option for Network-Assisted MPTCP Deployments: Plain Transport Mode. `https://tools.ietf.org/html/draft-boucadair-mptcp-plain-mode-08`.

[14] Broadband Commission. The state of broadband 2012: Achieving digital inclusion for all. `http://www.itu.int/net/pressoffice/press_releases/2012/61.aspx`.

[15] Broadband Commission. The state of broadband 2014: broadband for all. http://www.broadbandcommission.org/Documents/reports/bb-annualreport2014.pdf.

[16] United States Census Bureau. 2010 census urban and rural classification and urban area criteria, 2010. https://www.census.gov/geo/reference/ua/urban-rural-2010.html.

[17] Igor Canadi, Paul Barford, and Joel Sommers. Revisiting broadband performance. In *Proc. of IMC*, 2012.

[18] Juan Camilo Cardona, Rade Stanojevic, and Rubén Cuevas. On weather and Internet traffic demand. In *Proc. of PAM*, March 2013.

[19] Pedro Casas, Bruno Gardlo, Raimund Schatz, and Marco Melia. An educated guess on QoE in operationa networks through large-scale measurements. In *Proc. of SIGCOMM Workshop Internet-QoE.*

[20] Kuan-Ta Chen, Chun-Ying Huang, Polly Huang, and Chin-Laung Lei. Quantifying Skype user satisfaction. In *Proc. of ACM SIGCOMM*, 2006.

[21] Kuan-Ta Chen, Polly Huang, and Chin-Laung Lei. How sensitive are online gamers to network quality? *Commun. ACM*, 49(11), November 2006.

[22] Marshini Chetty, Richard Banks, AJ Brush, Jonathan Donner, and Rebecca Grinter. You're capped: understanding the effects of bandwidth caps on broadband use in the home. In *Proc. of SIGCHI*, 2012.

[23] Kenjiro Cho, Kensuke Fukuda, Hiroshi Esaki, and Akira Kato. The impact and implications of the growth in residential user-to-user traffic. In *ACM SIGCOMM Computer Communication Review*, volume 36, pages 207–218. ACM, 2006.

[24] David R. Choffnes and Fabián E. Bustamante. Taming the torrent: A practical approach to reducing cross-ISP traffic in peer-to-peer systems. In *Proc. of ACM SIGCOMM*, 2008.

[25] Cisco. Cisco visual networking index: Forecast and methodology 20122017.

[26] Comcast Business Class. Service level agreement. `http://business.comcast.com/pdfs/cbc-trunks-sla-110922.pdf`.

[27] Editor David Belson. The state of the Internet. Technical report, Akamai, 2015. Q3.

[28] B.S. Dhillon. *Reliability, Quality, and Safety for Engineers*. CRC Press, 2004.

[29] Lucas DiCioccio, Renata Teixeira, Martin May, and Christian Kreibich. Probe and pray: Using upnp for home network measurements. In *PAM*, 2012.

[30] Lucas DiCioccio, Renata Teixeira, and Catherine Rosenberg. Characterizing home networks with HomeNet profiler. Technical Report CP-PRL-2011-09-0001, Technicolor, September 2011.

[31] Marcel Dischinger, Krishna P. Gummadi, Andreas Haeberlen, and Stefan Saroiu. Characterizing residential broadband networks. In *Proc. of IMC*, 2007.

[32] Florin Dobrian, Asad Awan, Ion Stoica, Vyas Sekar, Aditya Ganjam, Dilip Joseph, Jibin Zhan, and Hui Zhang. Understanding the impact of video quality on user engagement. In *Proc. of ACM SIGCOMM*, 2011.

[33] Thad Dunning. *Natural experiments in the social sciences: a design-based approach.* Cambridge University Press, 2012.

[34] Electronic Frontier Foundation. Open Wireless Movement. https://openwireless.org/.

[35] P. Enriquez, A. Brown, and D. A. Patterson. Lessons from the PSTN for dependable computing. In *Proc. of Workshop on Self-Healing, Adaptive and self-MANaged Systems*, 200.

[36] FCC. 2013 measuring broadband America February report. `http://data.fcc.gov/download/measuring-broadband-america/2013/Technical-Appendix-feb-2013.pdf`.

[37] FCC. Measuring Broadband America. `http://www.fcc.gov/measuring-broadband-america`.

[38] FCC. In the matter of preserving the Open Internet broadband industry practices, December 2010.

[39] FCC. In the matter of protecting and promoting the Open Internet, February 2015.

[40] Federecal Communication Commission (FCC). In the matter of reliability and continuity of communication networks. PS Docket 11-60.

[41] Fon. How it Works — Fon. https://corp.fon.com/en/how-it-works.

[42] Thomas Gilovich, Dale Griffin, and Daniel Kahneman. *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press, 2002.

[43] W.H. Green. *Econometric Analysis*. Prentince Hall, 2003.

[44] Krishna P. Gummadi, Harsha V. Madhyastha, Steven D. Gribble, Henry M. Levy, and David Wetherall. Improving the reliability of internet paths with one-hop source routing. In *Proc. USENIX OSDI*.

[45] Kate Harrison. How to break up with your landline. Forbes, August 2012. http://www.forbes.com/sites/kateharrison/2012/08/08/how-to-break-up-with-your-landline/.

[46] Stacey Higginbotham. Why the consumer is still held hostage in peering disputes. http://bit.ly/1KbBBhl, July 2014.

[47] John Horrigan and Ellen Satterwhite. Americans' perspectives on online connection speeds for home and mobile devices. http://www.fcc.gov/encyclopedia/broadband-speed, 2010.

[48] IDC. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east, 2013.

[49] American National Standards Institute. Nema mg-1: Motors and generators.

[50] S. Isaacman and M. Martonosi. The C-LINK system for collaborative web usage: A real-world deployment in rural Nicaragua. In *Workshop on Networked Systems for Developing Regions*, 2009.

[51] ITU. The impact of broadband on the economy. https://www.itu.int/ITU-D/treg/publications/bbreports.html, 2012.

[52] ITU. New global broadband study: national plans and competitive markets are crucial. http://www.itu.int/net/pressoffice/press_releases/2013/27.aspx, July 2013.

[53] Nicholas Jackson. United Nations Declares Internet Access a Basic Human Right. http://www.theatlantic.com/technology/archive/2011/06/united-nations-declares-internet-access-a-basic-human-right/239911/, June 2011.

[54] Dipendra Jha, John P. Rula, and Fabián E. Bustamante. exploring xfinity: A first look at provider-enabled community networks. In *PAM*, 2016.

[55] Christian Kreibich, Nicholas Weaver, Boris Nechaev, and Vern Paxson. Netalyzr: Illuminating the edge network. In *Proc. of IMC*, 2010.

[56] S. Shunmuga Krishnan and Ramesh K. Sitaraman. Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs. In *Proc. of IMC*, 2012.

[57] D. Richard Kuhn. Sources of failures in the public switched telephone network. *IEEE Computer*, pages 31–36, April 1997.

[58] William Lehr, Steven Bauer, Mikko Heikkinen, and David Clark. Assessing broadband reliability: Measurement and policy challenges. In *Proc. of the Research Conference on Communication, Information and Policy Challenges*, 2011.

[59] Jack Loechner. Over the top services outpace pay tv. `http://www.mediapost.com/publications/article/254307/over-the-top-services-outpace-pay-tv.html`, July 2015.

[60] Thomas Robert Malthus. An essay on the principle of population and a summary view of the principle of population. London: Pickering, 1798.

[61] Matthew Mathis, Jeffrey Semke, Jamshid Mahdavi, and Teunis Ott. The macroscopic behavior of the tcp congestion avoidance algorithm. 27(3), 1997.

[62] Sascha D. Meinrath, James Losey, and Benjamin Lennett. A growing digital divid: Internet freedom and the negative impact of command-and-control networking. *IEEE Internet Computing*, 2011.

[63] MLabs. Network diagnostic tool. `http://www.measurementlab.net/run-ndt/`.

[64] Sarah Morris. Spain govt to guarantee legal right to broadband. `http://www.reuters.com/article/spain-telecoms-idUSLH61554320091117`, November 2009.

[65] Search Networking. The Holy Grail of five-nines reliability. `https://goo.gl/3JZRPG`.

[66] Ashkan Nikravesh, David Ke Hong, Qi Alfred Chen, Harsha V. Madhyastha, and Z. Morely Mao. QoE inference without application control. In *Proc. of SIGCOMM Workshop Internet-QoE*, August 2016.

[67] Office of Communication (Ofcom). UK fixed broadband speeds, November/December 2010. Technical report, London, UK, March 2011.

[68] Office of Communication (Ofcom). The consumer experience. Technical report, Ofcom, London, UK, January 2015.

[69] Huseyin Oktay, Brian J. Taylor, and David D. Jensen. Causal discovery in social media using quasi-experimental designs. In *Proc. of the First Workshop on Social Media Analytics*, 2010.

[70] John S. Otto, Mario A. Sanchez, David R. Choffnes, Fabian E. Bustamante, and Georgos Siganos. On blind mice and the elephant – understanding the network impact of a large distributed system. In *Proc. of ACM SIGCOMM*, 2011.

[71] John S. Otto, Mario A. Sánchez, John P. Rula, and Fabián E. Bustamante. Content delivery and the natural evolution of dns: Remote dns trends, performance issues and alternative solutions. In *Proc. of IMC*, 2012.

[72] Christoph Paasch and Olivier Bonaventure. Multipath tcp. *Queue*, February 2014.

[73] Jitendra Padhye, Victor Firoiuy, Dan Towsley, and Jim Kurose. Modeling TCP throughpout: A simple model and its empirical validation. In *Proc. of ACM SIGCOMM*, 1998.

[74] Kyoung Soo Park, Vivek S. Pai, Larry Peterson, and Zhe Wang. CoDNS: Improving DNS performance and reliability via cooperative lookups. In *Proc. USENIX OSDI*, 2004.

[75] Vern Paxson. Strategies for sound Internet measurement. In *Proc. of IMC*, October 2004.

[76] J. San Pedro, D. Proserpio, and N. Oliver. Mobiscore: Towards universal credit scoring from mobile phone data. 2015.

[77] B. Peirens and et al. Link bonding with transparent Multipath TCP. `https://tools.ietf.org/html/draft-peirens-mptcp-transparent-00`, July 2016.

[78] Policy by the numbers. International broadband pricing study: Updated dataset. `http://policybythenumbers.blogspot.com/2013/05/international-broadband-pricing-study.html`.

[79] Alan Ritacco, Craig Wills, and Mark Claypool. How's my network? A Java approach to home network measurement. In *Proc. of IEEE ICCCN*, 2009.

[80] Claudio Rossi, Narseo Vallina-Rodriguez, Vijay Erramilli, Yan Grunenberger, Laszlo Gyarmati, Nikolaos Laoutaris, Rade Stanojevic, Konstantina Papagiannaki, and Pablo Rodriguez. 3gol: Power-boosting adsl using 3g onloading. In *Proc. ACM CoNEXT*, 2013.

[81] SamKnows. Accurate broadband information for consumers, governments and ISPs. `http://www.samknows.com/`.

[82] SamKnows. Samknows & the FCC American broadband performance measurement. http://www.samknows.com/broadband/fcc_and_samknows, June.

[83] Mario A. Sánchez, John S. Otto, Zachary S. Bischof, and Fabián E. Bustamante. Trying broadband characterization at home. In *Proc. of PAM*, 2013.

[84] Mario A. Sánchez, John S. Otto, Zachary S. Bischof, David R. Choffnes, Fabián E. Bustamante, Balachander Krishnamurthy, and Walter Willinger. Dasu: Pushing experiments to the Internet's edge. In *Proc. of USENIX NSDI*, 2013.

[85] Experian Marketing Services. Cross-device video analysis. http://www.experian.com/assets/marketing-services/brochures/cross-device-video-analysis-2014.pdf.

[86] Skype. Plan network requirements for skype for business. https://technet.microsoft.com/en-us/library/Gg425841.aspx.

[87] Philip Smith. Bgp multihoming techniques. goo.gl/gVKYwmcontent_copyCopyshortURL, October 2004.

[88] Joan E. Solsman. Cogent: Comcast forced netflix with clever traffic clogging. http://cnet.co/1l3aDw1, May 2014.

[89] Joel Sommers, Paul Barford, Nick Duffield, and Amos Ron. Accurate and efficient SLA compliance monitoring. In *Proc. of ACM SIGCOMM*, 2007.

[90] Speedtest.net. The global broadband speed test. http://www.speedtest.net/.

[91] Rade Stanojevic, Vijay Erramilli, and Konstantina Papagiannaki. Cognitive bias in network services. In *Proc. of HotNets*, 2012.

[92] Frost & Sullivan. The advanced capabilities of voip and sip trunking exceeds traditional voice services, February 2014. `http://tinyurl.com/z349mxt`.

[93] Srikanth Sundaresan, Walter de Donato, Nick Feamster, Renata Teixeira, Sam Crawford, and Antonio Pescapè. Broadband internet performance: a view from the gateway. In *Proc. of ACM SIGCOMM*, 2011.

[94] Srikanth Sundaresan, Nick Feamster, Renata Teixeira, and Nazanin Magharei. Measuring and mitigating web performance bottlenecks in broadband access networks. In *Proc. of IMC*, October 2013.

[95] Srikanth Sundaresan, Nick Feamster, Renata Teixeira, Anthony Tang, W. Keith Edwards, Rebecca E. Grinter, Marshini Chetty, and Walter de Donato. Helping users shop for ISPs with internet nutrition labels. In *Proc. of HomeNets*, 2011.

[96] World Bank. IC4D 2009: Extending reach and increasing impact. `http://go.worldbank.org/NATLOH7HV0`.

[97] Arvind Krishnamurthy Xiao Sophia Wang, Aruna Balasubramanian and University of Washington David Wetherall. Demystifying page load performance with wprof. In *Proc. of USENIX NSDI*, 2013.

[98] Yang Xu, Chenguang Yu, Jingjiang Li, and Yong Liu. Video telephony for end-consumers: measurement study of Google+, iChat, and Skype. In *Proc. of IMC*, 2012.

[99] Mariya Zheleva, Paul Schmitt, Morgan Vigil, and Elizabeth Belding. The increased bandwidth fallacy: performance and usage in rural Zambia. In *Proc. of the 4th Symposium on Computing for Development*, 2013.