An Eight Years Perspective on the Internet Broadband Infrastructure in the USA

Justus Fries*, Trinh Viet Doan*, Rohit Gupta*, Vaibhav Bajpai†
*Technical University of Munich, Germany
[justus.fries | trinhviet.doan | rohit.gupta]@tum.de
†CISPA Helmholtz Center for Information Security, Germany
vaibhav.bajpai@cispa.de

***Abstract*—The broadband market in the US consists of a variety of access technologies and Internet Service Providers. However, the lack of broadband Internet access in remote regions and lack of ISP choice at high-speed tiers (above 100 Mbps) result in a digital divide that the Federal Communications Commission (FCC) has been aiming to close. To this end, we analyze a dataset collected by the Measuring Broadband America (MBA) program to provide a comprehensive view of broadband performance (reliability, throughput, and latency) across the US from 2012–2019. We also build coverage maps on reliability, throughput, and latency to identify potential underserved areas that upcoming Low Earth Orbit (LEO) satellites can cover to improve market options and diversity in the near future. Throughput speeds and latencies have improved over the years, although the observed throughput is lower in some specific states. The data shows that geostationary satellites can already serve as an alternative with reliable download speeds in areas where coverage or market competition is lacking, despite the inherently higher latency.**

I. INTRODUCTION

In this paper, we seek to provide an empirical grounding on the state of digital divide in one of the most developed countries in the world, the USA, for which broadband data at the grass-roots level is readily available. In the US, the broadband market offers a diverse mix of wireline access technologies, such as cable, DSL, or fiber but also wireless technologies such as fixed wireless, 4G and 5G, or geostationary satellite. However, in the market structure, Internet Service Providers (ISPs) may hold regional monopolies [1], [2], which eventually result in a lack of fitting service options for customers. According to the "Internet Access Services Report" by the Federal Communications Commission (FCC) for late 2018 [3], 81% of census blocks in the US have a choice between three providers, with 18% only being able to choose between two providers when considering speeds of at least 25 Mbps downstream. Additionally, for higher download speeds of at least 100 Mbps, 28% of the census blocks have no provider offering such a service speed, while 44% only have a single choice. This, along with other reports [4] that show similar results, indicates that the broadband market in the US lacks alternatives and competition, especially concerning higher speed tiers.

In addition, there is a lack of (high-speed) broadband Internet access in some US regions, primarily rural areas and tribal lands but also in urban areas, resulting in a digital divide. The 2020 Broadband Deployment Report by the FCC [5] shows improvements through fiber, LTE, and 5G deployments that may help close the digital divide in the future. For instance, between 2016 and 2018, the number of users in the US with no access to broadband with 25 Mbps downstream and 3 Mbps upstream (25/3 Mbps) has declined by 30%, with nearly every user in the US having access to fixed services (including satellites) of such speeds as of December 2018. If satellite is excluded, the number of users with broadband access drops to only around 75% of the population in rural areas and tribal lands. Note that the estimates of the report rely on self-reported data by ISPs (through FCC Form 477 [6], [7]), which are aggregated to census blocks. Therefore, the underlying data over-represents the actual coverage achieved in reality, making it difficult to claim that the digital divide is continuing to close over time [5]. Further, global emergencies that cause surges in Internet traffic such as the COVID-19 pandemic show that Internet access has to be improved still, as the shift to work from home and remote schooling has widened the digital divide, especially in already underserved areas.

Platforms that measure the broadband performance of home users can help in understanding and closing this gap. The Measuring Broadband America (MBA) program [8] has deployed around 8k SamKnows probes [9] in home networks across the country, covering most of the US broadband market [10]. These probes have performed active measurements regularly to assess the state of broadband since 2011. Although the FCC publishes reports and the collected data every single year, and previous work has analyzed shorter snapshots of the data [11], [12], [13], [14], a comprehensive longitudinal view with a focus on digital divide is missing. To close this gap, we analyze more than 1.7 TB of broadband data collected by MBA (§ I) between 2012 and 2019 with two main goals:

Goal 1 – is to provide a historical view of performance characteristics, namely reliability, throughput, and latency, to understand how the fixed-line broadband infrastructure is evolving and how well it would compare in light of upcoming 5G deployments. While reports by the FCC primarily cover changes from year to year, we focus on providing a longitudinal view over the decade as well as including metrics beyond broadband speed. We further quantify the differences between access technologies.

ISBN 978-3-903176-48-5

Goal 2 – is to build reliability, throughput, and latency coverage maps to identify underserved areas, where high-speed and low-latency broadband access is lacking. Such maps can help both the FCC and ISPs in determining underserved areas to both extend the reach of current deployments and to fill these gaps with upcoming Low Earth Orbit (LEO) satellite [15], [16], [17], [18], [19] deployments.

In pursuit of these goals, we find that over the course of the years between 2012 and 2019 (§ III), broadband performance in the US regarding download speeds has improved. Throughput has increased significantly across all access technologies (median: 17.2 Mbps in 2012 → 70.6 Mbps in 2019). However, as of 2019, only about 35% of all samples achieve download speeds at higher speeds above 100 Mbps. Further, the latency has decreased over the years (median: 26.8 → 22.5 ms), especially under load: We observe reductions of latency under load by 38.7% (median: 107.6 → 66 ms), likely due to improved Active Queue Management (AQM) measures. The results show that when comparing loaded and unloaded conditions, the inflated latency has regressed over time, i.e. the impact of bufferbloat has decreased. In terms of reliability, we find low packet loss and low download failure rates in general, although values vary over the years and across different access technologies. In particular, cable and fiber perform the best, whereas DSL is on the lower end regarding reliability, throughput, and latency.

Although we find high-speed cable access (above 100 Mbps) being available in most US states (§ IV), some states in the Midwest, the Northeast, South and Southeast can only choose from a small number of ISPs, which indicates regional market dominance. States where DSL is the primary access technology, appear to suffer from worse broadband performance. We use auxiliary data from FCC Form 477 [7], [6] as a ground truth to identify ISPs and regions that are underrepresented in the MBA program for the deployment of future measurement probes. For more details, we refer to our study's code repository[1] to facilitate reproduction [20] of analyses and future work.

II. MEASURING BROADBAND AMERICA

In this section, we describe the measurement platform and how the FCC validates the dataset. We also evaluate the representativess of the collected dataset and describe metrics selected for our study.

Measuring Broadband America (MBA) is a program by the FCC that measures broadband connectivity in the US since 2011. It deploys SamKnows [9] probes in the homes of volunteers, with around 8k probes all across the US to cover more than 80% of the US broadband market [10]. All probes that are included in the annual FCC MBA reports support the home network's maximum throughput speeds, with most probes (83%) supporting speeds of up to 1 Gbps [21].

The probes are distributed across the country based on state and region market share statistics of ISPs in the US [10]. These statistics are provided by the FCC's Form 477 data [6], [7], which requires all ISPs to file data about their offered services twice a year. The data is updated and complemented with voluntarily data provided by ISPs. The MBA program further considers the four Census regions (Northeast, Midwest, South, West) and population density for the probe deployment, as well as different speed tiers (up to 3 Mbps, between 3 and 10 Mbps, above 10 Mbps). As a result, the probes are distributed over 48 states and connected via different broadband access technologies. 8,631 probes were initially deployed across 49 states, with 7,322 probes remaining active as of 2019. The data has been collected for 26 ISPs over the course of the program.

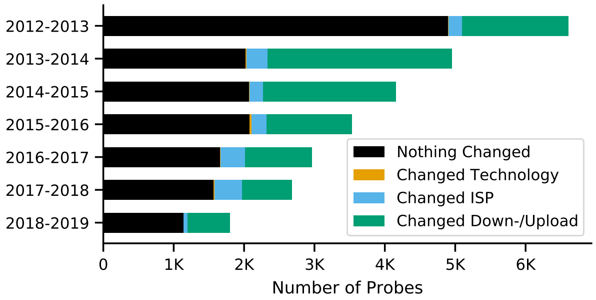

Fig. 1: Changes in probe properties based on metadata.

Fig. 1 shows the changes of probes regarding ISP, access technology, and throughput (based on differences in validated metadata) between two consecutive years. Overall, most probes do not experience any changes, although some probes change throughput speeds, while a very small fraction changes ISPs (due to rebranding and acquisitions) or access technologies.

The SamKnows probes measure a variety of broadband metrics [22] over IPv4. The tests [22] covered in this paper include: multi-threaded download speed, UDP latency, UDP latency under load, and UDP packet loss. The tests connect to the closest measurement servers (based on round-trip time); measurement servers are hosted both off-net (outside of ISP boundary) and on-net (within ISP boundary). The off-net measurement points were hosted by Measurement Lab (M-Lab) [23] and changed to Level 3 Communications (CenturyLink) across ten cities in the US, while the on-net measurement targets are hosted by the ISPs themselves. The FCC publishes the raw data for each month [24], along with annual reports on major findings [25]. For controllability reasons, the annual MBA reports only cover measurements to the off-net locations, i.e., latency results might be inflated due to geographical distance. Thus, we also consider the on-net measurements in our analysis.

Validated Data — is published and released by the FCC, covering a month in a year, along with the yearly published report. This Validated Data consists of probe specific information, such as its ISP, access technology, advertised throughputs, and the region hosting the probe along with the probe's census block, which corresponds to the approximate deployment location of the probe. Mergers and acquisitions between ISPs as well as access technology changes are reflected in the Validated Data, e.g., Charter merging with Time Warner Cable. Dunna *et al.* [26] further present an approach to sanitize the MBA metadata using auxiliary data filed by the FCC.

---

[1] https://github.com/justus237/ifip-net-2022-fcc-mba

Due to AT&T leaving the MBA program for 2019 and satellite ISPs leaving for 2018, Validated Data were carried forward, since probe measurements are not stopped but rather continue to be carried out in less intrusive ways. For both, certain filters were used in order to exclude probes that likely changed their ISP. If a probe appears in newer Validated Data, it is not carried forward. For satellite ISPs this is implemented by looking at the latency characteristics of the probes over time. If a probe exhibits latencies above 500 ms for a certain amount of time, it is excluded from the carried over data. No filters based on download speeds were used. For AT&T probes, a download speed change filter was used instead. If a probe's average download speed changed by more than 15 Mbps, it was removed. The advertised download speeds confirm this.

Representativeness — Using FCC Form 477 [6] data, which provides information on broadband service deployments and offerings in the US, we provide a first-hand approximation of the representativeness and coverage of the MBA Validated Data. Note that while Form 477 covers more states than the MBA dataset, it only provides information on the census blocks that an ISP can potentially provide Internet access in, not necessarily how many households actually have connectivity. We use the Form 477 data on deployments together with population estimates and compare it to the probe metadata from the MBA dataset. The number of states covered by each of the ISPs in the MBA dataset and in the Form 477 data is shown in Table I. All providers cover less states in the MBA dataset compared to reported data for Form 477. The biggest difference can be seen for satellite ISPs: Hughes does not have data for 15 states, while Wildblue/ViaSat does not for 30 states. However, satellite providers likely offer similar performance across states. Verizon does not have MBA data for 24 theoretically measurable states, Charter for 13, Comcast for 12. For other providers the differences are below 6 states. The state coverage of MBA appears appropriate, considering the size of the US. However, deployment of additional probes in states that are not covered by MBA but are present in Form 477 can improve representativeness even further.

Moreover, we group the Form 477 data to ISP, technology, and state combinations (triplets) that are also found in the MBA dataset. Disregarding Optimum, which appears to not be present in Form 477, two of those triplets are represented in the MBA dataset but not in the Form 477 dataset, both consisting of a single probe; these are discarded for further analysis. Using this data, we show the potential footprint that an ISP can have based on the data found in the MBA dataset and compare it to the potential footprint of an ISP in the complete Form 477 data, including all access technologies but limited to states found in the MBA dataset. States in the MBA program cover 326,134,176 people. The potential coverage is shown in Table I. A major difference is the footprints of Satellite ISPs, as they are not measured in every state in the MBA dataset. Further, Verizon, which was not measured in 24 states where connectivity is reportedly possible, has a footprint of 11.3% instead of a theoretical maximum of 17.6%, Charter and Comcast have a negligible decrease in their theoretical footprint. As a result, the triplets that are measured by the MBA data are generally representative when considering footprints of ISPs with a certain access technology in a state, although these footprint estimates do not reflect actual market share.

TABLE I: Metadata comparison for MBA and Form 477, showing number of states covered along with potential population coverage based on ISP, technology, and state triplets.

	MBA		**Form 477**	
	Num. of States	**Triplet Coverage**	**Num. of States**	**Triplet Coverage**
AT&T	20	39.1%	21	41.4%
CenturyLink	33	16.0%	37	16.4%
Charter	28	32.2%	41	33.3%
Cincinnati Bell	3	0.6%	4	1.0%
Comcast	27	34.9%	39	36.1%
Cox	14	6.8%	19	6.9%
Frontier	25	11.6%	31	11.7%
Hughes	34	86.2%	49	100.0%
Mediacom	16	2.0%	22	2.1%
Optimum	4	-	-	-
Verizon	9	11.3%	33	17.6%
Wildblue/ViaSat	19	55.6%	49	100.0%
Windstream	17	2.7%	18	2.8%

A. Metrics

The MBA dataset from April 2012 to December 2019 amounts to roughly 1.7 TB of raw data over 93 months. We aggregate the data for the following metrics:

Latency, Latency under Load, and Packet Loss — The *latency* describes the round-trip time between the probe and a nearby target. The test sends small, regularly spaced UDP packets to the destination, with each UDP packet composing of an 8 byte sequence number and an 8 byte timestamp. Results are aggregated hourly up to the 99th percentile. The hourly statistics cover the minimum, mean, and maximum round-trip time, as well as the standard deviation. *Latency under load* is measured by sending UDP packets to the measurement servers in 500 ms intervals at the same time as the downstream and upstream throughput measurements (cf. below) are conducted, which is implicative of bufferbloat [27]. Both latency and latency under load are aggregated daily by calculating the median of the hourly mean.

The latency test also records the number of successfully delivered and lost packets; a packet is considered lost if the test does not receive a response to the UDP latency measurement within three seconds [28]. We use these records to determine *packet loss* rate daily for each probe. However, note that packet loss should not necessarily be considered bad: Aiming to achieve lower latencies for modern real-time applications may be achieved with smaller buffers to avoid bloat, which leads to more dropped packets at the bottleneck [29]. While minimal levels of packet loss are generally desired, each application is impacted differently by packet loss. As such, there is no threshold for packet loss which defines good or poor performance; some use cases are able to handle up to

1% packet loss without impacting user experience, whereas for others, as much as 2–3% packet loss are considered tolerable.

Downstream Throughput and Download Failures — are measured using three (eight, since 2016) concurrent HTTP `GET` connections to a target. Effects of TCP (such as slow start) are accounted for by using warm-up periods with small data transfers, which are not included in the dataset. Previous work [30] has shown that concurrent TCP connections better approximate the achievable throughput, as single connections are limited by the TCP receive window [31]. Data is aggregated daily and we only consider results with at least one successful measurement and less than half failed ones. Throughput tests are used to compute the *download failure* rate. A download is considered failed if any of the underlying TCP connections cannot be established, times out, is reset, or if the warmup period does not receive any data. The failure rate for each probe across all measurement targets is computed daily.

Note that while the FCC MBA datasets also cover measurements performed over IPv6, we exclusively focus on IPv4 in this paper and leave observations over IPv6 for future work.

III. LONGITUDINAL VIEW

To study the attrition rate, we evaluated how frequently probe owners change ISPs, access technology, or contract between years. While the overall number of probes decreases over the years, majority of probes retain the same ISPs, observed by the high number of probes with no changes or only different speeds (see Fig. 1). Changes in measured download speed are described in § III-B. Overall, the number of ISP changes from one year to the next is rather small; this is likely explained by the market structure of ISPs in the US, where consumers have limited ISP choices (see § IV), as well as customers adjusting their usage behavior to their current broadband speed [32]. Since neither technology nor the ISP change often for probes, the published Validated Data for one specific month in the year (September/October) by the FCC is used to label and analyze the raw data across the whole respective year in the following.

A. Reliability

Packet Loss – Comparing the reliability of different access technologies over the years, we find that for cable access probes, packet loss is low (<1%) across all years for both unloaded and loaded measurements. Under loaded conditions, 99% of the samples exhibit a packet loss under load of at most 1.1% (2018) in every year. 2019 represents an outlier with a 99th percentile of 45.1% packet loss. For comparison, in 2018 three cable probes had an average packet loss above 45%, which increases to 11 in 2019. On the other hand, packet loss rates are lower under unloaded conditions: the 50th (99th) percentile is 0% (<7.1%) packet loss across all years.

We find that packet loss over DSL is generally higher in comparison to cable for loaded conditions. Under load, the highest number of probes with no packet loss occurred in 2013 (3.7%). All other years have less, with 2012, 2016, 2017, 2018, and 2019 having less than 1% of probes with no packet loss. Additionally, the 99th percentiles are at least 36.6% and at most 52.6% packet loss across all years. In terms of unloaded conditions, the percentage of probes with a packet loss rate of 0% is below 50% in 2018 and 2019. The 99th percentile is the highest at 5.1% packet loss in 2016, with 2018 and 2019 at around 4.6%.

Packet loss over fiber is generally slightly higher than over cable. Under load, the 99th percentile under load is highest in 2019 with 14.0%; for unloaded conditions, the highest 99th percentile occurred in 2015 with 3.8%. As such, these packet loss rates are about two times worse relative to cable.

Download Failure Rate – We further compare the download failure rates of the different access technologies over the years. For cable, the download failure rate is at 0% for most of the probes (>94.7%) every year except for 2013 (75.6%).

For DSL, the failure rate is higher than for cable (similar to packet loss). However, we observe improvement over the years: 2013 has the lowest share of probes with 0% failure rate (41.8%), followed by 2012 (67.3%), and 2014 (81.8%). Since 2015, the fraction of probes with on average 0% failure rate stays above 95.4%, which indicates a general improvement of reliability via DSL over the years. However, this slightly decreases from 98.9% (2018) to 96.0% in 2019.

For fiber, we find that more than 98.9% of the probes have no failures for every year, similarly to cable, 2013 is lower with 86.8% of probes not experiencing failures; this also means that the overall median failure rate is consistently zero. Therefore, fiber exhibits even more reliable behavior than cable with respect to downloads.

Takeaway: *Overall, broadband has become more reliable over the years, although cable and fiber appear more reliable than DSL. Cable shows lower packet loss (half as seen over fiber); while fiber has lower failure rates than cable. We see DSL improving in terms of download failure rates since 2012: almost all probes have a median failure rate of 0% after 2015.*

B. Throughput

Using the metadata provided by the FCC, we find that *advertised* downstream as well as upload throughput speeds increase each year: The median advertised upload speed has increased from 2 Mbps in 2012 to 5 Mbps in 2019, while the median download speed has increased from 15 Mbps to 50 Mbps in the same timeframe. To verify whether this advertised throughput is actually reached, we study the *achieved* downstream throughput.

The achieved downstream throughput of probes for each year is shown in Fig. 2. The throughput at the 50th percentile is 17.2 Mbps in 2012, which more than quadruples until 2019 (70.6 Mbps). Similarly, the 75th percentile increases from 23.4 Mbps in 2012 to 127.5 Mbps in 2019. In 2018 only 26.3% of the samples show a download throughput exceeding 100 Mbps, while in 2019 this improves drastically to 35.0%. The 99th percentiles increase from 83.8 Mbps (2012) to 803.6 Mbps (2019). As the maximum throughput observed in 2019 is 928.4 Mbps, and most download speeds do not exceed

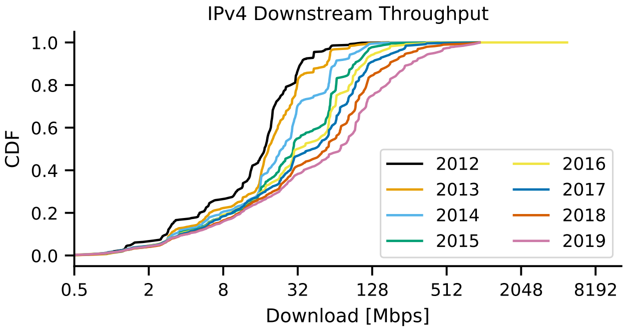

Fig. 2: CDF of the daily downstream throughput by year. The median throughput over IPv4 has improved from 17.2 Mbps in 2012 to 70.6 Mbps in 2019.

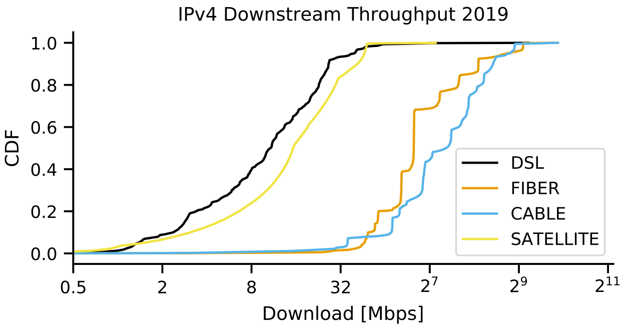

Fig. 3: CDF of the throughput by access technology. Fiber and cable provide the highest throughput speeds, while throughput of DSL and satellite is comparably low.

100 Mbps, 5G [33] deployments that can provide up to 1 Gbps can improve the potential throughput significantly.

Nielsen's Law of Internet Bandwidth [34] predicts a yearly speed increase of 50% for high-end users. To investigate this, we calculate the relative differences between the yearly 95th percentile download speeds to consider high-end customers. However, we do not find consistent evidence supporting Nielsen's Law in the sampled MBA datasets: From 2014–2015 and 2015–2016, the increases are only 21.3% and 21.9%, respectively, whereas the highest increases are from 2013–2014 (58.5%) and from 2018–2019 (60.2%). All other yearly increases are between 34%–38%. Thus, Nielsen's Law does not hold in the average case, although the increases in 2013–2014 and 2018–2019 do surpass the projected growth of 50%.

By Access Technology — Fig. 3 shows that the median throughput is the lowest for DSL with 10.7 Mbps, followed by satellite with 15.3 Mbps. Fiber and cable achieve significantly higher throughputs (medians at 99.9 Mbps and 165.1 Mbps, respectively) in comparison. We find the yearly IPv4 throughput improvement to be mainly driven by probes using cable or fiber as their access technology, especially from 2012 to 2015. We also notice accumulation of samples around specific throughput thresholds, which likely represent different speed tiers offered.

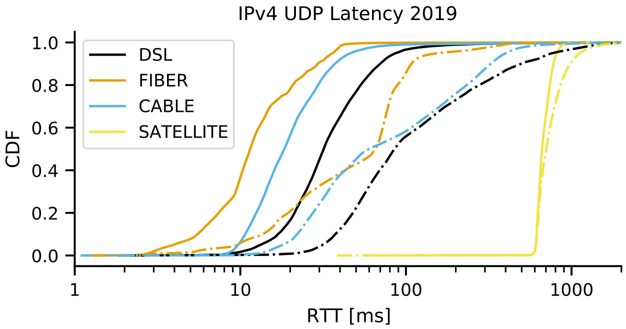

Fig. 4: CDF of the UDP Latency by access technology. The dashed lines represent latency under load. Satellite latency is consistently at around 650 ms.

Takeaway: *The achieved throughput increases significantly over the years, with cable and fiber reaching the highest downloads speeds in 2019. Throughput over DSL is much lower in comparison to other technologies, which indicates that upcoming 5G deployments can improve throughput in regions where broadband is primarily or exclusively offered over DSL.*

C. Latency / Latency under Load

Latency under load has improved over the years, with 2019 having the lowest median latency under load (66 ms) improved from 107.6 ms in 2012, i.e., a reduction of 38.7%. The high latency under load in 2012 is likely caused by bufferbloat, for which improved AQM implementations (see RFCs 8033, 8289, 8290) have been proposed, standardized, and deployed in the following years. The unloaded IPv4 latency, on the other hand, exhibits a more stable pattern with slight improvements from year to year, in which the medians range from 22.5 ms to 26.8 ms, with the lowest latency observed in 2019.

By Access Technology — Access links and technologies are known to have large impacts on latency [35], [36]. To study this, Fig. 4 shows the unloaded latency (solid lines) along with the latency under load (dashed lines) for the different access technologies in 2019. We see that fiber and cable show the lowest unloaded latency, with medians of 11.5 ms and 18.4 ms in 2019. The median latency for DSL is higher with 31.9 ms, likely due to links applying (more) parity bits to the transmitted data to enable Forward Error Correction (FEC) [30] or DSL interleaving on the last mile [37]. However, under loaded conditions, our observations are different: The median latency under load over fiber is nearly six times higher (65.5 ms) when compared to the unloaded median. The latency over other access technologies deteriorate less under load in comparison, with the medians increasing by a factor of 3.2 for cable (to 59.7 ms) and 2.7 for DSL (to 85.3 ms).

Takeaway: *Overall, we observe that the daily latencies decrease over the years, with improvements most visible for latency under load, which suggests that the impact of bufferbloat on latency has decreased over time. On the other hand, unloaded latency has improved slightly.*

IV. COVERAGE MAPS

Having observed the longitudinal growth of broadband in the USA, we now build coverage maps to compare different US states in terms of their broadband performance. Each point on a coverage map represents the median (metric) for a specific probe. The measurements allow determining potential areas that experience poor broadband performance (creating a digital divide), where the extension of fixed-line broadband services or deployment of LEO satellites [15], [16], [17], [18] and 5G [33] can benefit consumers in terms of availability. Further, in conjunction with the Form 477 data, we highlight areas that are under-sampled in the the FCC MBA program, where additional probes can be deployed to improve representativeness.

A. Reliability

Packet Loss — Fig. 5 shows the medians of daily packet loss rates under load for each probe by location in 2019, distinguishing between loaded (top) and unloaded conditions (bottom) over IPv4. For packet loss under load, we observe the Northeast to have the highest amount of probes (78.6%) with low (<1%) packet loss, followed by the rest of the East Coast (overall 72.0% with low packet loss). The West Coast on the other hand has only 53.0% of the probes with low packet loss. The southern central states fare better (61.7%) than the northern ones (48.2%). Out of all access technologies, DSL has the highest median packet loss between 2.5% (Midwest) and 4.3% (Northeast). We also observe the Northeast to have higher packet loss under load over DSL in general compared to other regions. On the other hand looking at all non-DSL access technologies combined, the Northeast has lower packet loss compared to other regions.

For packet loss under unloaded conditions, packet loss is generally low (<1%) across the country. However the Northeast fares the best out of all regions when it comes to the relative amount of probes with low packet loss. The West Coast is only slightly worse than the East Coast, however, the overall differences are minimal and all states have over 90% of their probes exhibit low packet loss.

Download Failure Rate — Fig. 6 presents the median of daily failure rates for all probes in 2019 over IPv4. While the failure rate is generally low for the majority of probes, a few probes exhibit higher failure rates, primarily in the Central U.S. Most of the probes with medium to high failure rates (between 20% and 35%) are served by Windstream, which offers broadband access over DSL. Additionally, probes served by Mediacom in the Midwest show an unexpected amount of high failure rates.

Takeaway: *Overall, download failure rates are low across probes, with some exceptions in the Central regions. Probes connected via DSL exhibit the highest failure rates and show low reliability (packet loss): While packet loss under unloaded conditions is low (<1%) across the country, packet loss under load is the lowest on the East Coast.*

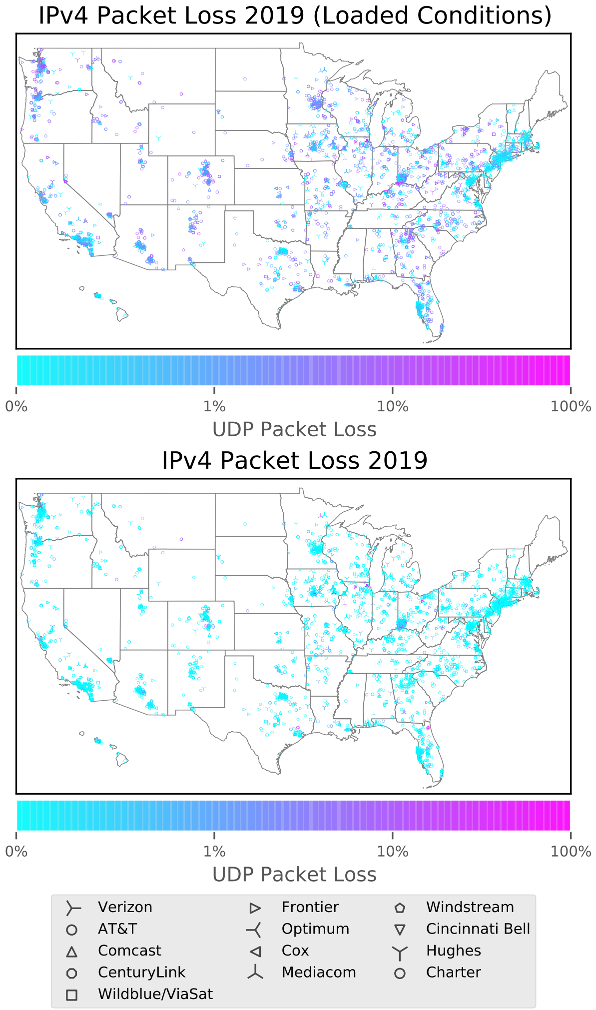

Fig. 5: Packet loss by location. Under load, probes at the coasts are affected least by higher packet loss.

B. Throughput

For coverage maps of download speeds, Fig. 7 shows median downstream throughput measured in 2019 for each probe over IPv4. We see that the probe locations and higher throughput speeds roughly follow the US population density, with higher downstream throughputs visible at the Coasts, especially the East Coast. Lower download speeds are most common in the Great Lakes region, the Mid-Atlantic, the Southeast and Hawaii. Fig. 3 shows that a number of probes with lower download speeds below the FCC speed benchmark of 25 Mbps [5] are connected via DSL. This is also visible on the coverage map, where DSL (AT&T, CenturyLink, Cincinnati Bell, Frontier, Hawaiian Telcom, Verizon, or Windstream) has lower throughput across the country. We further see that satellite ISPs provide lower download speeds, although they surpass throughput over DSL for most percentiles (see Fig. 3).

While there are other ISPs offering higher download speeds in these areas with lower DSL speeds, the number of options might be limited for the consumers: Fig. 9 shows the download

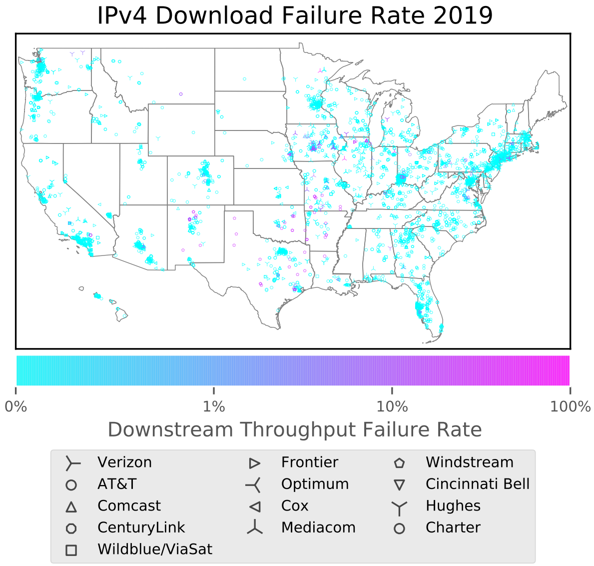

Fig. 6: Failure rates of downloads by location. The overall download failure rate is low (<1%), although some probes in the Midwest and South exhibit more failures.

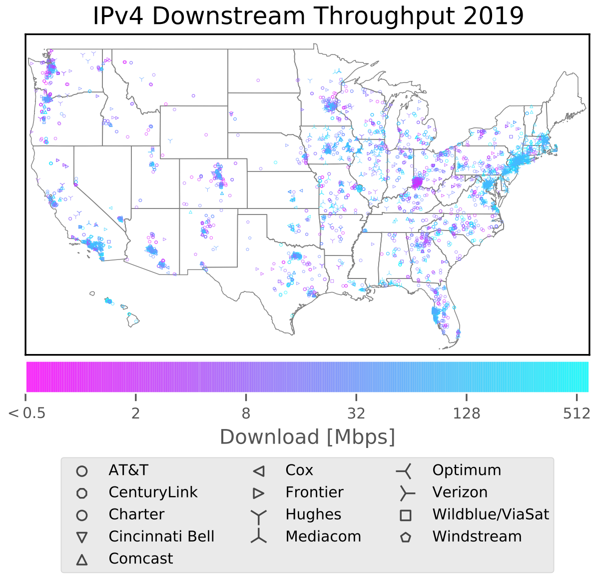

Fig. 7: Downstream throughput by location. Many probes in the East North Central, and Middle and South Atlantic divisions achieve lower throughputs, especially over DSL.

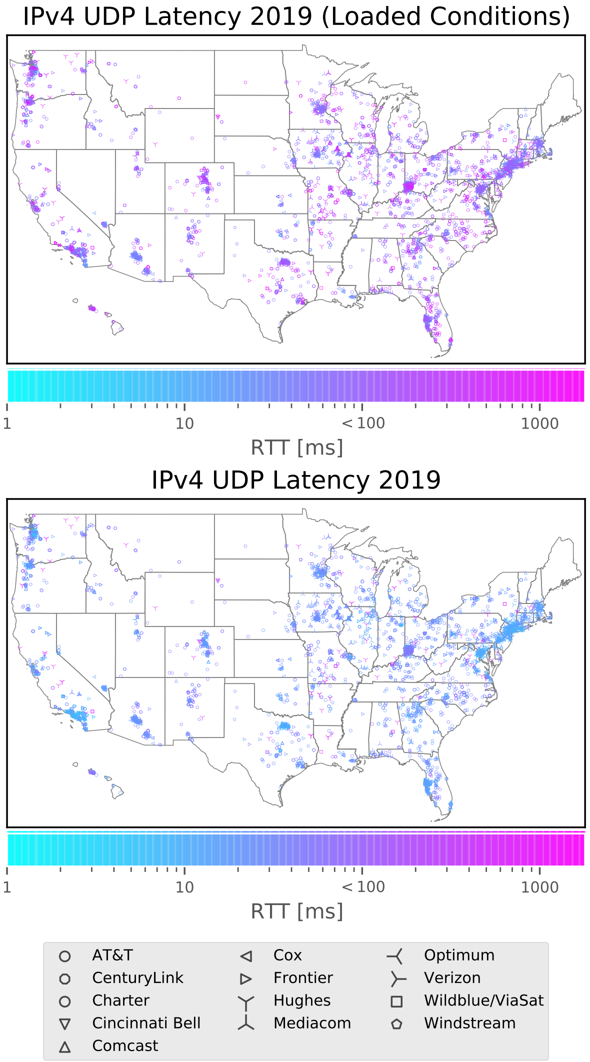

Fig. 8: UDP latency by location. The unloaded latency for most probes is at <250 ms, but increases substantially under load. Higher latencies can be seen in the Midwest.

speeds over IPv4 for probes by state and ISP (ISPs with the same access technology are grouped together). Darker gray cells indicate combinations of ISP/access technology and state that are offered (according to Form 477) but where probes are missing in the MBA dataset. For these cells, probes should be deployed for more representative measurements in the future, while for lighter gray cells, the FCC and ISPs may consider expanding to/in these regions.

Considering access technologies, for cable and fiber ISPs, blue strands are visible in Fig. 9 (which represent higher download speeds); this supports the observation that cable and fiber generally provide the highest download speeds. Further, DSL is seen to provide download speeds below the FCC benchmark [5] of 25 Mbps in nearly all states, with only AR, IA, MN, NV and TX achieving higher median downstream throughput over DSL; most of those are served by Windstream.

Users in states such as CA, CT, FL, GA, LA, OH, or VA have a variety of providers to choose from, even at higher speed tiers above 100 Mbps with a mix of fiber and cable connectivity. However, in other states such as CO, ID, MT, ND, SD, or WY users have few or no providers with download speeds above 100 Mbps, which indicates a possible lack of competition in the broadband market. Here, deploying LEO satellites [15], [16], [17], [18] can particularly benefit customers by contributing additional options in terms service availability.

Takeaway: *Throughput is the lowest for probes located in the Rocky Mountains and Plains regions. Areas where DSL is the*

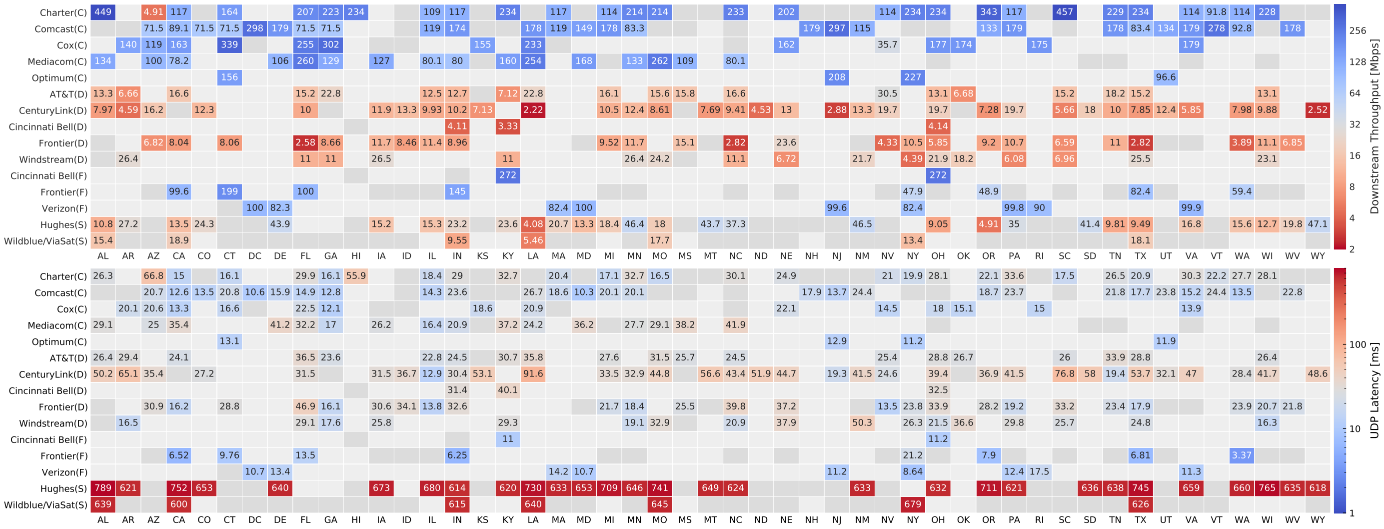

Fig. 9: IPv4 Downstream Throughput (top) and unloaded latency (bottom) by ISP, access technology, and state for 2019. Darker gray cells indicate triplets (§ II) that exist in Form 477 (ISP offers services in that state), but are not measured by MBA.

primary access technology suffer from slower speeds, as DSL achieves much lower throughput compared to cable and fiber. In those areas, probes struggle to achieve the 25/3 Mbps speed benchmark that the FCC uses to evaluate broadband, indicating that the digital divide is not yet closed in many regions.

C. Latency

Regarding latency measurements, Fig. 8 shows median latency under load (top) and under unloaded conditions (bottom) over IPv4 by probe as of 2019. Under loaded conditions, probes experience higher latencies across the country. When grouped by region, probes connected via fiber achieve the lowest latencies (median of 22.9 ms for the Midwest), primarily provided by Cincinnati Bell. This is followed by cable probes in the West (45.2 ms) and Northeast (61.2 ms). However in the South, Cable probes have inflated latencies under load (104 ms), higher on average than DSL connections (100.1 ms), which fare worse than other access technologies in all other regions. However DSL probes also exhibit high median latencies (up to 1026 ms for certain states).

In unloaded scenarios, the latency is generally low in all regions, which indicates that no regions exhibit excessively high latencies. This is also reflected in Fig. 9 (bottom), showing the median daily latency by ISP, access technology, and state. Most ISPs achieve latencies below 20–30 ms, although some DSL providers consistently exhibit latency values above 30 ms across all states, occasionally up to 76.8 ms. Satellite ISPs have the highest latencies, in the range of 600–789 ms, as expected due to their high propagation delays.

Takeaway: *In all regions, latency in unloaded scenarios is low. However, under load, probes across the whole country measure higher latency especially in the South.*

V. CONCLUSION

We provided an empirical grounding of broadband infrastructure in the US, using a dataset with eight years of data collected by the FCC Measuring Broadband America (MBA) program. Focusing on the longitudinal evolution and coverage maps to evaluate the digital divide, we studied several performance indicators. In the process, we also demonstrated the representativeness of the dataset by comparing it with the Form 477 dataset and identified key areas where the FCC can target deployment of more probes for increased representation.

The data shows that packet loss appears to be higher over DSL and satellite links, but download failure rates have improved substantially over the years. The median advertised download speed is roughly more than three times higher in 2019 than in 2012 (from 15 to 50 Mbps). The highest download throughputs are achieved at the coasts, with lower download speeds primarily being seen over DSL, and in the Midwest and Northeast. Moreover, we find that only few ISPs have throughput measurements that exceed 100 Mbps in some states, which indicates that they might be underserved; Low Earth Orbit (LEO) satellite broadband deployments could improve coverage in these regions. Further, we see that download speeds are at most 127.5 Mbps for 75% of the probes, which means that a speed of up to 1 Gbps by upcoming 5G deployments can improve the performance of US broadband drastically in terms of throughput. We observe that latency has improved over the years, especially under loaded conditions. In unloaded scenarios, the median latency has reduced by roughly 4 ms (from 22.5 ms to 26.8 ms) or 16%. The median latency under load has reduced by 38.7% from 107.6 ms (2012) down to 66 ms (2019). Nevertheless, the impact of bufferbloat has decreased over the years. Fiber links show the best results, despite showing less reliability when load is added to the network. While the state of broadband in the US has improved over the years, several areas, especially in the Midwest and Northeast but also in some Southern states, lack access to reliable, low-latency, high-speed broadband. Overall, the observed trends indicate that the divide is closing, although the pace is slow in some states.

While geostationary satellites do help with bridging the divide in these underserved areas, they only go as far as providing basic connectivity. However, even with geostationary satellites, low latency broadband is still missing in many of these regions, which upcoming LEO satellite deployments can target in the upcoming years to continue closing the digital divide.

VI. ACKNOWLEDGEMENTS

This work was partly supported by the Volkswagenstiftung Niedersächsisches Vorab (Funding No. ZN3695).

REFERENCES

[1] T. Simonite, “The Tech Antitrust Problem No One Is Talking About,” 2020, [Online; accessed 2022-04-08]. [Online]. Available: https://www.wired.com/story/tech-antitrust-problem-no-one-talking/

[2] E. Falcon, “Broadband Monopolies Are Acting Like Old Phone Monopolies. Good Thing Solutions to That Problem Already Exist,” 2019, [Online; accessed 2022-04-08]. [Online]. Available: https://www.eff.org/deeplinks/2019/05/broadband-monopolies-are-acting-old-phone-monopolies-good-thing-solutions-problem

[3] Federal Communications Commission, “Internet Access Services: Status as of December 31, 2018,” Sep 2020, [Online; accessed 2022-04-08]. [Online]. Available: https://docs.fcc.gov/public/attachments/DOC-366980A1.pdf

[4] ——, “FCC Adopts First Consolidated Communications Marketplace Report,” 2018, [Online; accessed 2022-04-08]. [Online]. Available: https://www.fcc.gov/document/fcc-adopts-first-consolidated-communications-marketplace-report-0

[5] ——, “2020 Broadband Deloyment Report: New FCC Report Shows Digital Divide Continuing to Close,” 04 2020, [Online; accessed 2022-04-08]. [Online]. Available: https://www.fcc.gov/document/new-fcc-report-shows-digital-divide-continuing-close-0

[6] ——, “Fixed Broadband Deployment Data from FCC Form 477,” 2020, [Online; accessed 2022-04-08]. [Online]. Available: https://www.fcc.gov/general/broadband-deployment-data-fcc-form-477

[7] ——, “Form 477 Census Tract Data on Internet Access Services,” 2019, [Online; accessed 2022-04-08]. [Online]. Available: https://www.fcc.gov/general/form-477-census-tract-data-internet-access-services

[8] ——, “Measuring Broadband America,” Jan 2021, [Online; accessed 2022-04-08]. [Online]. Available: https://www.fcc.gov/general/measuring-broadband-america

[9] V. Bajpai and J. Schönwälder, “A Survey on Internet Performance Measurement Platforms and Related Standardization Efforts,” *IEEE Communications Surveys and Tutorials*, 2015. [Online]. Available: https://doi.org/10.1109/COMST.2015.2418435

[10] Measuring Broadband America, “Sample Size and Volunteer Selection,” 2011, [Online; accessed 2022-04-08]. [Online]. Available: https://transition.fcc.gov/cgb/measuringbroadbandreport/technical_appendix/4SampleSizeandVolunteerSelection.pdf

[11] S. Sundaresan, N. Feamster, R. Teixeira, and N. Magharei, “Measuring and Mitigating Web Performance Bottlenecks in Broadband Access Networks,” in *IMC*, 2013. [Online]. Available: https://doi.org/10.1145/2504730.2504741

[12] S. Sundaresan, N. Feamster, and R. Teixeira, “Home Network or Access Link? Locating Last-Mile Downstream Throughput Bottlenecks,” in *PAM*, 2016. [Online]. Available: https://doi.org/10.1007/978-3-319-30505-9_9

[13] Z. S. Bischof, F. E. Bustamante, and R. Stanojevic, “The Utility Argument - Making a Case for Broadband SLAs,” in *PAM*, 2017. [Online]. Available: https://doi.org/10.1007/978-3-319-54328-4_12

[14] C. Smith-Salzberg, F. Fund, and S. S. Panwar, “Bridging the digital divide between research and home networks,” in *IEEE INFOCOM Workshops*, 2017, pp. 444–449. [Online]. Available: https://doi.org/10.1109/INFCOMW.2017.8116417

[15] M. Handley, “Delay is Not an Option: Low Latency Routing in Space,” in *HotNets 2018*. ACM, 2018. [Online]. Available: https://doi.org/10.1145/3286062.3286075

[16] ——, “Using ground relays for low-latency wide-area routing in megaconstellations,” in *HotNets 2019*. ACM, 2019. [Online]. Available: https://doi.org/10.1145/3365609.3365859

[17] D. Bhattacherjee and A. Singla, “Network topology design at 27, 000 km/hour,” in *CoNEXT*. ACM, 2019, pp. 341–354. [Online]. Available: https://doi.org/10.1145/3359989.3365407

[18] G. Giuliari, T. Klenze, M. Legner, D. A. Basin, A. Perrig, and A. Singla, “Internet Backbones in Space,” *CCR*, vol. 50, no. 1, pp. 25–37, 2020. [Online]. Available: https://doi.org/10.1145/3390251.3390256

[19] S. Kassing, D. Bhattacherjee, A. B. Águas, J. E. Saethre, and A. Singla, “Exploring the ”Internet from space” with Hypatia,” in *IMC*. ACM, 2020. [Online]. Available: https://doi.org/10.1145/3419394.3423635

[20] V. Bajpai, A. Brunström, A. Feldmann, W. Kellerer, A. Pras, H. Schulzrinne, G. Smaragdakis, M. Wählisch, and K. Wehrle, “The Dagstuhl Beginners Guide to Reproducibility for Experimental Networking Research,” *CCR*, vol. 49, pp. 24–30, 2019. [Online]. Available: https://doi.org/10.1145/3314212.3314217

[21] S. Bauer, D. D. Clark, and W. H. Lehr, “Gigabit Broadband Measurement Workshop Report,” *CCR*, 2020. [Online]. Available: https://doi.org/10.1145/3390251.3390259

[22] Federal Communications Commission, “Measuring Broadband America - Open Methodology,” 2020, [Online; accessed 2022-04-08]. [Online]. Available: https://www.fcc.gov/general/measuring-broadband-america-open-methodology

[23] C. Dovrolis, K. Gummadi, A. Kuzmanovic, and S. Meinrath, “Measurement Lab: Overview and an Invitation to the Research Community,” *CCR*, 2010. [Online]. Available: https://doi.org/10.1145/1823844.1823853

[24] Federal Communications Commission, “Measuring Broadband Raw Data Releases - Fixed,” 2020, [Online; accessed 2022-04-08]. [Online]. Available: https://www.fcc.gov/oet/mba/raw-data-releases

[25] ——, “Measuring Broadband America Program - Fixed Reports Data and Related Materials,” Jan 2021, [Online; accessed 2022-04-08]. [Online]. Available: https://www.fcc.gov/reports-research/reports/measuring-broadband-america/measuring-broadband-america-program-fixed

[26] A. Dunna, Z. S. Bischof, and R. Fontugne, “Sanitizing a view of consumer broadband in the united states,” in *TMA*. IFIP, 2020. [Online]. Available: http://dl.ifip.org/db/conf/tma/tma2020/tma2020-camera-paper79.pdf

[27] J. Ye, K. Leung, V. O. K. Li, and S. H. Low, “Combating Bufferbloat in Multi-Bottleneck Networks: Equilibrium, Stability, and Algorithms,” in *INFOCOM*. IEEE, 2018, pp. 648–656. [Online]. Available: https://doi.org/10.1109/INFOCOM.2018.8486251

[28] R. Padmanabhan, P. Owen, A. Schulman, and N. Spring, “Timeouts: Beware Surprisingly High Delay,” in *IMC*, 2015. [Online]. Available: https://doi.org/10.1145/2815675.2815704

[29] N. Feamster and J. Livingood, “Measuring internet speed: current challenges and future recommendations,” *Commun. ACM*, vol. 63, no. 12, pp. 72–80, 2020. [Online]. Available: https://doi.org/10.1145/3372135

[30] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè, “Measuring home broadband performance,” *Commun. ACM*, vol. 55, no. 11, pp. 100–109, 2012. [Online]. Available: https://doi.org/10.1145/2366316.2366337

[31] S. Bauer, D. Clark, and W. Lehr, “Understanding Broadband Speed Measurements,” in *TPRC*, 2010. [Online]. Available: https://ssrn.com/abstract=1988332

[32] Z. S. Bischof, F. E. Bustamante, and R. Stanojevic, “Need, Want, Can Afford: Broadband Markets and the Behavior of Users,” in *IMC*, 2014, pp. 73–86. [Online]. Available: https://doi.org/10.1145/2663716.2663753

[33] A. Narayanan, E. Ramadan, J. Carpenter, Q. Liu, Y. Liu, F. Qian, and Z.-L. Zhang, “A First Look at Commercial 5G Performance on Smartphones,” in *The Web Conference*, New York, NY, USA, 2020, p. 894–905. [Online]. Available: https://doi.org/10.1145/3366423.3380169

[34] Nielsen Norman Group, “Nielsen’s Law of Internet Bandwidth,” 1998, [Online; accessed 2022-04-08]. [Online]. Available: https://www.nngroup.com/articles/law-of-bandwidth/

[35] J. Rula, Z. Bischof, and F. Bustamante, “Second Chance: Understanding diversity in broadband access network performance,” in *ACM SIGCOMM Workshop on Crowdsourcing and Crowdsharing of Big (Internet) Data*, 2015. [Online]. Available: https://doi.org/10.1145/2787394.2787400

[36] T. Høiland-Jørgensen, B. Ahlgren, P. Hurtig, and A. Brunstrom, “Measuring Latency Variation in the Internet,” in *CoNEXT*. ACM, 2016. [Online]. Available: https://doi.org/10.1145/2999572.2999603

[37] V. Bajpai, S. J. Eravuchira, and J. Schönwälder, “Dissecting last-mile latency characteristics,” *Computer Communication Review*, vol. 47, no. 5, 2017. [Online]. Available: https://doi.org/10.1145/3155055.3155059