Revisiting Broadband Performance

Igor Canadi
University of Wisconsin-Madison
canadi@cs.wisc.edu

Paul Barford
University of Wisconsin-Madison
pb@cs.wisc.edu

Joel Sommers
Colgate University
jsommers@colgate.edu

ABSTRACT

Understanding the empirical characteristics of broadband performance is of intrinsic importance to users and providers, and has been a significant focus of recent efforts by the Federal Communications Commission (FCC) [9]. A series of recent studies have reported results of empirical studies of broadband performance (*e.g.,* [11, 15, 22]). In this paper, we reappraise previous empirical findings on broadband performance. Our study is based on a unique corpus of crowd-sourced data consisting of over 54 million individual tests collected from 59 metropolitan markets over a 6 month period by Speedtest.net. Following analytic approaches from prior studies, our results confirm many of the raw performance results (upload/download/latency) for ISPs in specific US markets. However, the size and scope of our data enable us to examine the details of characteristics that were not identified in prior studies, thereby providing a more comprehensive view of broadband performance. Furthermore, we also report results of broadband performance characteristics in 35 metropolitan markets outside of the US. This not only provides an important baseline for future study in those markets, but also enables relative comparison of broadband performance between markets world wide.

Categories and Subject Descriptors

C.2.3 [**Network Operations**]: Network management; C.4 [**Performance of Systems**]: Performance attributes; C.4 [**Performance of Systems**]: Measurement techniques

General Terms

Experimentation, Measurement, Performance

Keywords

Access networks, Broadband access

1. INTRODUCTION

High speed broadband connectivity to the home is the foundation for rich media applications and services. Over the past several years, the impact of streaming media enabled by broadband to the home has been dramatic. For example, the growing phenomenon of "cord cutting" — doing away with cable television subscriptions in lieu of streaming — has important implications for service providers, television, advertising and other industries [20].

The increasing reliance by consumers on high speed broadband connectivity has made its performance and reliability a matter of public interest. To that end, the US Federal Communications Commission (FCC) has initiated efforts to develop a national broadband strategy that promotes competition among providers and uses data-driven standards to ensure that service providers do not abuse their market power [8]. The "data driven" mandate led to the FCC's consumer broadband test site, which enables users to evaluate their broadband performance with several different tools [9].

While broadband performance test sites have been available for some time (*e.g.*, [3,13]) measuring, characterizing, and understanding broadband performance in the large presents several significant challenges. The first is perspective: with nearly two billion Internet users world wide, the task of gathering "representative" data and extracting meaningful information even with a metropolitan area focus is daunting. Second, taking measurements of last mile performance necessitates instrumentation at an end host. While many tools (*e.g.,* [7,15]) and systems (*e.g.,* [4]) for end host-based performance measurement exist, deploying them broadly and then using them to conduct systematic study intrinsically relies on user participation, which is obviously challenging. Third, different test methods and the diversity of infrastructure in the last mile complicate the interpretation and comparison of measurements. For example, users who run tests from their laptops may be limited by wireless interference instead of their broadband connection. We argue that the collective perspective offered by multiple tests methods is the best way to gain an understanding of broadband performance.

In this paper we present results of a study of broadband performance that seeks to reconfirm results from prior studies and to expand the knowledge base in this critical area. Our study is based on a unique data set provided by `speedtest.net` [3]. Speedtest is a popular end host-based performance measurement application that directs participating users to local servers in over 700 metropolitan markets world wide. We evaluate a data corpus of over 54 million performance tests collected over a 6 month period in 59 metro markets (24 US, 35 non-US).

While the goals of our work are aligned with prior broadband performance studies (*e.g.,* [11, 15, 17, 22]), our approach expands on prior studies in several important ways. Most importantly, we provide a crowd-sourced perspective on broadband performance. While Speedtest data are limited in several ways (see Section 2), the size, diversity and rich details of these data enable us to reconsider results from prior studies. Moreover, we identify new

IMC'12, November 14–16, 2012, Boston, Massachusetts, USA.

characteristics of broadband performance that to the best of our knowledge have not been previously reported. The scope of the Speedtest data also enables us to examine broadband performance from a world-wide perspective. Prior studies have typically focused on broadband performance in the US or Europe. We report results that compare and contrast broadband performance for major geographic areas all over the world.

The first component of our analysis considers broadband performance in 25 metropolitan areas in the United States. The metro areas represent a broad cross section of the population in terms of market size and geographic location. Our analyses focus on download, upload and latency measurements and consider recent results reported by Sundaresan *et al.* [22] as a point of comparison. We begin by reporting results of a small case study that compares Speedtest measurements to SamKnows measurements [4], which were the basis for many of the results in [22]. Next, we consider the question of whether or not users experience consistent performance from their providers, in many cases our findings support observations of prior work (*e.g.*, in terms of lower latencies for cable vs. DSL), but in others they are not (*e.g.*, in terms of provider quality ranking). We make no argument that in cases where our findings differ, that ours are "more valid". Rather, we believe that our findings show that performance comparisons are complex and dynamic, and benefit from on-going study.

The second component of our analysis takes advantage of the crowd-sourced nature of our data to dig into the details of local connectivity. We focus on two characteristics that emerge in our data that were not evident or highlighted in prior studies. The first characteristic is a tight clustering of measurements that show lower performance. While we can not absolutely validate the causes, this could be explained by older service plans that have not been upgraded, older equipment, or rate limiting. The second characteristic considers the question of latency versus distance from a test server. While this may seem obvious, prior studies identify buffers as a potential dominant factor in latency [17]. Our geographic analysis confirms that there is a strong correlation between latency and distance.

The final component of our analysis considers broadband performance in 35 metropolitan areas outside of the US. Our analysis methodology is consistent with studies of US and European broadband performance. Perhaps as expected, we find that broadband performance in Hong Kong is superior to the US on average. We find that performance in Europe is on par with the US on average. We find that performance in South American and African sites is below the US, while performance is relatively lower in India and Indonesia. These results highlight the relative penetration of high speed broadband world wide and have implications for economic development and competitiveness.

The remainder of this paper is organized as follows. In Section 2, we describe the details of the Speedtest measurement methodology, our data set and its limitations. In Section 3, we describe the details of the evaluations that we conduct on our data. The results of our analyses are reported in Section 4. We discuss prior studies that inform our work in Section 5. We summarize, conclude and discuss future work in Section 6.

2. DATA

In this section we describe the Speedtest data set used in our study. We provide an overview of the measurement algorithm used by Speedtest, the deployment of Speedtest servers around the world, and the specific types of data that are collected in their infrastructure. We list the 59 metro areas in which the data were collected. We also discuss limitations with the Speedtest data and how

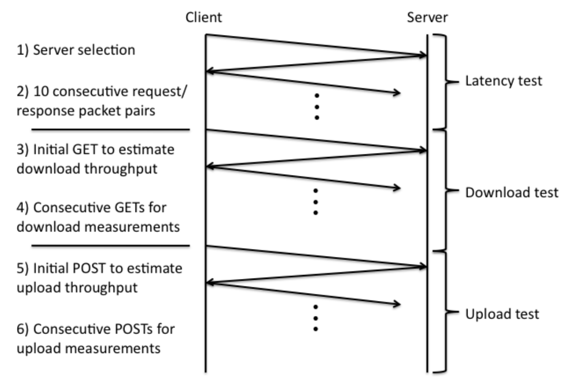

Figure 1: The Speedtest performance measurement protocol, which is activated when clients start a test.

they may affect our comparisons with prior broadband performance studies.

2.1 Speedtest Measurements

`Speedtest.net` [3] is an openly available, Flash-based bandwidth/performance measurement application that is managed and maintained by Ookla, Inc. [2]. Over 3 billion performance tests have been run since 2006 when Speedtest came on line, and there are hundreds of thousands of tests performed per day, globally.

The testing protocol that is initiated by clients running the web-based Speedtest application is shown in Figure 1. When a client initiates a test, a request is automatically sent to the Speedtest server that is identified as being geographically nearest to the client. Alternatively, clients can specify a server for performing a test. There are Speedtest servers deployed in over 700 locations world wide.

The Speedtest wiki [1] explains the details of each test. Measurements of latency, download, and upload performance are conducted via HTTP (TCP port 80). Latency tests use a minimum of 10 RTT ping-style measurements and report the mean result. The first step in a download test is the transfer of a fixed-size file from server to client. This transfer is used to determine the size of a target file that will be used in the actual performance tests. The target file is fixed at a smaller size when the initial transfer indicates a lower bandwidth path. Larger size target files are used when the initial transfer indicates higher bandwidth. The second step in a test consists of repeated transfers of the target file using up to 8 parallel HTTP connections. Care is taken to avoid client caching. Throughput is sampled at up to 30 times per second; the top 10% and bottom 30% are discarded and the mean of the remaining samples is used as the overall throughput estimate. This approach is motivated by an attempt to obviate burst effects due to OS overhead and other host-system effects, and produce a maximum throughput estimate that corresponds to the expected case for the user. Test runs are designed to be relatively short in order to enhance user experience. The use of HTTP (and by extension, TCP) in the Speedtest protocol is prompted by a desire to gather performance measures similar to what many applications would experience.

Upload tests come after download tests. They are designed to follow a protocol that is similar to the download tests.

Each test results in a log entry stored at the local Ookla server that includes a unique test identifier, timestamp, upload and download throughput estimates (in Kb/s), latency estimate (in milliseconds), anonymized client IP address, browser user agent string, client geographic coordinates (longitude / latitude), server name and coordinates, geographic distance between client and server, and ISP name. While Speedtest can be run from either wireline or mobile clients, log entries are easily distinguished by device type and access type.

2.2 Data Sets Considered

The data we use in this study were collected from Speedtest servers located in 59 metro areas over a period of six months from June 1, 2011 through November 30, 2011. The goal in selecting metro areas was to provide a broad perspective on broadband performance. The criteria that were used for selection included geographic location of servers, metro market sizes and socio-economic information. The result were 24 metro areas in the US and 35 outside of the US.

For US, we selected 24 geographically diverse metro areas that can be grouped by population size: small (*e.g.,* Lawrence, KS; Sumter, SC; Grand Forks, ND), medium (*e.g.,* Portland, ME; Grand Rapids, MI; San Francisco, CA) and large (*e.g.,* New York, NY; Los Angeles, CA; Chicago, IL).

For areas outside of the US, we selected ten metro areas in Europe, and five in Asia, Pacific, Middle East, South America and Africa respectively for a total of 35. Similar to the US, these selections were based on a geographic diversity and market size with an attempt to select areas for each category that had roughly the same population. In order to gain a measure of consistency in our evaluations, we only include tests (directed to that server) that are conducted within a 200 miles radius. The full set of metro areas, along with summary statistics of the associated Speedtest results can be found in Table 2.

Our data indicates that roughly 300,000 users per day initiated tests to the 59 servers considered in this study. As noted above, we also have the browser user agent string. Figure 2 shows a breakdown of client operating system and browser types in our data. Further, each test entry includes the client's geographic coordinates. These coordinates are determined from GPS, if available, or by using the MaxMind GeoIP geolocation database [19], which is generally accurate at the city level. Locations of test clients in our data set are shown in Figure 3 (we only show test locations that are within a 200 mile radius of the Speedtest servers included in our study).

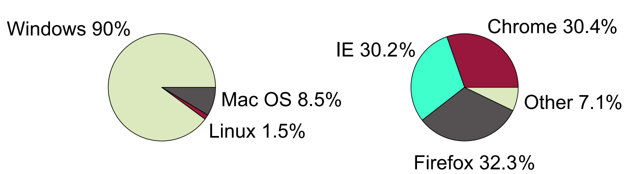

Figure 2: Profile of Speedtest users's Web browsers and Operating systems

2.3 Discussion

We believe that the data provided by Speedtest offer a unique and valuable perspective that enhances the general understanding of broadband performance. The broad use of Speedtest around the world enables detailed profiles of broadband performance to be established for diverse markets. The fact that Speedtest has been using substantially the same measurement methodology for years enables longitudinal study of broadband performance and comparisons across markets. The web-based delivery of tests precludes the need for deployment and management of dedicated systems or software at client sites and encourages broad use of the tool.

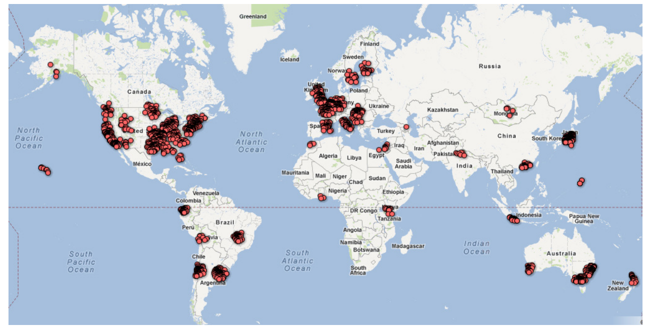

Figure 3: Locations of measurements

While the crowd-sourced nature of the data set is compelling, there are limitations that are important to recognize since they could have an influence on the findings in our study. Speedtest data are assembled based on diverse users running the test application. While some information about client hosts is available (*e.g.,* host operating system via the browser user agent string), this information cannot be verified. Furthermore, host systems are likely to vary widely in terms of hardware and software configurations. There also is no way to dictate when or how tests are run. Indeed, Speedtest may well be invoked when users believe their network performance is poor or otherwise problematic. This may bias results of performance tests toward below average operating conditions. We have no way of establishing baselines for performance or assessing testing bias in each metro area that we consider other than appealing to statistical characterizations and the relatively large numbers of tests conducted over the six month period from which our data is collected. Nevertheless, many of our conclusions described below are consistent with prior work, suggesting that any biases are fairly limited in nature or that prior work suffered from similar or the same biases.

An additional limitation that is specific to our objective of reproducing prior results is the fact that we cannot identify specific users in the Speedtest data. While this does not preclude all analyses, it does mean that can cannot evaluate the performance observed by specific users, which was the focus of certain aspects of prior broadband studies.

3. EVALUATION METHODOLOGY

Analyzing and drawing conclusions from a large, multidimensional data set presents a number of challenges. Toward the goal of reconfirming prior results, many of our evaluation techniques are taken directly from past studies which enables us to draw meaningful comparisons.

To assess the high level characteristics of throughput and latency, we provide summary statistics (average and standard deviation) for each metro area. We also use simple summary statistics for country-level aggregates and present the results using bubble plots, similar to prior work [15]. The location of each bubble is the average throughput performance and the area of bubbles is proportional to the number of measurements. When comparing different US operators, we also use bubble plots to present their average performance and popularity amongst users.

We consider characteristics of specific service providers using scatter plots of download and upload speed for each measurement.

Similar graphical analyses have been used in prior work, such as [22]. When plotting, we make every point a semi-transparent circle, which makes it easy to assess the areas of high and low density: high density areas appear darker, while low density areas appear lighter. We can also more easily distinguish different data points. Due to the large number of measurements for every ISP, in some cases we had to randomly sample the data used in the plots to create manageable figures. Every plot contains roughly 30,000 points that were (uniform) randomly selected from the full dataset. Finally, while scatter plots are a natural way to capture data from a single source, it is difficult to compare two scatter plots. Thus, to compare operators and show correlation between parameters, we use various types of line plots and bar charts.

The Speedtest measurement method estimates upload/download performance from a user to the closest server. If a user is geographically distant from the server, the measurement may reflect wide area effects such as congestion. For this reason, we only consider measurements within a radius of 200 miles of a given server. Latency measurements are potentially the most sensitive to distance (although we recognize that geographic distance may not correlate perfectly with network latency). In our latency analyses, unless otherwise noted, we discard measurements that are more than 50 miles away from the server.

Finally, we note that our data analysis was facilitated through development of a MySQL database on which we ran queries using a series of Python scripts. Statistical analysis and plotting were done using R.

4. BROADBAND PERFORMANCE RESULTS

In this section we report the results of our analysis. Our primary comparisons are with the dataset and findings of the recent study of Sundaresan *et al.* [22], but we also compare our results and data to the prior work of Kreibich *et al.* [15]. We expand on these prior works by examining new geographic regions, and discuss other findings specific to our dataset.

4.1 Alternative Broadband Testing Protocols

The study by Kreibich *et al.* uses data from their Netalyzr system [15], and the recent study by Sundaresan *et al.* [22] reported findings from empirical broadband performance data based on *SamKnows* deployments and from their own BISMark system. Below, we discuss these measurement methods and specifically compare the SamKnows testing method with Speedtest.

4.1.1 SamKnows

SamKnows operates a FCC-sponsored broadband access network study, and has developed a test methodology focused on measuring different aspects of a broadband Internet connection from the gateway's point of view [4]. The tests are executed from a device called a "whitebox", which is a device situated between the access network modem and home router. These devices are deployed in the homes of volunteers, and they contain software that attempts to run a connection test every two hours. The whitebox has the ability to monitor both wired and WiFi network traffic and does not perform tests when it senses any user activity, in order to obtain an accurate measure of access network capabilities.

To measure download and upload speeds, whiteboxes open three concurrent connections with a SamKnows server. First, a warm-up sequence is initiated in which each connection repeatedly downloads small chunks of the target payload. The warm-up period is used to avoid effects of TCP slow start, and network congestion. When three consecutive chunks are downloaded at the same speed, the real testing begins. The test is run for a fixed period of time, or a fixed-size payload. After the test ends, each connection reports reports the throughput it obtained, and the speeds from each connection are summed.

4.1.2 Comparison with SamKnows

Although there are some differences between SamKnows and Speedtest methodologies in the number of parallel connections, and methods of aggregating the raw results, we believe that the biggest difference in reported values is due to where and when the test is executed. SamKnows deploys their whiteboxes between a router and a modem, and the monitoring capabilities designed into the devices help to avoid cross-traffic effects from home networks. Speedtest, on the other hand, can be run from clients connected to home networks over wireless connections. It can also be run when there is activity on the line created by other clients.

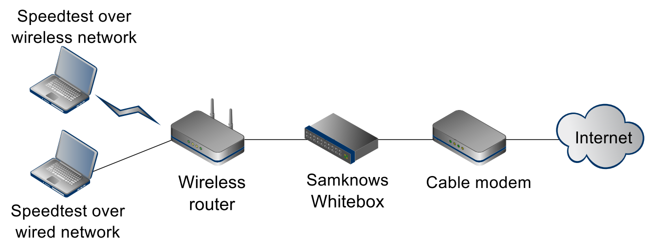

Figure 4: Setup of the network we used to compare Speedtest and Samknows test methodology.

To better understand and evaluate the mentioned differences, we ran both Speedtest and Samknows tests from a network connected to Charter's cable Internet service in Madison, WI. The operator's promised speed was 15 Mb/s download and 3 Mb/s upload with Powerboost [6]. Figure 4 shows the network setup that was used. End hosts were connected to a router through both wired and wireless connections. The router was connected to a SamKnows whitebox, which was in turn connected to the cable modem. We ran a total of 12 tests over the course of two weeks, six in the day time (9am – 4pm) and six in the night time (8pm – 2am). Every test consisted of three runs of each of the following: (1) executing Samknows download and upload test binaries on the whitebox; (2) running a Speedtest from a Macbook Pro connected to the router over 100 Mb/s wired Ethernet; and (3) running a Speedtest from a Macbook Air connected to the router over wireless 802.11g Ethernet.

Figures 5 and 6 show averages of the three runs for each measurement. SamKnows almost always reported higher download speeds than Speedtest over the wired network. However, the difference between the two was very small. Executing Speedtest experiments over the wireless network, on the other hand, shows much slower download performance in all tests except one. The wireless network we were testing on supports download speeds up to 54 Mb/s, so it should not be a bottleneck. The one test where we achieved the full potential of the Internet connection is intriguing; the cause for this surprisingly good performance is presently unclear. SamKnows tests also show higher upload speeds and more consistent results than Speedtest in almost every case. A clear limitation with our Speedtest data is that we have no way of separating the measurements conducted over a wireless network from the ones over a wired network.

To examine the behavior of the Speedtest application in the presence of cross traffic in the home network, we started a download of a large file over the wireless network and simultaneously ran Speedtest from another machine over the wired network. After executing the Speedtest measurement, we executed a SamKnows performance test from the whitebox. Note that while the whitebox usually does not start a test if the line is not idle, it is possible to disable that check, which we did for this experiment. Our results show that a single large file download affects Speedtest, which reports an average of 11 Mb/s, while the SamKnows whitebox reports the correct speed of 15 Mb/s. When we started multiple large file downloads and ran the Speedtest application and whitebox measurement, the results were poor for both Speedtest and SamKnows. The whitebox reported approximately 7.5 Mb/s, and the Speedtest application gave an even lower estimate.

Figures 7 and 8 show the standard deviation of the three runs, presenting the consistency of the results over short time spans. The standard deviation of SamKnows was the lowest for both upload and download throughput tests. In the data point where Speedtest over the wireless network matched the performance of the wired network, the standard deviation of the wireless tests was low and exhibited similar variability as the wired Speedtest and SamKnows measurements. In these experiments, we did not find any significant differences between tests performed during the night and during the day.

4.1.3 Netalyzr

Netalyzr is a Java-based tool that executes various types of probing in order to diagnose access network behavior and performance. Of relevance to our study is the fact that it uses UDP probes to estimate both latency and throughput. It uses a train of small UDP probes sent 100 millisecond intervals to measure latency, then employs a slow-start-like algorithm to detect maximum upload and download throughput. Due to a limitation of the Java environment, there are limits to the maximum throughput detected by Netalyzr (about 20 Mb/s). The client-side tool connects to a back-end server running on Amazon EC2.

A notable similarity between the Speedtest method and Netalyzr is that the measurements are initiated from a computer *behind* the access network gateway/router. This computer may be connected to the local network via WiFi or wired Ethernet, and test traffic is subject to interference with other traffic on the local network. While these end-host based measurements are subject to errors induced by interference (*e.g.*, cross-traffic interference, WiFi contention, etc.), it is arguable that these types of endhost-based measurements give a more accurate perspective of performance that *users* experience. Moreover, end-host based measurements are significantly easier to collect since no special measurement hardware is required. As a result, platforms like Netalyzr and Speedtest add a rich and complementary perspective on broadband performance.

4.2 Revisiting previous findings from US markets

One of the goals of our work is to reassess empirical findings from previous studies [11, 15, 22] in light of the dataset obtained from Speedtest.

Sundaresan *et al.* [22] analyzed SamKnows data from 8 of the biggest ISPs in the US market: Comcast, AT&T (SBC Internet Services), TimeWarner (Road Runner), Verizon, Cox, Qwest, Charter and Cablevision (Optimum Online). The advantages of their data were two fold. First, they have speed measurements from one user taken periodically every hour over the course of few months. Due to privacy reasons, we were not able to connect measurements with specific clients. This limitation prevented us from repeating some of their findings. Second, their tests were conducted from the gateway, thus avoiding bottlenecks and losses in home network. A study by Bauer *et al.* found that the SamKnows method was the most accurate broadband testing method [5], but we note that they found that the Speedtest method also gives a reasonably accurate characterization of access network performance. Furthermore, the Speedtest measurements capture real user experience, and measure the throughput speeds that users actually obtain, as opposed to what they might obtain in perfect conditions.

Figure 9 shows average speeds obtained by users of different ISPs. The center of each circle represents the average download and upload performance obtained by users of a specific ISP. The area of the circle is proportional to the number of measurements that were performed by users of a given ISP on Speedtest's website.

One of the questions addressed in the prior work of Sundaresan *et al.* was whether users achieve consistent performance. Their conclusion was that most of the ISPs provide consistent speeds, except for some ISP-based exceptions (Cablevision, and to a lesser extent, Cox), whose large fraction of users have inconsistent download speeds. We cannot directly compare our results with this prior work due to the inability to identify particular clients or users in our data. Instead, we indirectly evaluate consistent via scatterplot analysis in Figures 10 and 11. Figure 10 shows results for Cablevision, and Figure 11 shows results for Charter; we note that both of these ISPs are cable providers. Every point in the graph is one measurement from a user in the United States. In our measurements, we found that Cablevision exhibited the best upload speed consistencies (as defined by variability in the measurements), and we see that patterns in the download vs. upload scatter plots are very similar. These results suggest that users may indeed obtain consistent performance from these ISPs, although we cannot directly verify that. In terms of download speeds, both ISPs exhibit high variability: in both plots, download speeds are spread across the large interval, while upload speeds are largely localized in small intervals. Our results at least confirm the notion that upload speeds are more consistent than download speeds.

To compare cable operators with DSL operators, in Figures 12 and 13 we show the results of measurements for AT&T and Qwest operators, respectively. In general it seems that DSL operators show better clustering of points around their service plans. They also show higher variability in upload rates than cable operators.

In addition to download and upload throughput, our data also contain latency measurements between the user and Speedtest's server. Sundaresan *et al.* measure and make conclusions based on *last-mile latencies*, which is the latency from the gateway to the first hop in network. One of their results is that cable operators generally exhibit lower latencies than DSL operators. Although our dataset includes end-to-end latency measurements, we observe the same result. Figure 14 shows the latency distribution for cable and DSL operators. 60% of all the measurements for the cable operators are lower than 20 ms. The only exception is Cox, whose majority of users are in the 20–29 ms range. DSL users see higher latencies and more variance. Almost half of the measurements on Qwest reported latencies larger than 60 ms. Verizon's users with latency below 20 ms are their FiOS users.

Figure 15 shows average download throughputs for measurements with specific latencies. It is clear that the average download drops as latency increases. One explanation for this effect could be that measurements with bigger latencies were conducted under higher load. This explanation may correspond to findings of the study by Sundaresan *et al.*, which attribute the effect to large modem buffers. Another possible explanation is the well-known bias of TCP against longer round-trip times, which may result in lower throughputs achieved.

Sundaresan *et al.* also looked into how time of day affects user's performance. They observed the biggest difference between peak

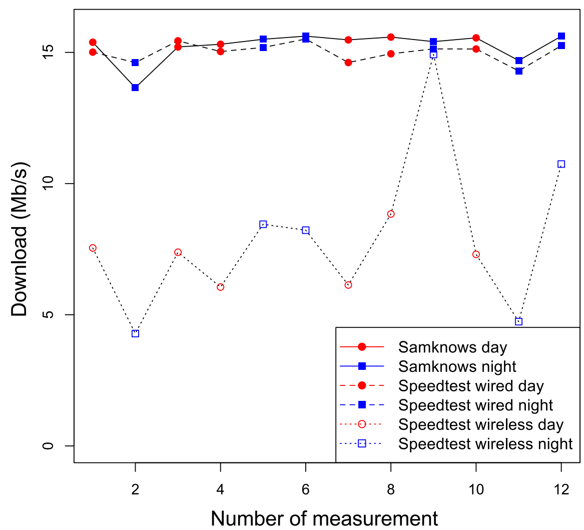

Figure 5: Download throughput as reported by Speedtest on wired and wireless network and SamKnows.

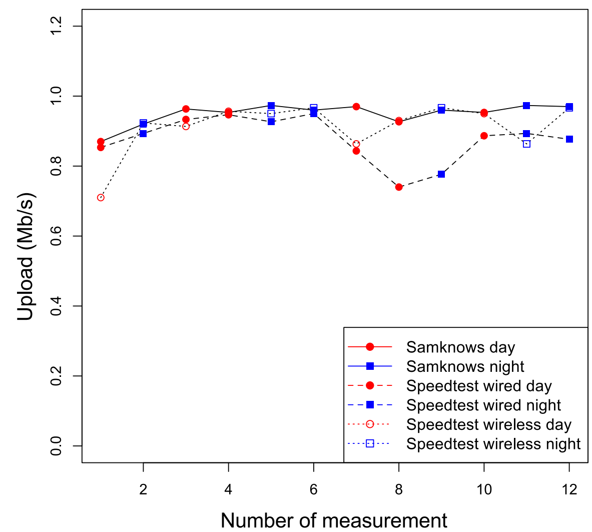

Figure 6: Upload throughput as reported by Speedtest on wired and wireless network and Samknows.

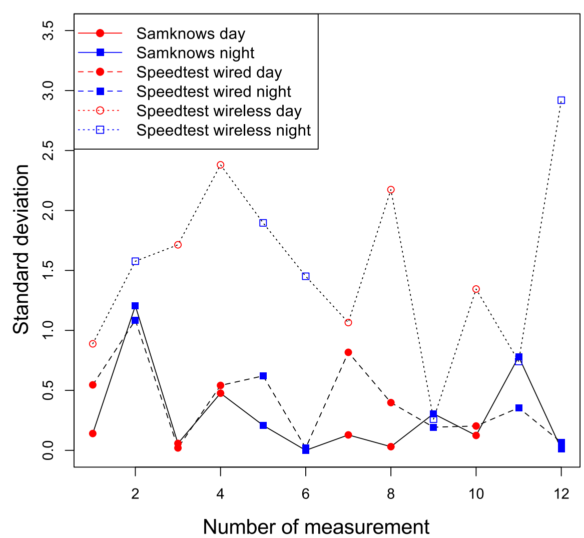

Figure 7: Standard deviation of download throughput of three consecutive runs of Speedtest on wired and wireless network and Samknows

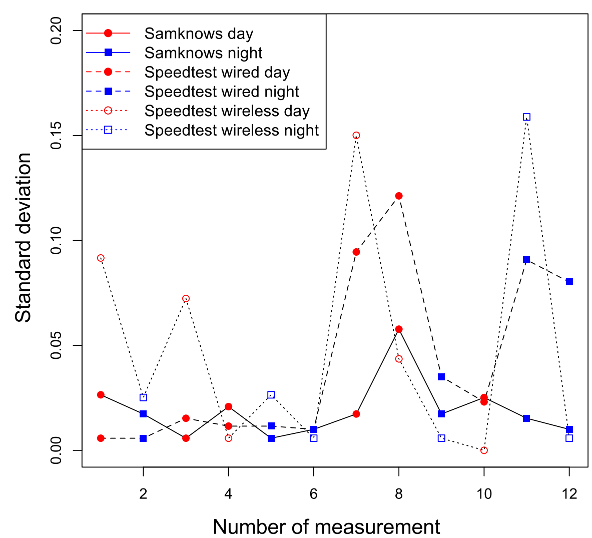

Figure 8: Standard deviation of upload throughput of three consecutive runs of Speedtest on wired and wireless network and Samknows

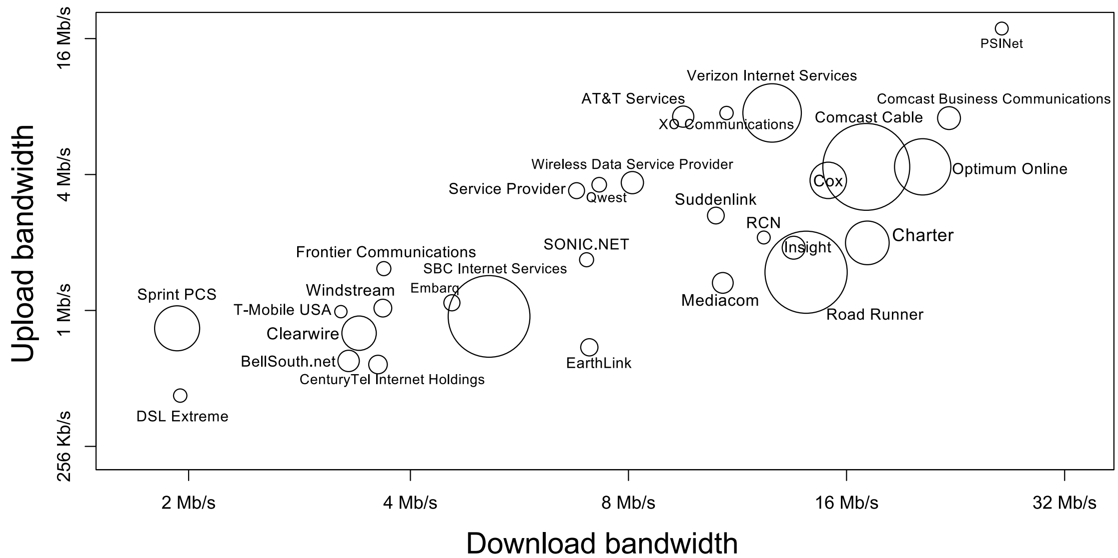

Figure 9: Average download/upload bandwidths for most frequent ISPs from US markets. Circle areas are proportional to number of measurements from that ISP.

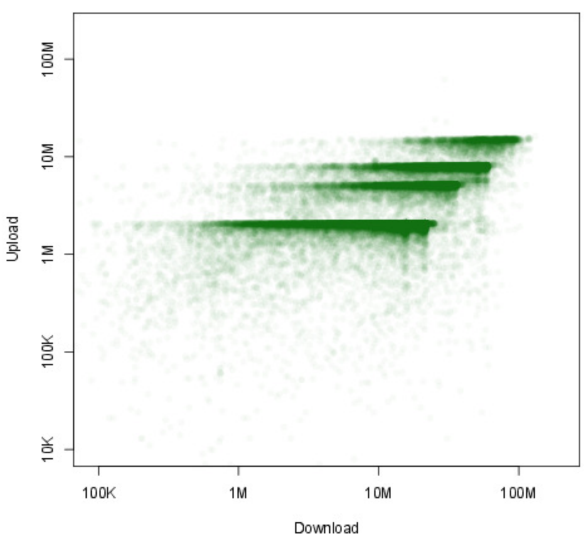

Figure 10: Scatter plot of download vs. upload throughput for Cablevision.

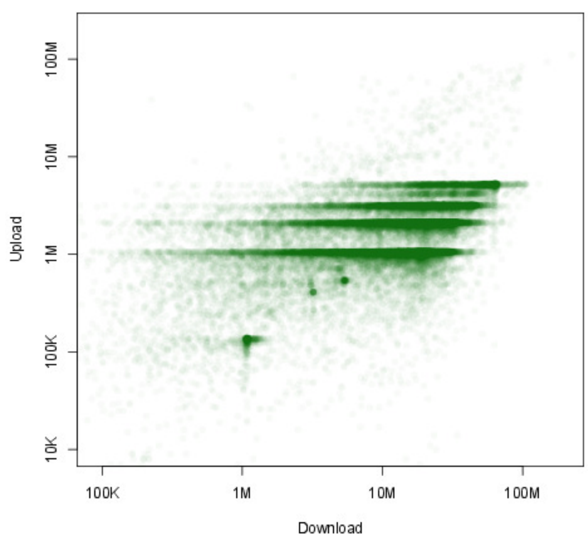

Figure 11: Scatter plot of download vs. upload throughput for Charter.

and worst performance of 40%. They also showed that some operators have consistent performance throughout, while the others exhibit degraded service during peak hours. To confirm their finding, we had to be careful due to potential time biases in our data. For example, during working hours most Speedtest experiments are likely conducted from businesses and universities, while during the evenings and nights, users run the experiments from their home networks. To overcome this limitation, we consider only experiments conducted on a Sunday. We limited our analysis to the Los Angeles metro area and we consider two operators, Cox and Time Warner. In prior work, Cox exhibited significant degradation during peak hours, while Time Warner showed consistent performance. Figure 16 shows the average download speed during specific hours of the day for Cox and Time Warner in the Los Angeles metro area. Our dataset confirms both findings by Sundaresan *et al.*. First, download speeds depend on time of day. Second, some operators have significant differences between peak and worst performance. Cox's peak average performance is two times better than the av-

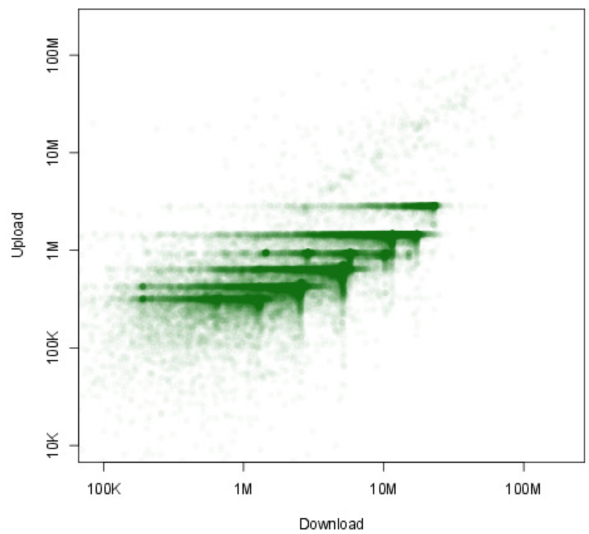

Figure 12: Scatter plot of download vs. upload throughput for AT&T.

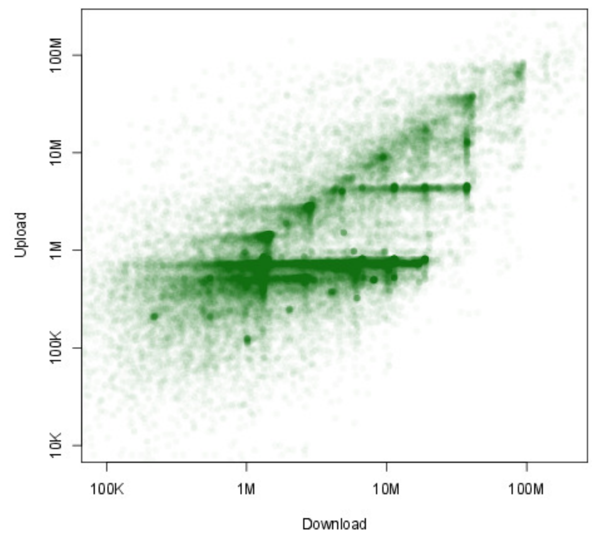

Figure 13: Scatter plot of download vs. upload throughput for Qwest.

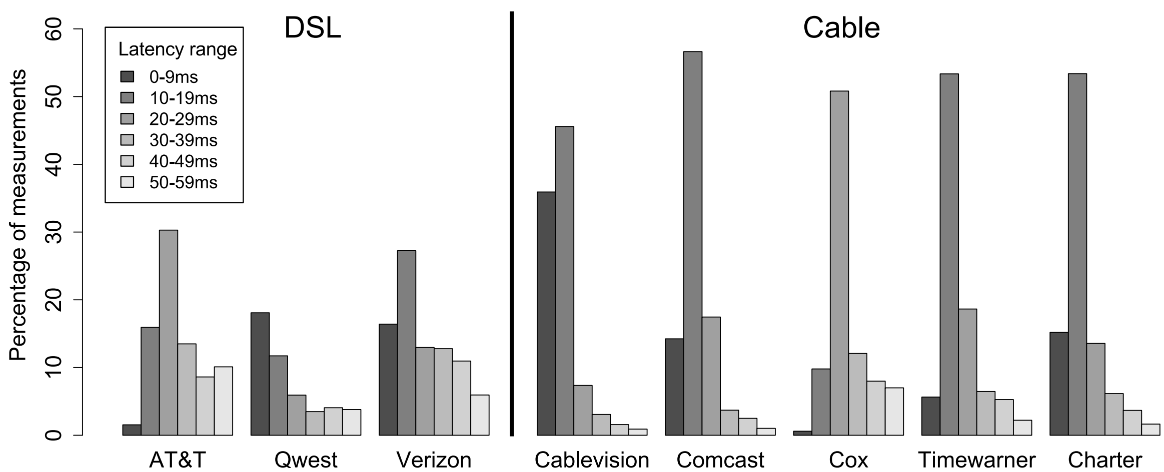

Figure 14: Latency distribution for operators from US markets.

erage download speed during busy times. Time Warner showed a slowdown of only 26%.

4.3 New Findings

In addition to analyzing prior results in light of our Speedtest data, we observed features not previously reported. Here we report on two of those: persistent low performance tests, and the impact of distance from the measurement server and network latency. Because of its wider perspective, these characteristics are more easily measured and observed with a dataset such as Speedtest, as compared with purpose-built and more narrowly deployed system like SamKnows.

4.3.1 Persistent low performance tests

During our analysis of download/upload scatter plots, we observed a number of high density dots in low performance areas, suggesting persistent low performance. For example, in Figure 11, we observed many measurements with a download speed a 1 Mb/s and upload speed of 120 kb/s. Comcast, Cox and Verizon also had similar artifacts with low performance, each with somewhat different specific upload / download measurements. We hypothesize that these artifacts represent "cheap and slow" service plan offerings. While operators may no longer offer these low performance plans, it is clear that some of the customers have not yet switched to new service plans. Because the rates are much slower, the tests easily saturate the rate limits, yielding little variability in the resulting measurements.

However, we also observed more interesting high density points, which we could not explain by cheap service offerings. Speedtest's servers in Australia (Sydney, Melbourne, Perth) reported many measurements of 1.3 Mb/s download speed and 210 kb/s upload speed for all of Australia's biggest Internet operators (iiNet, Internode, Optus, Telstra, TPG). As an example, Figure 17 shows the scat-

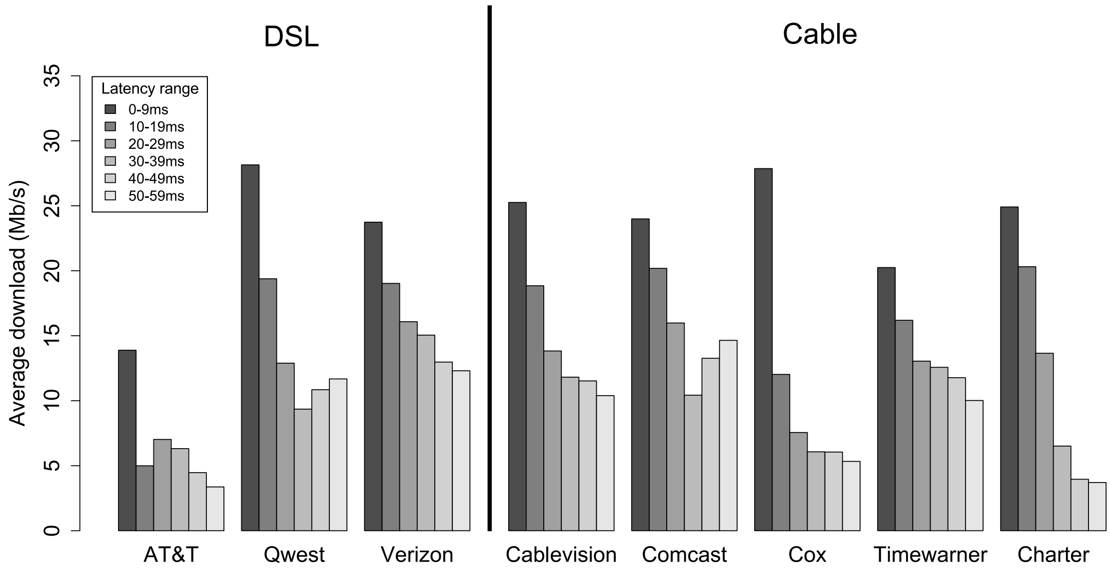

Figure 15: Average download speed with respect to latencies for operators from US markets.

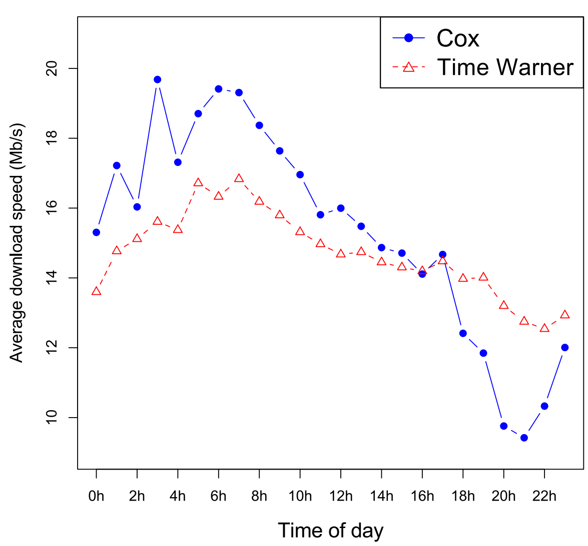

Figure 16: Average download performance for specific time of day. Analysis of Cox and Time Warner in Los Angeles, CA metro area.

ter plot of download and upload speeds for Optus Internet. The measurement that is consistent with other Australian operators is circled; other high-density points representing persistently low performance tests are not common to all the Australian operators.

One explanation for this phenomenon is that once a user exceeds the service plan's download quota, there is a resulting reduction in speed. However, websites for the above network operators state that speeds are usually reduced to 128, 256 or 512 kb/s, which we do not see in our results. The other possible explanation is that the patterns are related to old equipment or old configuration. Customers might not be achieving the internet speeds they are paying for because of old equipment that does not support faster speeds.

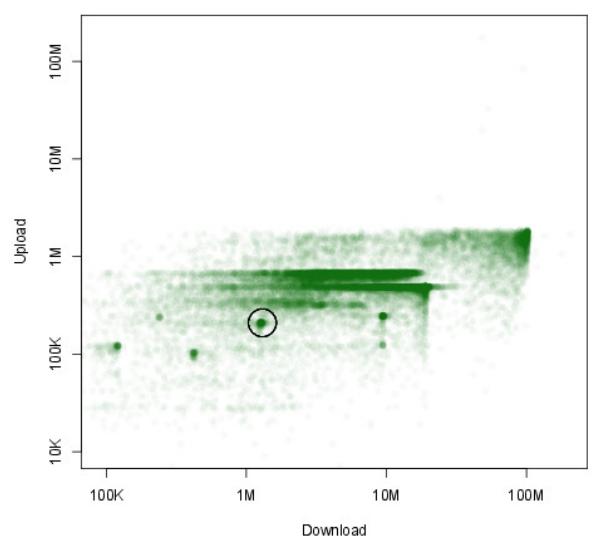

Figure 17: Scatter plot of download vs. upload throughput for Optus Internet in Australia

In almost all download/upload scatter plots, there are other high density points not connected to service plans offered by operators. Figure 11 shows two clusters below the common upload speed of 1 Mb/s and between 1 and 10 Mb/s download speeds. There are much more high density clusters in Figure 13, again mostly at low upload speeds. We intend to further investigate these data points in future work.

4.3.2 Impact of distance on latency

It is commonly assumed that as you increase geographic distance between end-points, network latency between them increases. With our dataset, we were able to provide empirical support for, and quantify that effect. We considered measurements in a radius of 600 miles from the server. Figure 18 plots the relation of latency and distance for AT&T, Charter and Comcast. We generally observe that as distance increases, so does network latency. This effect is especially true for Charter and Comcast on distances up to 150 miles. After 150 miles, both Charter's and Comcast's average latencies decrease, which is a surprising effect. There is a sudden increase in latency for measurements on AT&T's network around a distance of 100 miles. After that, the latency is almost constant up to 300 miles, where it starts increasing with distance again. An implication of these results is that although latency generally does correlate with geographic distance, network topology and connectivity of a provider clearly has an important impact on broadband performance for end-users.

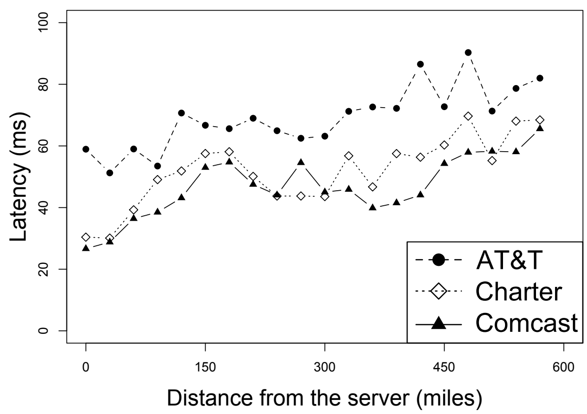

Figure 18: Impact of distance to latency for US operators.

4.4 Markets Beyond US

Figure 19 shows average download and upload speeds for all countries represented in our dataset. Note that the countries identified in the figure are from the perspective of the client host, not the Speedtest server. In cases where a Speedtest server is close to a country border, it is certainly possible for a client host to be within 200 mi and in a different country. Since we are interested in broadband access performance, we consider the country from which a test is initiated.

We observe that Hong Kong has the most developed Internet infrastructure by far with cheap and easily accessible fiber connections. Most European countries and the US have average download speeds larger than 5 Mb/s. Countries in the speed range from 2 Mb/s to 5 Mb/s are mostly South American or African. Both India and Indonesia have average download speeds lower than 2 Mb/s.

It is interesting to compare our findings with the findings of Kreibich *et al.* [15]. For most of the countries, the average speeds are twice as high in our dataset than what was reported in their study. There are also numerous relative performance differences. For example, Japan's average download speeds were reported to be almost two times faster than Hong Kong. In our measurements, we observe Hong Kong connections to be faster than those from Japan. We hypothesize that the number of users connecting over fiber in Hong Kong has increased, leading to average speeds that are now much faster than any other country.

Table 1: Number of measurements with download speed greater than 40 Mb/s and upload speed greater than 15 Mb/s broken down by servers.

Metro area	Fiber	Total
Hong Kong	428003	1476203
Budapest	110186	2197648
Tokyo	73388	641759
Los Angeles	62923	4448778
New York, NY	58075	2161745
San Francisco, CA	37944	1778681
Paris	30465	798004
Manchester	21396	2793875
Stockholm	21212	177703

On a global level, we examined fiber adoption rates in various countries and latencies of fiber connections. We also studied broadband connection infrastructural trends in developing countries.

4.4.1 Fiber adoption

For the purpose of the fiber connection analysis, we assume that every measurement with download speed greater than 40 Mb/s and upload speed greater than 15 Mb/s is over a fiber optic infrastructure.

Table 1 shows the number of high speed measurements broken down by individual metro area (which corresponds to a Speedtest server). We report only areas in which we observe at least 20,000 such measurements. In addition to Hong Kong, we also see high fiber adoption rates in Hungary, Japan, France, United Kingdom and Sweden. Note that in addition to home user fiber connections, these results also include measurements from universities and businesses.

4.4.2 Fiber connection consistency and latency

We compare operators offering fiber connections in different markets. In Hong Kong, the operator with the largest number of users is City Telekom. We compare it with DIGI, which offers fiber connection in Hungary, and Verizon FiOS, the most common high speed ISP in our US dataset.

Figures 20, 21 and 22 show scatter plots of measurements faster than 40 Mb/s download and 15 Mb/s upload for the above operators in their respective markets. Not only does Hong Kong have the largest fraction of high-speed internet users as shown above, but operators in Hong Kong market also offer higher speeds than operators from European and US markets, like DIGI and Verizon FiOS. Interestingly, City Telekom offers completely symmetric speeds for both upload and download. In prior work, Sundaresan *et al.* found that US operators provide better upload than download speed consistency. On Figure 20 for City Telekom in Hong Kong, we observe the opposite effect; there are more non-optimal measurements with upload rather than download degradation. DIGI, on the other hand, shows the same effect as US operators; very few measurements have below optimal upload, while there are lots of measurements with degraded download speeds.

We have shown in earlier sections that average download speed is lower on measurements with higher latencies for both cable and DSL operators. Is this also true for fiber connection? To study this question, we analyzed latency and download speeds from City Telekom and DIGI. We consider only measurements above 40 Mb/s download and 22 Mb/s upload speeds. We cut off measurements

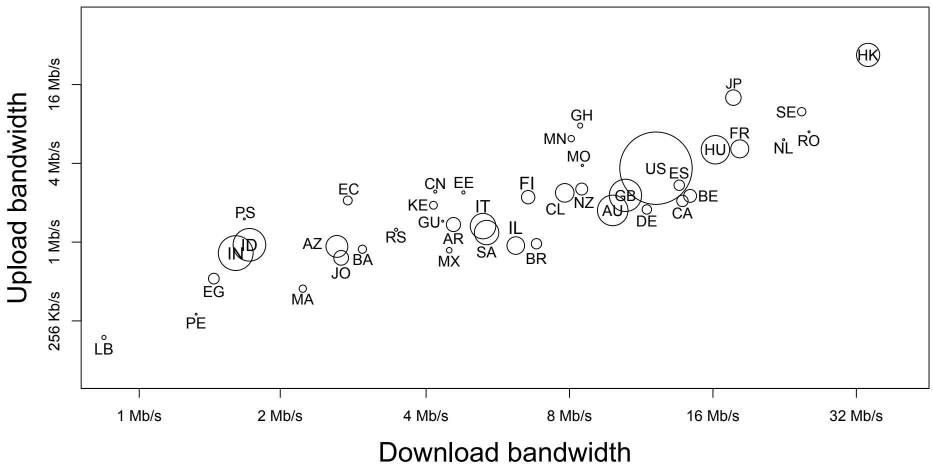

Figure 19: Average download/upload bandwidths for countries with most measurements. Countries identified are from the perspective of client hosts, not Speedtest servers. Circle areas are proportional to number of measurements initiated from that country.

with upload speeds between 15 and 22 Mb/s because we do not want to include data from the DIGI's suboptimal service plan we observed at bottom left part of the Figure 21.

Figure 23 shows both the latency distribution and the relation of average download speed to latency for DIGI and City Telekom. Users of Fiber connection see much lower latencies than users of cable and DSL operators. Most of the users have latencies in the range of 3 – 5 ms. The average download speed indeed decreases with latency. The effect is most noticeable in latency range from 1 – 5 ms, but is also true for latencies beyond 5 ms.

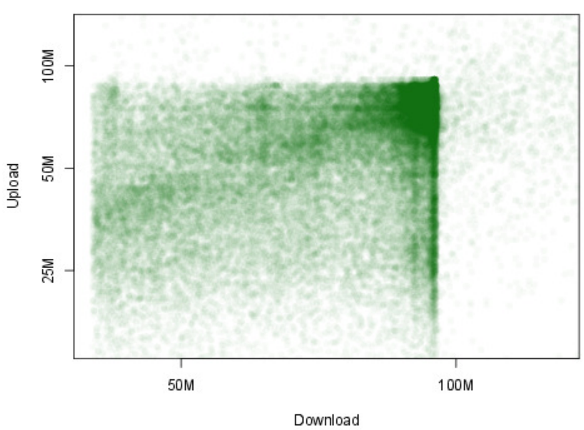

Figure 20: Scatter plot of download vs. upload throughput for City Telekom in Hong Kong.

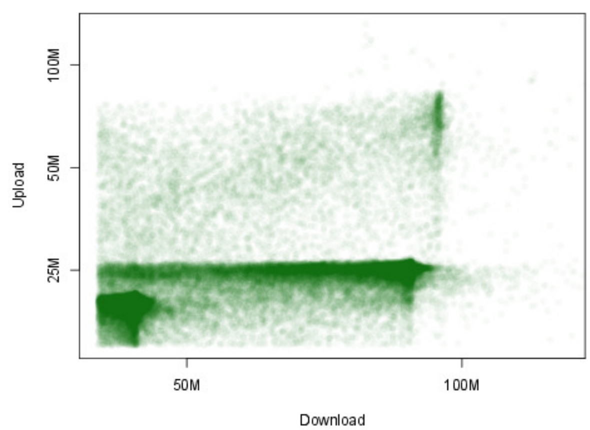

Figure 21: Scatter plot of download vs. upload throughput for DIGI in Hungary.

5. RELATED WORK

There is a growing body of studies that examine broadband access network performance. A number of earlier works focused on broadband access speeds in the context of peer-to-peer applications, *e.g.*, Lakshminarayanan *et al.* [16], and measurement methods for basic characterization of broadband access networks, such as the asymmetry in upload and download speeds, *e.g.*, Dischinger *et al.* [11] and Croce *et al.* [10]. For example, the authors of [11] found that download speeds exceed upload speeds by a factor of 10 in some cases, but that measured bandwidths matched speeds advertised by ISPs quite well at all times of day. They also found that DSL access links exhibit large latencies compared with cable modem access links. In contrast, as can be seen in the Appendix and in Figure 9, we find that instances in which there is an order of magnitude difference between download vs. upload to be the rare case and that in general, the gap has narrowed significantly.

While the aforementioned works employed active probe-based measurements, other studies have analyzed passive measurements collected from service provider networks. These measurements en-

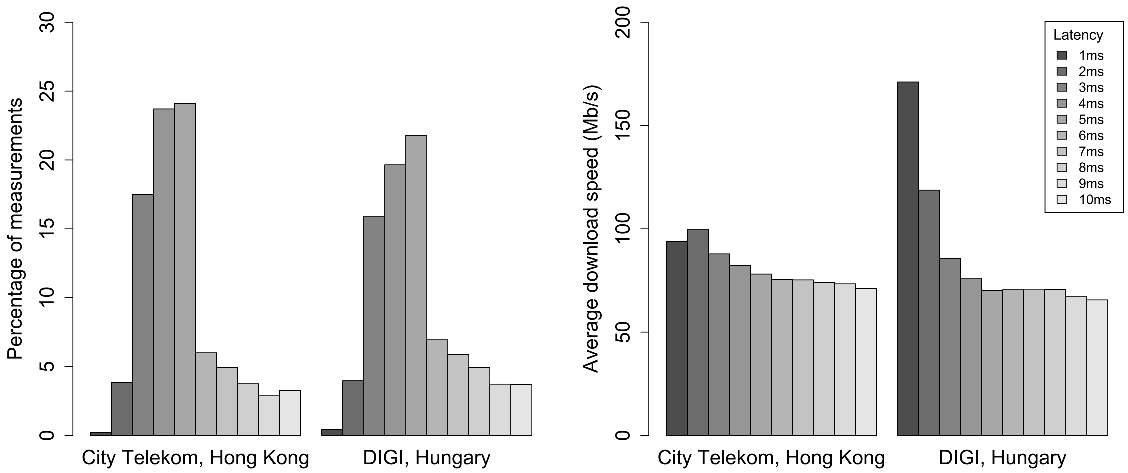

Figure 23: Latency distribution and relation of latency and average download speed for Fiber operators.

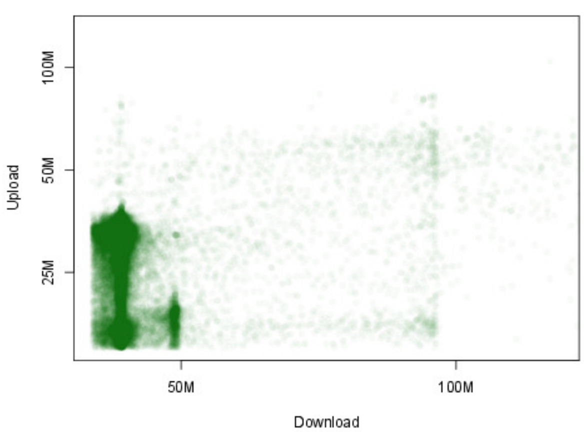

Figure 22: Scatter plot of download vs. upload throughput for Verizon FiOS users in US.

able analysis of specific application-layer behavior, which is beyond the scope of active probe-based measurements *e.g.,* from Speedtest.net. For example, the study by Cho *et al.* examined residential broadband traffic in Japan. Among other findings, their study showed that 63% of traffic was peer-to-peer, with many "heavy hitters", especially on fiber-connected access links. A more recent study by Maier *et al.* examined traffic from about 20,000 residential DSL customers from a large ISP in Europe [17]. In contrast to the earlier study in Japan, they found that HTTP traffic strongly dominates peer-to-peer traffic. They also found that delays from the home network to the ISP's gateway often exceed delays in the wide area, and that users rarely consume the full capacity of their access links (an observation also made in an earlier work by Siekkinen *et al.* [21]). Maier *et al.* suggested that users' achievable throughputs were often limited by suboptimal TCP settings.

More recent measurement studies to assess broadband access speeds and performance have been fueled, in part, by a sponsored project by the US Federal Communications Commission to assess broadband speeds and coverage, which is being managed by SamKnows [4, 9]. The recent study by Sundaresan *et al.* [22] used data collected by SamKnows for the FCC study, as well as measurement data produced by their own BISMark system, to evaluate a variety of characteristics of broadband access networks. For their study, both data sets are collected from processes running directly on the gateway router. In the BISMark system, upload and download throughputs are actively measured using a single-threaded HTTP connection, and other techniques are used to both actively and passively measure link characteristics (*e.g.,* the ShaperProbe tool is employed to measure capacity [14]). In contrast, the SamKnows throughput measurement method employs parallel TCP streams in order to be more likely to saturate the upload and download capacity. Among other issues, the authors examined ISP traffic shaping policies, differences among local access providers, and effects due to oversized buffers and effects of various modem models. As discussed in Section 4.1.1, differences between Speedtest vs. Samknows measurements are primarily due to SamKnows gateway-based deployment and cross-traffic avoidance. We make no argument that one is better than the other — simply that the provide different perspectives on broadband performance. However, Speedtest data clearly provides a broader perspective since it does not require whitebox deployment.

Similarly, work by Kreibich *et al.* on Netalyzr [15] has exposed effects of overbuffering on edge devices (*i.e.*, the "bufferbloat" problem [12]). Similar to Speedtest, Netalyzer is an applet that runs on client nodes and accesses dedicated servers to assess performance. Netalyzer differs from Speedtest by *(i)* providing measurements to users beyond latency and bandwidth, including *e.g.,* DNS response time, path MTU, and IPv6 support among others, *(ii)*, using EC2 for servers (instead of local servers), and *(iii)* using carefully constructed UDP streams instead of TCP transfers to measure bandwidth. Among other things, the study in [15] reports upload and download bandwidths in markets around the world. Their results show a more narrow range of bandwidths than observed in our data, most likely reflecting the nearly 3 year difference in data gathering.

This highlights the need for continuing analysis of broadband performance.

While devices deployed by SamKnows are available upon request by residential broadband users, many users rely on publicly available bandwidth testing services. Among these, Speedtest [3] is one of the most widely deployed. The Measurement Lab project also makes available the NPAD and NDT tools for network diagnostic and throughput testing [18]. These latter tools employ a single TCP connection for assessing throughput speeds. Bauer *et al.* examined the accuracy of various broadband access testing tools and found that while in-gateway measurement systems such as the one employed by SamKnows are the most accurate, the methods employed by the Speedtest application are also quite accurate [5]. Moreover, they found that tools that only use a single TCP connection for measuring throughputs tend not to be very accurate.

6. SUMMARY AND CONCLUSIONS

In this paper we revisit the issue of broadband performance using crowd-sourced data from `speedtest.net`. The objectives of our work are threefold: *(i)* to reconfirm prior results on broadband performance in the US, *(ii)* to expand on prior studies by investigating broadband performance details afforded by our data set, and *(iii)* to compare and contrast broadband performance in markets around the world. Our data set was collected over a 6 month period in 2011 from 59 metro markets around the world.

Our analysis of US markets focuses on reconfirmation of prior work by Sundaresan *et al.*. To begin, we report results of a case study that compares Speedtest measurements to SamKnows measurements. Our results show that there is high correlation between reported performance when Speedtest clients use wireline Ethernet, but Speedtest results are substantially lower when wireless Ethernet is used. The results of our evaluation of broadband performance in US markets are consistent with many of the prior performance studies. However, several of our results differ, for example in terms of service provider rankings. Our data reveal several additional features of broadband performance including tight clusters of lower performance (which we attribute to older service plans, older equipment or throttling), and the correlation between latency and distance to a server. Our analysis of non-US markets shows a broad spectrum of performance with Hong Kong at the high end, and India and Indonesia at the lower end.

In future work, we plan to use a broader set of Speedtest data to consider longitudinal characteristics of broadband deployments world wide. We also plan to investigate instances of anomalous conditions such as outages or step function jumps in performance that will provide a perspective on the robustness of broadband networks.

Acknowledgements

We thank our shepherd and the reviewers for their helpful suggestions and comments.

This work was supported in part by NSF grants CNS-0716460, CNS-0831427 and CNS-0905186, and CNS-1054985. Any opinions, findings, conclusions or other recommendations expressed in this material are those of the authors and do not necessarily reflect the view of the NSF.

We thank Ookla, Inc. and Andrew Bassett for generous access to the `Speedtest.net` performance data.

We thank Sam Crawford and SamKnows for the use of white-boxes in our experiments.

We also thank Anton Kapella for his fruitful suggestions and input on this work.

7. REFERENCES

[1] Ookla Documentation Wiki. http://wiki.ookla.com, 2011.
[2] Ookla, Inc. http://www.ookla.com, 2011.
[3] Speedtest.net. http://www.speedtest.net, 2011.
[4] SamKnows: Accurate broadband performance information for consumers, governments, and ISPs. `http://www.samknows.com/broadband/index.php`, 2012.
[5] S. Bauer, D. Clark, and W. Lehr. Understanding broadband speed measurements. In *38th Research Conference on Communication, Information and Internet Policy*, September 2010.
[6] S. Bauer, D. Clark, and W. Lehr. Powerboost. In *Proceedings of the ACM SIGCOMM workshop on Home networks*, 2011.
[7] R. Carlson. Network Diagnostic Tool. `http://www.internet2.edu/performance/ndt/`, February 2012.
[8] US Federal Communications Commission. National Broadband Plan. `http://www.broadband.gov/plan/`, 2011.
[9] US Federal Communications Commission. Measuring Broadband America. `http://www.fcc.gov/measuring-broadband-america`, 2012.
[10] D. Croce, T. En-Najjary, G. Urvoy-Keller, and E.W. Biersack. Capacity estimation of adsl links. In *Proceedings of ACM CoNEXT Conference*, 2008.
[11] M. Dischinger, A. Haeberlen, K.P. Gummadi, and S. Saroiu. Characterizing residential broadband networks. In *Proceedings of ACM Internet Measurement Conference*, 2007.
[12] J. Gettys and K. Nichols. Bufferbloat: Dark buffers in the internet. *Queue*, 9(11), November 2011.
[13] Google. Measurement Lab. `http://www.measurementlab.net`, 2012.
[14] P. Kanuparthy and C. Dovrolis. ShaperProbe: end-to-end detection of ISP traffic shaping using active methods. In *Proceedings of ACM Internet Measurement Conference*, 2011.
[15] C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson. Netalyzr: Illuminating the edge network. In *Proceedings of ACM Internet Measurement Conference*, 2010.
[16] K. Lakshminarayanan and V.N. Padmanabhan. Some Findings on the Network Performance of Broadband Hosts. In *Proceedings of ACM Internet Measurement Conference*, 2003.
[17] G. Maier, A. Feldmann, V. Paxson, and M. Allman. On Dominant Characteristics of Residential Broadband internet Traffic. In *Proceedings of ACM Internet Measurement Conference*, 2009.
[18] M. Mathis, J. Heffner, P. O'Neil, and P. Siemsen. Pathdiag: Automated TCP Diagnosis. In *Proceedings of Passive and Active Measurement Conference*, 2008.
[19] MaxMind. The MaxMind IP Geolocation Servicel. `http://www.maxmind.com`, 2012.
[20] S. Perez. Nielsen: Cord Cutting and Internet TV Viewing on the Rise. `http://techcrunch.com`, February 2012.
[21] M. Siekkinen, D. Collange, G. Urvoy-Keller, and E. Biersack. Performance limitations of ADSL users: a case study. In *Proceedings of Passive and Active Measurement Conference*, 2007.
[22] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè. Broadband internet performance: A view from the gateway. In *Proceedings of ACM SIGCOMM '11*, Toronto, Canada, August 2011.

APPENDIX

The following table shows summary statistics and basic information on the Speedtest measurements included in our study.

Table 2: Number of measurements, average download and upload speeds and latency broken down by metro areas.

Region	Metro Area	Country Code	Tests	Download Mean (Stdev) Kb/s	Upload Mean (Stdev) Kb/s	Latency Mean (Stdev) millisec.
Europe	Barcelona	ES	323431	13384 (18471)	2704 (7264)	76 (161)
	Brussels	BE	495738	14380 (14404)	2374 (5701)	51 (164)
	Budapest	HU	2197648	16519 (27293)	5188 (12439)	50 (168)
	Frankfurt	DE	242569	12053 (16779)	1834 (3904)	118 (211)
	Manchester	UK	2793875	10718 (14322)	2328 (6176)	56 (119)
	Paris	FR	798004	19562 (49113)	5598 (24270)	59 (149)
	Rome	IT	1740709	5401 (9747)	1365 (6760)	74 (142)
	Sarajevo	BA	212119	2981 (5738)	893 (4145)	96 (267)
	Stockholm	SE	177703	26509 (40140)	10542 (23238)	50 (405)
	Turku	FI	515003	6484 (7953)	2273 (5183)	94 (155)
Asia	Hong Kong	HK	1476203	33955 (43730)	26988 (36826)	37 (137)
	Jakarta	ID	2852275	1743 (6250)	976 (4659)	133 (327)
	New Delhi	IN	3242544	1632 (3397)	844 (2045)	165 (278)
	Tokyo	JP	641759	18085 (29084)	13022 (18372)	70 (267)
	Ulaanbaatar	MN	109354	8261 (17064)	6326 (13437)	75 (200)
Africa	Accra	GH	68753	8561 (19282)	7900 (20609)	114 (291)
	Bamako	ML	19	244 (145)	143 (117)	489 (507)
	Cairo	EG	308805	1468 (3910)	539 (2771)	132 (347)
	Casablanca	MA	141558	2256 (3640)	451 (1933)	173 (240)
	Nairobi	KE	164759	4237 (11049)	1955 (6469)	164 (338)
Pacific	Auckland	NZ	381377	8689 (16535)	2605 (9791)	58 (167)
	Melbourne	AU	961180	12285 (18398)	1416 (4372)	56 (138)
	Perth	AU	353578	6254 (18409)	2028 (11248)	72 (168)
	Sydney	AU	1184639	9464 (17653)	2019 (9378)	50 (129)
	Tamuning	GU	12803	4137 (5869)	1384 (3899)	115 (198)
North America	Anchorage, AK	US	61144	4948 (7038)	1672 (6360)	144 (504)
	Bellingham, WA	US	433021	11978 (12793)	2436 (5181)	61 (153)
	Burlington, VT	US	101058	16554 (36229)	3963 (8671)	115 (249)
	Chicago, IL	US	1365027	13487 (22697)	4174 (12340)	53 (134)
	Dallas, TX	US	815938	10824 (17556)	3717 (9973)	76 (173)
	Flagstaff, AZ	US	55350	11822 (17760)	3260 (6548)	103 (197)
	Grand Forks, ND	US	84126	11774 (14269)	3192 (5425)	74 (161)
	Grand Rapids, MI	US	136568	7432 (10921)	2506 (5258)	95 (168)
	Honolulu, HI	US	127632	8536 (9948)	1834 (4526)	94 (130)
	Idaho Falls, ID	US	37079	6904 (12399)	2297 (4915)	131 (209)
	Lawrence, KS	US	170725	10525 (24911)	3553 (10067)	261 (432)
	Lexington, KY	US	368759	9800 (10433)	1887 (4865)	63 (140)
	Little Rock, AR	US	173791	7279 (9936)	1981 (4205)	90 (161)
	Los Angeles, CA	US	4448778	11377 (18100)	3215 (10473)	54 (139)
	Medford, OR	US	51780	13180 (15873)	3266 (10605)	74 (159)
	Miami, FL	US	740662	15977 (16262)	4114 (10271)	39 (112)
	New York, NY	US	2161745	16749 (20415)	4584 (8367)	50 (139)
	Pensacola, FL	US	143703	7916 (9085)	1961 (3617)	96 (145)
	Philadelphia, PA	US	489278	12736 (18220)	5474 (10544)	77 (189)
	Portland, ME	US	152908	9709 (11956)	2713 (6447)	91 (252)
	Provo, UT	US	31563	8648 (19535)	4427 (10258)	98 (189)
	San Francisco, CA	US	1778681	12384 (22480)	4764 (13072)	55 (147)
	Springfield, MO	US	458499	7561 (13952)	2047 (9096)	75 (143)
	Sumter, SC	US	78590	10560 (17325)	2460 (5852)	126 (234)
South America	Brasilia	BR	286197	6983 (8207)	995 (2751)	66 (180)
	Buenos Aires	AR	528919	4713 (9768)	1407 (5386)	72 (169)
	La Paz	BO	44925	752 (1428)	295 (1005)	315 (341)
	Quito	EC	213365	2800 (7660)	2123 (5885)	129 (294)
	Santiago	CL	935391	7928 (13115)	2415 (8278)	92 (359)
Middle East	Amman	JO	598767	2672 (4868)	759 (3263)	183 (292)
	Baku	AZ	1280475	2658 (4955)	948 (3383)	87 (196)
	Kabul	AF	1998	768 (1205)	585 (2025)	409 (689)
	Riyadh	SA	1595746	5487 (10381)	1212 (6274)	76 (175)
	Tel Aviv	IL	870621	6058 (6031)	943 (3133)	78 (197)