Second Chance: Understanding diversity in broadband access network performance

John P. Rula Zachary S. Bischof Fabián E. Bustamante
Dept. of EECS, Northwestern University
Evanston, IL
{john.rula, zbischof, fabianb}@eecs.northwestern.edu

ABSTRACT

In recognition of the increasing importance of broadband, several governments have embarked on large-scale efforts to measure broadband services from devices within end-user's homes. Participants for these studies were selected based on features that, a priori, were thought to be relevant to service performance such as geographic region, access technology and subscription level. Every new-year deployment since has followed the same model, ensuring that the number of measurement points remains stable despite the natural churn.

In this paper, we start to explore the issue of vantage point selection in residential broadband networks by leveraging the publicly available datasets collected as part of the FCC Broadband America study. We present the first analysis of the variation of performance in edge networks and diversity of individual vantage points. We explore the underlying causes of this diversity through a factor analysis of contextual factors within an ISP such as the geographic location of subscribers. The goal of this analysis is to inform additional deployments in ongoing studies, and guide the design and deployment of future investigations into broadband networks.

Categories and Subject Descriptors

C.2.5 [**Communication Networks**]: Local and Wide-Area Networks; C.4 [**Performance of Systems**]: Measurement techniques

Keywords

Broadband networks, measurement techniques

C2B(1)D'15, August 17-21 2015, London, United Kingdom

 ISBN 978-1-4503-3539-3/15/08…$15.00

DOI: http://dx.doi.org/10.1145/2787394.2787400

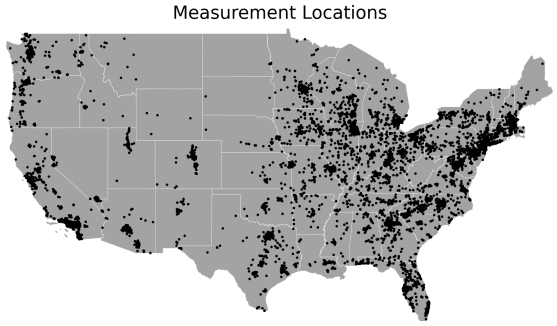


Figure 1: Map of all measurement locations used in FCC Broadband America study circa 2012.

1. INTRODUCTION

In recognition of the increasing importance of broadband access in relation to education and economic growth, many governments have launched or funded programs that aim to survey broadband availability and characterize the performance of these services. A common approach to studying broadband access is to distribute instrumented devices to end users within their home network [8]. For example, the company SamKnows in collaboration with government organizations in the US, the EU, Singapore, Brazil, and Canada [8] has begun distributing residential gateways to thousands of broadband subscribers for national broadband characterization. These devices measure the performance of selected users' broadband connections through active and passive measurements.

Measurement locations for these studies are typically 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 1, were selected from a pool of over 145,000 volunteers [4] based on features that, a priori, were thought to be 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 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.

From the relatively small scale of these efforts, the validity of the inferences made from such studies will ultimately depend on the integrity of the sampling procedure used and the features considered. For instance, the FCC's National Broadband Plan is only able to monitor 10,000 end-hosts, a small fraction (0.01%) of the over 87 million broadband customers in the United States alone.

In this paper, we start to explore the issue of vantage point selection in residential broadband networks by leveraging the publicly available datasets collected as part of the FCC Broadband America study. We present the first analysis of the variation of performance in edge networks and diversity of individual vantage points. We explore the underlying causes of this diversity through a factor analysis of contextual factors within an ISP such as the geographic location of subscribers. The goal of this analysis is to inform additional deployments in ongoing studies, and guide the design and deployment of future investigations into broadband networks.

The remainder of this paper is structured as follows. We describe the FCC Broadband America study and the specifics of its dataset used in our study in Section 2. We present our analysis of broadband access network diversity in Section 3 and our factor analysis of contextual elements on experiences performance in Section 4. Section 5 presents relevant related work, and Section 6 concludes with a discussion and future work.

2. BACKGROUND AND DATASET

In this section we give an overview of the ongoing FCC Broadband America study whose goal is to characterize the state of residential broadband connectivity in the United States. We describe the publicly available dataset used in our study and detail the specific performance metrics we used in our analysis.

2.1 FCC Broadband America

In 2011, the FCC in cooperation with SamKnows, distributed instrumented home routers, which conduct and report network measurements, to US broadband customers. The stated goal of the FCC Broadband America study is to:

> *"serve as consistent benchmarks on the health of broadband Internet access services in the United States, and to better chart progress towards the FCC goal of continuing to evolve the speeds and quality of service at which broadband access is commonly available to the American public"* [4].

The study initially placed 10,000 instrumented routers at locations across the United States. These routers were placed using *stratified sampling* to select individual placement from a pool of 145,000 volunteers. In stratified sampling, a population is first segmented into separate groups with the goal of separating a population into homogeneous subgroups. Individual samples are randomly selected from within each group. Volunteers of the FCC study were grouped into subpopulations based on each participant's geographic area, including region of the country, state and city, as well as the subscriber ISP and subscription capacity. The number of samples taken from each group were proportional to the national market share of broadband subscribers they represented.

2.2 Dataset

For our study, we utilized data collected during January 2012. While public data is available between 2011 to 2013, our goal is to analyze the diversity in broadband performance at any particular snapshot, and not to conduct a longitudinal analysis.

Due to the scope of this paper, we limited our analysis to two major performance metrics: **landmark latency** and **packet loss**. These metrics were chosen due to their ubiquity as common network performance characteristics and their importance in residential broadband quality of experience [11].

As part of its continued experimentation, each router sends up to 600 UDP packets to preselected measurement servers every hour; less if the link is in use for part of the hour. These servers are hosted either internally by the hosting ISP or at external locations hosted by the MLab. Results are reported as hourly statistical summaries of latency measurements (e.g. mean, min, max, and standard deviation). We restricted our analysis to measurements targeting servers located within the ISP, identified by the server domain, to prevent any interference from cross-ISP links or intermediary networks.

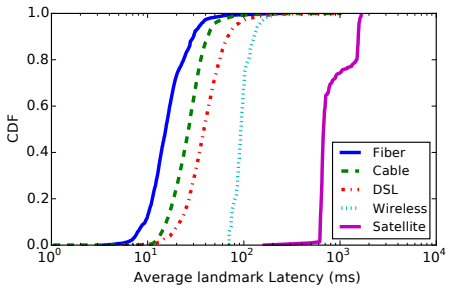
Landmark Latency. Each of the UDP packets sent are echoed back to the client by the measurement server, with their round trip time representing the latency between clients and each measurement server. We qualify these measurements as landmark latency, since they reflect latency measurements to fixed landmarks, and also to clarify our measurements against other common latency quantities used in broadband characterization such as last-mile latency.

Packet Loss. The loss rate of a residential broadband network has a significant impact on the quality of experience and usability of broadband access links. High rates of packet loss can severely limit link throughput and greatly increase website page load times. From the series of UDP packets sent, the number of missing responses is used to calculate the packet loss rate of the network during that hour.

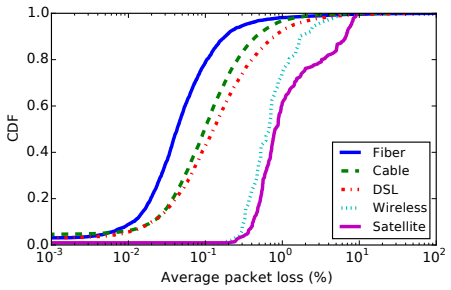
In the following sections, we use the outlined metrics from this dataset to analyze the diversity in performance seen across an ISP, and attempt to discern the underlying factors that contribute to any performance variance.

3. DIVERSITY IN BROADBAND NETWORKS

We now analyze the diversity of performance experienced across an ISP for various network performance metrics. While this work represents only a preliminary look, we seek a fundamental understanding of the nature of diversity and heterogeneity in broadband access network performance. Understanding the causes and distribution of this diversity can help inform sampling procedures and vantage point



(a) Landmark latency.



(b) Packet loss.

Figure 2: Performance characteristics of all users across broadband networks in the FCC study grouped by access technology.

selection in future characterization studies. We investigate the diversity of ISP performance across three levels (i) performance variation across the entire ISP (ii) the impact of individual vantage points on the overall performance diversity and (iii) the impact of random sampling on the measured performance of an ISP.

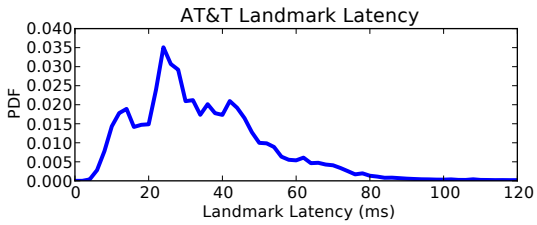
3.1 Characterizing Network Diversity

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.

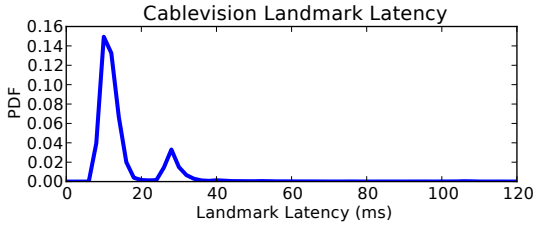
The effect of this growth pattern is a large amount of diversity within the performance of a large ISP's network. Each smaller ISP acquired differed in the quality of their design and construction of the underlying infrastructure. As we see in Figure 2, the underlying access technology can be an indicator of the received internet performance.

The figure, which displays all users in the FCC Broadband America study grouped by access technology, shows the distinct performance characteristics between different technologies (Fig. 2a), 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.

Performance heterogeneity is also visible within the same technology and the same provider. Figure 3 shows the probability distribution of landmark latency for AT&T and Cablevision's networks. As the figure shows, the diversity in ISP performance can manifest itself through either a wide variance in overall performance (Fig. 3a), or through distinct modes of performance as is the case with Cablevision (Fig. 3b). These modes indicate clustered groups of performance within an ISP; understanding the source of these modes is critical to informing the selection of vantage points in future deployments. In the next section we explore whether the source of this diversity is between vantage points or within the vantage points themselves.



(a) AT&T – DSL



(b) Cablevision – Cable

Figure 3: RTT performance in AT&T and Cablevision's broadband networks. Performance diversity can manifest itself through either a wide variance in performance as seen in AT&T, or through distinct modes of operation as seen in Cablevision's network.

3.2 Diversity in Vantage Points

We next investigate the variance in performance between vantage points, and how this affects overall diversity of an ISP's network. We use this to infer whether the source of diversity originates from within each vantage point (e.g. large individual variation) or between vantage points.

To see the impact of individual vantage points, we compare the distribution of overall latency from all measurements to the distribution of vantage point mean performance. The difference between these two distributions give insights into the source of ISP diversity. For instance, given the multiple performance modes observed in the previous section, if each vantage point experiences a large variance in performance, and thus accounts for overall ISP performance,

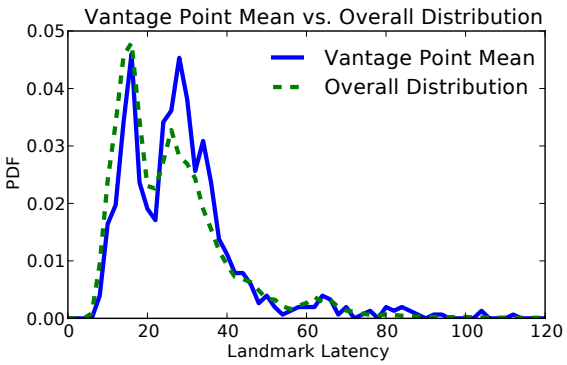


Figure 4: Distribution of vantage point mean RTT for vantage points in Time Warner's broadband network compared to the overall distribution seen from all measurements from all clients. The mean vantage point latency closely resembles the overall distribution, with a KS distance of 0.125.

then plotting the distribution of each vantage points average metric would reduce the number of visible modes.

We employ the Kolmogorov-Smirnoff distance to test the similarity between distributions. The Kolmogorov-Smirnoff distance (KS-distance) is defined as the maximum distance between the cumulative distribution functions of two distributions, and can be used as a non-parametric significance test between distributions. This allows us to measure the similarity between two distributions with multiple modes of operation, and in a normalized fashion.

To illustrate the impact of KS-Distance values, Figure 4 shows the probability distributions for mean vantage point latency compared to the overall latency distribution. The mean vantage point latency distribution is nearly identical to the overall performance distribution in the case of Time Warner cable and is able to completely capture the three visible modes of performance at 19, 35 and 65 milliseconds. The KS distance between the two distributions is 0.125.

Using this distance between mean vantage point, and overall distributions, we find broadband network diversity originates from the diversity in individual vantage points. The distribution of vantage point averages closely resembled the overall distribution for nearly all ISPs studied. Table 1 displays the KS distance for these two distributions, showing small KS distances for nearly all ISPs studied.

3.3 Sampling Variance

In the following paragraphs we look at the variation in performance obtained through random sampling from the known population of FCC vantage points.

To measure the potential variance due to sampling, we randomly selected groups of vantage points, with sample sizes at powers of 2 intervals, from the entire population of FCC vantage points. For each sampled set, we compared the distance between the sampled distribution and the overall performance distribution using the Kolmogorov-Smirnov (KS) distance. We repeated this process 1000 times for each increasing power of 2 less than the size of the population.

Provider	Technology	KS Distance
AT&T	DSL	0.117
CenturyLink	DSL	0.155
Qwest	DSL	0.156
Verizon	DSL	0.150
Windstream	DSL	0.168
Brighthouse	Cable	0.185
Cablevision	Cable	0.255
Charter	Cable	0.151
Comcast	Cable	0.178
Cox	Cable	0.196
Insight	Cable	0.246
Mediacom	Cable	0.183
TimeWarner	Cable	0.125
Frontier	Fiber/DSL	0.221
Verizon	Fiber	0.195
Clearwire	Wireless	0.195

Table 1: Distance between overall distribution of landmark latency and the average vantage point latency of each subscriber for each provider in the study. Satellite providers were removed due to the small numbers of total vantage points.

Figure 5 shows the variance in random samples of a population for a representative ISP for each access technology. In the figure, each marker denotes the average KS distance for each of the 1000 runs, and the error bars representing the standard deviation of each iteration. The steep rate of convergence for each ISP can be seen in the figure. For instance, for the fixed line broadband providers, subsampled distributions are able to approach KS distances of 0.2 at approximately a 1% sampling, and experience greatly diminishing returns after the sample reaches 10%. This is not the case for wireless provider Clearwire, whose sampled distributions are significantly higher than other providers even at a 50% sample rate.

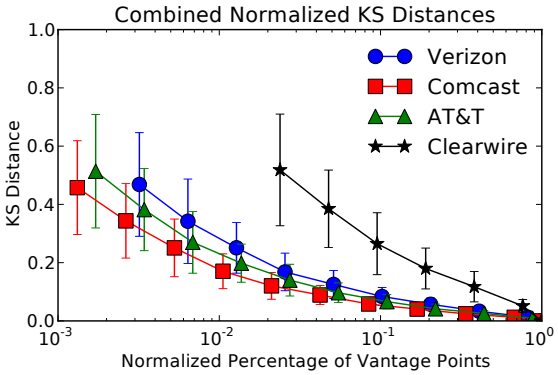


Figure 5: Kolmogorov-Smirnov distances for a random sampling of a normalized percentage of an ISPs population. A representative ISP for each technology is shown in the figure. Each marker indicates the average distance of 1000 runs and the errors bars the standard deviation of the set.

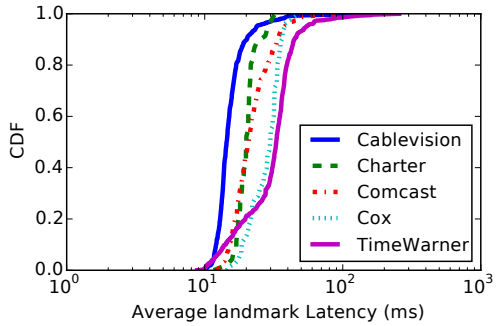


Figure 6: Average landmark latency for cable service subscribers in the Northeast with a capacity of at least 10 Mbps.

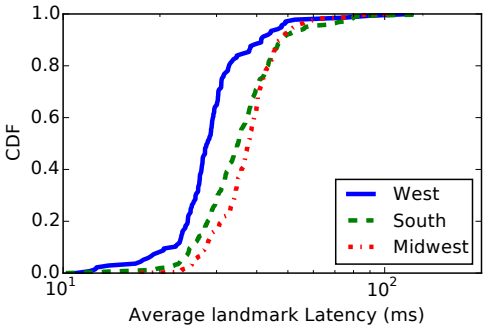


Figure 7: Average landmark latency for AT&T subscribers with a capacity of at least 10 Mbps.

4. FACTOR ANALYSIS

We conducted a factor analysis of the features chosen for the stratified sampling used in the FCC Broadband America study. The strata used include the **service provider**, **access technology**, **geographic region**, and **subscription capacity**. For service capacity, using each participant's subscription speed, k, volunteers were split into following three bins: $k < 3Mbps$, $3Mbps \leq k < 10Mbps$, and $k \geq 10Mbps$. In our analysis of the how these strata affect performance, we limit our focus to their impact on service latency and packet loss rates.

For our study, we use techniques similar to those in [1, 6] to account for confounding factors. We first divide users into disjoint sets based on attributes of their broadband connection. Each connection is described using a tuple in the form of $(provider, access\ technology, capacity, region)$. To quantify whether a factor affects performance, we compare groups of users that match on all but the factor of interest. For example, to test whether a subscriber's region could affect performance, we compare two groups of subscribers; all users in both groups have the same ISP, the same access technology, and similar capacity but differ in region. This is then repeated for each similar pair of sample groups. Note that we exclude sets with a small number of users (less than 10) from these trials since they were unlikely to have statistical significance.

To determine whether or not any difference between the two groups could be due to chance, we use a two-tailed test to measure statistical significance. We again use the KS-distance to quantify the distance between two distributions. For this analysis, we consider a p-value smaller than 0.05 to be strong evidence against the null hypothesis that the two distributions are the same.

Service Provider. We first look at how a user's broadband provider affects service latency. In our analysis, we found that a subscriber's provider provided the highest information gain for predicting a user's network latency, closely followed by the type of access technology.

Next, 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 6 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. These two distributions had a KS-distance of over 0.70 with an extremely small p-value (less than 10^{-100}). The average KS-distance over all pairs within this group was 0.55.

Overall, we found that 75% of all comparisons between similar sets of subscribers had a p-value less than 0.05, meaning that different ISPs in the same region with the same access technology frequently showed statistically significant differences in latency performance.

On the other hand, there were also cases that showed little variation across providers. DSL users in the South region with speeds of less than 3 Mbps (not shown) showed the least diversity across ISPs. In this case, the average KS-distance was 0.25.

Access Technology. As shown in Sec. 3.1, there is a clear trend of both latency and packet loss being driven by access technology. 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.

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. In the case of Verizon, the difference between the distributions of these two technologies was great; the median latency nearly tripled from 15 ms to 40 ms with a KS-distance of 0.89 between these two groups.

Geographic Location Location can also play a large role in an ISP's performance, particularly for nationwide ISPs. As mentioned in Sec. 3.1, 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 7 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 (a KS-distance of 0.188), they both differ significantly from the West region – both have a KS distance above 0.40.

Overall, we observed a statistically significant difference in performance in 64% of comparisons when controlling for confounding factors. However, the average KS-distances of 0.29 was relatively low compared to the other factors in our analysis.

Subscription Rate Last, we look at the effect a user's subscription speed has on the connection's latency and packet loss performance. We found that 64% of comparisons between similar sets of users showed a statistically significant difference across capacity tiers. Interestingly, we found that as service capacity increased, the average network latency either decreased or stayed the same. We believe that this is likely due to the fact that download and upload capacity typically increased together and that the increased upload capacity minimizes the time waiting for transmission, resulting in a decreased latency. However, when looking at packet loss rates, only 28% of comparisons showed a significant difference.

5. RELATED WORK

In the following paragraphs we review the relevant literature looking at measurement location and bias in networks as well as previous attempts to measure and characterize residential broadband.

Active and passive monitor locations for network measurements. The location of active and passive monitors in networks has been studied extensively. One class of problem similar in nature to ours is the beacon placement problem within network tomography [5, 7]. Network tomography attempts to infer characteristics of the larger network from a collection of distributed monitors. Our goal is rather to characterize the overall performance of an end-host on a network edge, at the level of an ISP or autonomous system (AS) level.

Other avenues of research include the placement of passive monitors for recording and sampling network flows [3]. Suh et al. [9] studied the placement of these passive flow monitors within networks, discovering the NP-hardness of the problem, as well as the few total number of monitors needed to successfully record all flows within a network. The placement of network flow monitors are able to capitalize on common links traversed. Our work, in contrast, attempts to select measurement locations only from the network edge.

Residential Broadband Characterization. In addition to the governmental efforts by SamKnows and governmental agencies, several research efforts have looked into characterizing the quality and performance of residential broadband services. Sundaresan et al. [10] deployed instrumented home gateways to volunteer participants to perform network experiments through their BisMark system. Bischof et al. [2] attempted crowdsourcing residential broadband characterization through the use of network intensive applications such as BitTorrent clients. These efforts, along with the ongoing government studies we analyzed in this paper, have provided a first look into the performance of residential broadband connections. Our goal is to leverage these initial research efforts to inform the next generation of research efforts.

6. CONCLUSION AND FUTURE WORK

This work examined the diversity within residential broadband networks, highlighting the importance of participant selection in existing, and future, deployments. We explored the underlying factors that cause heterogeneity in access networks, including variations across technologies and geographic regions within the same provider. We plan on expanding our analysis of network and vantage point diversity, leveraging additional broadband datasets and performing longitudinal analysis, with the goal of informing future vantage point selection through a principled approach.

As broadband services continue to grow in importance, attracting the attention of users and policy makers, there is a pressing need to determine the most appropriate metrics, measurement approaches and sampling strategies to help us derive a meaningful picture of their state. This work is a first but critical step toward that goal.

7. REFERENCES

[1] Z. S. Bischof, F. E. Bustamante, and R. Stanojevic. Need, want, can afford – broadband markets and the behavior of users. In *Proc. IMC*, November 2014.

[2] Z. S. Bischof, J. S. Otto, M. A. Sánchez, J. P. Rula, D. R. Choffnes, and F. E. Bustamante. Crowdsourcing ISP characterization to the network edge. In *Proc. of W-MUST*, 2011.

[3] G. R. Cantieni, G. Iannaccone, C. Barakat, C. Diot, and P. Thiran. Reformulating the monitor placement problem: Optimal network-wide sampling. In *Proc. ACM CoNEXT*, 2006.

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

[5] J. D. Horton and A. López-Ortiz. On the number of distributed measurement points for network tomography. In *Proc. IMC*, 2003.

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

[7] R. Kumar and J. Kaur. Efficient beacon placement for network tomography. In *Proc. IMC*, 2004.

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

[9] K. Suh, Y. Guo, J. Kurose, and D. Towsley. Locating network monitors: complexity, heuristics, and coverage. *Computer Communications*, 29(10), 2006.

[10] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè. Broadband internet performance: a view from the gateway. In *Proc. ACM SIGCOMM*, 2011.

[11] S. Sundaresan, N. Feamster, R. Teixeira, and N. Magharei. Measuring and mitigating web performance bottlenecks in broadband access networks. In *Proc. IMC*, 2013.