Coverage is Not Binary: Quantifying Mobile Broadband Quality in Urban, Rural, and Tribal Contexts

Vivek Adarsh(✉)[1], Michael Nekrasov[1], Udit Paul[1], Tarun Mangla[2],
Arpit Gupta[1], Morgan Vigil-Hayes[3], Ellen Zegura[2], Elizabeth Belding[1]
[1]University of California, Santa Barbara
{vivek, mnekrasov, u_paul, arpitgupta, ebelding}@cs.ucsb.edu
[2]Georgia Tech, [3]Northern Arizona University
{tmangla3, ewz}@cc.gatech.edu, morgan.vigil-hayes@nau.edu

***Abstract*—Cellular network performance does not cleanly generalize. A variety of factors, such as location, terrain, signal quality and network load, affect the performance of services delivered over LTE networks. As a result, the presence of LTE coverage does not always equate to usable service; coverage can be of poor quality, or it can be congested and difficult to access. Given that reliance on LTE networks for Internet connectivity has exploded, it is critical to understand the quality of experience for applications delivered over these networks in a variety of scenarios. To this end, we develop a robust measurement suite that we use to conduct a unique measurement campaign in tribal, rural, congested urban and uncongested urban regions, representing a variety of under-provisioned, congested, and well-provisioned operational LTE networks run by four major providers. Our analysis confirms that the performance of LTE networks in tribal and rural areas is typically worse than even heavily congested urban networks. More specifically, in the regions that we study, LTE networks in under-provisioned (tribal/rural) areas have 9× poorer video streaming quality, 10× higher video start-up delay, undergo more than 10× the number of resolution switches, and lead to more than 2× slower Web browsing experience as compared to urban deployments. We show that throughput and latency are 11× and 3× worse in tribal and rural locations, despite identical LTE carrier subscription plans.**

***Keywords*—LTE, QoE, QoS, Rural, Congestion, Video streaming, Web browsing**

I. Introduction

LTE plays an increasingly critical role in providing pervasive Internet access. A 2019 Pew Research study reports that roughly one in five U.S. adults is "smartphone dependent," meaning they solely rely on mobile broadband for Internet access at home [1]. Individuals living in rural and tribal regions are particularly likely to rely on mobile broadband for Internet access [2]. As growing numbers of people depend on LTE networks as their primary means for accessing healthcare, financial, and educational services, it has become *critical to evaluate how well these networks service user applications*. Due to the COVID-19 pandemic, the urgency of evaluating the quality of experience for applications delivered over mobile broadband has skyrocketed as stay-at-home orders and rapid movement to online schooling increase the demand for applications that are known to be sensitive to network quality, such as video streaming and interactive video chat [3]. As a result, communities without access to usable, high speed broadband, such as many rural and tribal regions, are severely disadvantaged [4].

There is a need for targeted measurement campaigns that represent performance within challenged networks [5]; the Federal Communications Commission recently encouraged researchers to undertake campaigns to study and report the state of rural networks [6]. Our goal is to understand mobile quality of service (QoS) and quality of experience (QoE) performance profiles for common, and increasingly essential, applications such as video streaming and Web browsing, in a variety of network conditions. To do so, we undertook an extensive measurement campaign to collect 16 datasets of network traces in the Southwestern U.S. for four major telecom operators: AT&T, Sprint, T-Mobile and Verizon, gathering over 30 million LTE packets. To understand geographic performance discrepancies, we collected measurements of LTE networks in tribal, rural, and urban communities. While we anticipate that network performance in tribal and rural areas will differ from that in urban areas because rural and tribal networks are often under-provisioned [5], [7], the objective of our study is to quantify the severity of performance degradation in under-provisioned networks. Service quality is not a binary label, just like cellular coverage; for instance, application performance is subject to network conditions. Our goal to quantify network performance stems from the need to accurately indicate the behavior of different applications, and not simply label a region as "covered" or "not covered".

Our tribal and rural measurements were conducted in New Mexico. New Mexico is one of the least densely populated states in the U.S. and 10% of its land area belongs to one of the 23 sovereign tribes with territories in the state [8]. In the rural regions, there is a high concentration of smartphone dependent residents [9].

In addition to the tribal and rural contexts, we collect network traces from crowded events in urban locations in California, during which atypically high volumes of network utilization cause congestion. For comparison, we also collect traces from the same urban locations during typical operating conditions as a baseline. Our datasets have broad spatial and temporal variability, but can be classified into three categories: under-provisioned (rural and tribal), congested (congested urban), and well-provisioned (baseline urban).

While Web browsing is a critical component of daily Internet access, streaming video currently accounts for 65% of all downstream mobile traffic worldwide [10] (for instance, in the U.S., more than 80% of the population possess some form of video streaming subscriptions [11]). Therefore, we focus our analysis on understanding the QoE of Web browsing and streaming video for these regions. At each location, we collect extensive QoS and QoE measurements. Based on our analysis, we illustrate critical performance differences between the three location categories. Our key contributions and findings include:

- Collection of 16 network performance datasets from 12 locations across the Southwestern U.S., representative of three network conditions: under-provisioned (rural and tribal), congested urban and well-provisioned urban;
- Characterization of LTE traffic across all locations and network conditions in the datasets, through analysis of four QoS and six QoE metrics;
- Analysis of QoS and QoE data, which reveals that rural and tribal LTE networks consistently perform worse than the studied urban baseline deployments, and typically comparable or worse than congested urban networks.

II. Evaluation Metrics

Typically, a binary representation of cellular coverage (i.e., is an area covered or not) is used to characterize the state of Internet connectivity over LTE networks. However, from our own experiences, as well as that of others [12], such a simplistic characterization of networked services over LTE networks is insufficient. This becomes increasingly true in rural regions, as base station coverage areas are greater, and weaker signals are more commonly experienced. Even in well-covered urban areas, performance can suffer during times of atypically high usage, i.e. flash crowds due to a heavily attended community event [13], [14]. As application requirements place more load on the network, it becomes critical to determine, not only whether coverage exists in a region, but whether that coverage is of high enough quality to support the types of applications wanted, or needed, by the local users. As we have seen with recent shelter-in-place orders due to COVID-19, residents of regions with sub-standard Internet access are at risk of being left behind, educationally, economically and medically [4], [15]. To evaluate network quality, we turn to QoS and QoE metrics and use these metrics to analyze the ability of the networks to support the most accessed applications: Web browsing and video streaming traffic, which are applications of high, and still growing, usage. In this section, we describe the QoS and QoE metrics collected for this measurement study, which are summarized in Table I.

A. Quality of Service Metrics

Different applications have different network requirements, and QoS metrics capture the state of network performance. For instance, delay-sensitive Internet traffic, such as live streaming video and multimedia teleconferencing, requires low end-to-end delay to maintain interactivity; an application such as on-demand gaming is dependent on both end-to-end delay and achieved throughput. Barriers to achieving usable QoS in LTE networks include poor coverage quality and high network utilization. We measure four metrics to determine QoS.

Reference Signal Received Power (RSRP): RSRP is defined as the linear average over the power contributions (in Watts) of the resource elements that carry cell-specific reference signals within the measurement frequency bandwidth [16]. Although there are many key performance indicators (KPIs) related to received signal strength, we focus specifically on RSRP, as defined by 3GPP [17]. [18] demonstrates that RSRP has a significant impact on the mean opinion score (MOS) of video streaming; MOS varies significantly at RSRP values between -84dBm and -102dBm and declines rapidly below -104dBm. RSRP is used for a variety of LTE operations (e.g., cell selection, handover decisions [19], network quality assessment, etc.) and, as illustrated by [20], is widely accessible through mobile operating systems. Typically, RSRP is reported in dBm by the user equipment (UE) as the average power over several narrow-band control channels. We record instantaneous RSRP readings from the user equipment every one second through the Network Monitor application [21].

Throughput: Our network monitoring suite automates the collection of throughput measurements by fetching a pre-specified 500 MB file from an AWS instance, hosted in Virginia. For uniformity, we use the same instance for all measurement tests. To calculate the throughput, the client initiates iPerf threads over TCP to download the file. The large file size allows the data traffic to fill the pipe and to minimize the effect of slow start. We log the packet traces at the client during the

TABLE I: Overview of QoS and QoE metrics at each location, aggregated across available providers.

Type	Metric	Test Interval	# of Datapoints	Tools
QoS	RSRP	1 second	2160	Network Monitor
	Throughput	1 second	2160	iPerf
	Latency	1 second	960	HPing3
	Packet loss	N/A	N/A	tshark
QoE	Start-up delay	1 second	2160	Selenium, iframe API
	Video quality	1 second	2160	Selenium, iframe API
	Resolution switches	1 second	2160	Selenium, iframe API
	Rebuffering percentage	1 second	2160	Selenium, iframe API
	Buffer size	1 second	2160	Selenium, iframe API
	Page load time	N/A	300	Selenium, Chromium

iPerf tests to sample throughput over one second intervals for each location.

Latency: We measure round-trip times through pings, initiated by Hping3 [22], to the same Virginia-based AWS server. We configure Hping3 to use TCP packets instead of ICMP because ICMP packets were occasionally dropped at the server. The latency test runs for 120 seconds with one-second intervals between each ping. We measure latency twice during each measurement session: once before all the video stream and throughput tests (described above) and once immediately after. Hence we collect 240 latency data-points per operator, for 960 total at every location. Low round-trip times tend to be indicative of a better user experience for delay-sensitive applications.

Packet Loss: Packet loss in cellular networks can occur due to network congestion and/or transmission errors [23]. We analyze the synchronous packet traces from both the client and the server during throughput tests to compute packet loss using `tshark CLI`.

B. Quality of Experience Metrics

To measure QoE, we focus on streaming video and Web browsing, currently the most heavily used QoE-centric applications in mobile networks [24]. Internet video streaming services typically use Dynamic Adaptive Streaming over HTTP (DASH) [25] to deliver a video stream. DASH divides each video into time intervals known as segments or chunks, which are then encoded at multiple bit-rates and resolutions. We measure a variety of metrics associated with video streaming quality, as described below.

To assess the quality of Web browsing, we measure the page load times of some of the most frequently accessed Web pages. Numerous studies and media articles report its importance for user experience [26], [27], and consequently to business revenue. The QoE metrics we measure are summarized in Table I and are described below. Our approach for measuring the majority of these metrics is described in section III.

Start-up Delay: Start-up delay is the time elapsed from the moment the player initiates a connection to a video server to the time the application starts rendering video frames. This delay usually corresponds to how quickly the HTTP adaptive streaming client is able to fill the threshold buffer required for playback.

Video Quality: Video quality is the number of pixels in each dimension of video frame [28]. Video quality, or resolution, is an important component of QoE; a higher resolution results in a better visual experience, up to a point.

Resolution Switching: Frequent changes in video resolution can result in user frustration, particularly when the video quality is downgraded [24]. We compute the number of samples that had a different resolution from the prior sample in our video streaming sessions as a percentage of total number of samples collected during the video session. Since resolution switches occur in-between video *chunks* that are typically 4–5 seconds long [28], our analysis at one-second granularity is a lower bound estimate, if not better. Both the magnitude (difference in pre- and post-switch resolution) and the frequency of video resolution switches affect the quality of experience [24].

Rebuffering Percentage: A rebuffering event occurs when the application buffer waits to accumulate enough content to resume playback. Poor link quality and/or congestion can lead to an increase in video rebuffering events because they cause delays in packet delivery [29]. When rebuffering occurs, the user notices interrupted video playback, commonly referred to as *stalling*. Rebuffering events have a key influence on user satisfaction and significantly impact video abandonment [30]. We represent the rebuffering percentage as the amount of time the video stalls during the playback expressed as a percentage of total playback time.

Buffer Size: The streaming client employs a playout buffer or client buffer, whose maximum value is the buffer capacity (in seconds) to temporarily store chunks to absorb network delay variations (i.e., jitter). To ensure smooth playback and adequate buffer level, the client requests a video clip chunk by chunk (in seconds) using HTTP GET requests, and dynamically determines the resolution of the next chunk based on network conditions and buffer status. When the buffer level is below a minimum threshold, the client requests chunks as fast as the network can deliver them to increase the buffer level. The playback stalls when the buffer is empty before the end of the playback is reached.

Page Load Time: To compute load times, the fetching of Web pages is automated using Selenium [31]. We use the Tranco Top 25 list [32]. For evaluation, we log the navigation timings of a Web page starting from $navigationStart$ through the $loadEventEnd$ event [33]. These instances of event timings help in a fine-grained analysis of page load times. We download each webpage three times and average the results. The browser cache is automatically wiped out after each Web page load to reflect true load time for the next iteration.

III. Methodology and Datasets

It is our goal for our measurements to represent a range of network deployments that vary both by signal quality and offered load. We focus on the networks of the four major U.S. providers: AT&T, Sprint, T-Mobile and Verizon. In this section, we first describe our custom measurement suite and our measurement methodology. We then describe the details of our collected datasets.

A. Measurement Suite

We collect measurements from which we can derive QoS and QoE in a variety of locations in which we expect varying LTE performance. Capture of disparate performance will give us both a broad picture of QoS and QoE, and the opportunity to study performance in different population densities (urban vs. rural vs. tribal reservations) and usage scenarios (congested vs. typical usage). To collect these measurements, we use a custom-built measurement suite that captures both network-

TABLE II: Summary of Datasets

Location	Date (2019)	# LTE Packets	Type	Carriers*
New Mexico				
Tribal_1	May 28	3.18 Million	Tribal, Rural	V,A,T,S
Tribal_2	May 29	1.38 Million	Tribal, Rural	V,T
Tribal_3	May 28	2.03 Million	Tribal, Rural	V,A,T,S
Tribal_4	May 30	2.16 Million	Tribal, Rural	V,A,T,S
Tribal_5	May 30	2.27 Million	Tribal, Rural	V,A,T,S
Tribal_6	May 31	2.33 Million	Tribal, Rural	V,A,T,S
Rural_1	May 31	1.26 Million	Non-Tribal, Rural	V,T
Rural_2	Jun 01	2.83 Million	Non-Tribal, Rural	V,A,T,S
San Diego, CA				
Urban_1_Cong	Sep 22	2.25 Million	Urban, Congested	V,A,T,S
Urban_1_Base	Sep 28	1.92 Million	Urban, Baseline	V,A,T,S
Urban_2_Cong	Sep 29	2.51 Million	Urban, Congested	V,A,T,S
Urban_2_Base	Sep 30	1.97 Million	Urban, Baseline	V,A,T,S
Urban_3_Cong	Sep 21	2.65 Million	Urban, Congested	V,A,T,S
Urban_3_Base	Sep 30	2.13 Million	Urban, Baseline	V,A,T,S
San Francisco, CA				
Urban_4_Cong	Sep 25	2.18 Million	Urban, Congested	V,A,T,S
Urban_4_Base	Sep 26	2.08 Million	Urban, Baseline	V,A,T,S

*This column lists the mobile carriers present in each data set (some areas had no coverage for particular network operators). V: Verizon, A:AT&T, T:T-Mobile, S: Sprint.

and application-level metrics. For video, application-level metrics are measured by streaming YouTube videos; we choose YouTube because it has an extensive mobile reach of 88%, more than 3.5× that of its next best competitor [34]. For Web browsing, we utilize the Tranco ranking list as it addresses the stability and responsiveness issues faced by other Website ranking services [32].

We run our measurement suite on Lenovo ThinkPad W550s laptops, each of which is tethered to its own Motorola G7 Power (Android 9) via USB in order to measure cellular performance. The cellular plans on all our cellular user equipment (UE) have unlimited hot-spot data enabled to effectively achieve the same level of performance as we would on the mobile device. We run our measurement suite on laptops tethered to phones as this configuration gives us the same application performance while facilitating ease of programming and data extraction, and achieving higher level of unification for various application-level measurements. We record instantaneous RSRP readings from the UEs every one second through the Network Monitor application [21].

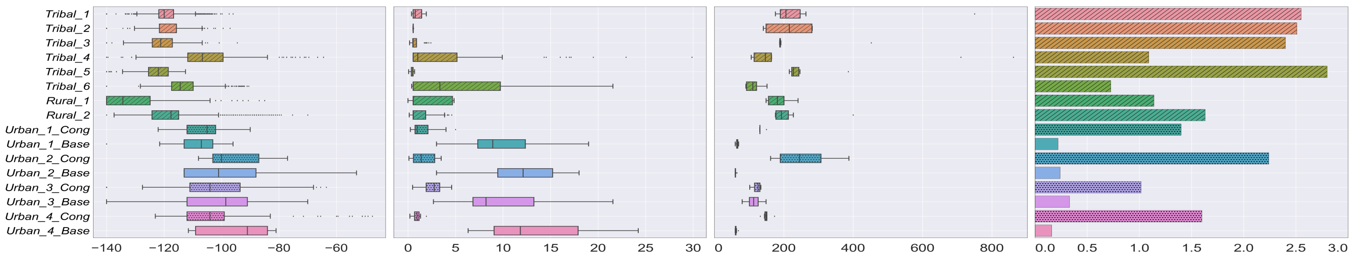
To collect video QoE metrics, we run a 3-minute clip of a Looney Tunes video three times across each of the four LTE providers at each location; we exclude from our results the sessions that experienced playback errors during execution. We chose this particular video due its mix of high and low action scenes, which results in variable bitrates for different segments in the video (typically, high action scenes have higher bitrate than low action scenes). To infer video QoE, we collect the input features (RSRP and throughput) synchronously, on a separate device, so as not to bias the video streaming measurements. After testing multiple playback durations, we observed that a 3-minute window was adequate for the playback to reach steady state, while long enough to capture rebuffering and/or resolution switches that occur. To execute this experiment, we first automate the loading and playback of the YouTube video on the Chrome browser using Selenium [35]. The video resolution is set to auto. Then we use YouTube's iframe API [36] to capture playback events reported by the video player. The API outputs a set of values that indicate player state (not started, paused, playing, completed, buffering) using the getPlayerState() function. The API also provides functions for accessing information about play time and the remaining buffer size. Similarly, we employ Selenium to automate the loading of Web pages on the Chromium browser to capture page load time measurements.

B. *Description of Datasets*

We collected 16 datasets from 12 locations across the Southwestern U.S. Eight of those datasets were collected from rural locations that had sparse cellular deployment. Six of the eight locations were under the jurisdiction of American Indian Tribal sovereignty, while two were from non-tribal rural regions. In the remaining text, we sometimes use the word "rural" to refer to both tribal and non-tribal rural areas. These eight collection points span an area of 21 square miles in New Mexico and were collected over a period of five days. In addition, we collect another eight datasets from four urban locations in California. For each urban location, we collect two datasets: one during a large event or gathering, in which we expect cellular network congestion to occur (these datasets are marked with **_Cong**) [13], [37]; and a second during typical operating conditions. We call the latter dataset the baseline for that location (marked with **_Base**). Hence, our traces are broadly classified into three categories: rural and tribal, congested urban, and baseline urban. The details of each dataset are summarized in Table II. In each location, we concurrently collected the complete set of traces on four major U.S. carriers (AT&T, Sprint, T-Mobile and Verizon) using four separate, equivalent UEs, as described in §III-A. The designation of each location as tribal, rural or urban ("type" column in Table II) is based on information gathered from the Census Bureau [38]. Through these measurement campaigns, we collect and analyze over 32.7 million LTE packets.

IV. Network Characterization

In this section, we analyze network performance characteristics in each measurement location, and by so doing we attempt to determine whether any generalizations based on network offered load exist. Note that in certain locations, one or more operators did not provide LTE coverage, as is indicated in Table II. Our assessment reveals several discernible differences in network performance across region types and network conditions (congested/uncongested).



a. RSRP (dBm) b. Throughput (Mbps) c. Latency (ms) d. Packet loss rate (%)

Fig. 1: Distribution of QoS metrics across different network conditions.

A. Quality of Service Analysis

We evaluate the relationship between spatially and temporally varying network conditions through four QoS metrics: RSRP, throughput, latency and packet loss, as described in §II-A. We analyze both the mean and median values and present median results aggregated across all telecom providers at each location. In addition, wherever applicable we report any performance deviations that fall beyond one standard deviation (68% confidence interval) because they may skew the distribution. Stated otherwise, we separately report the performance of each individual telecom operator if $x \leq \mu - \sigma$ or $x \geq \mu + \sigma$, where x is the median performance value for the outlier operator, μ and σ are the mean and standard deviation of the entire distribution. In the boxplots, the right and left edges of the box represent the first quartile (25^{th} percentile) and third quartile (75^{th} percentile), respectively, with the median line drawn within the box. The whiskers capture 5^{th} to 95^{th} percentile values.

RSRP: We observe a wide range of RSRP values on all levels: between datasets, within datasets, and between dataset types, as shown in Figure 1(a). The median RSRP value across all rural and tribal locations is -118dBm. We observe that rural and tribal locations report 15.06% and 19.55% lower RSRP values than urban congested and baseline urban locations, respectively.[1] This result is consistent with reports of limited LTE coverage in rural and tribal locations; these regions frequently have significantly sparser LTE base station deployment, and hence larger coverage areas that lead to regions with lower signal quality. One notable exception is Tribal_4 where the reported values are, on average, 12dBm higher than elsewhere in the rural region. This is likely due to the relatively denser deployment of base stations in the area; our wireless gear was physically closer to the eNodeBs (LTE base stations) that served the region (verified through CellMapper [39]).

Throughput: Figure 1(b) compares throughput across all locations, averaged across all providers present in each location. We observe high variability in throughput distributions, with values that range from less than 500 Kbps to about 30 Mbps. Congested urban measurements report a median throughput of 1.51 Mbps, while rural and tribal regions report 35% less median throughput, at 0.98 Mbps. Uncongested urban locations have by far the best average performance.

We observe a few outliers in our analysis: AT&T reports $30\times$ and $26\times$ the median throughput values in Tribal_4 and Tribal_6, faring considerably better than its competitors. In addition, Sprint performs $8\times$ better than the region's median. In Tribal_4, if we exclude the outlier (AT&T), the median value for the other operators is 0.56 Mbps, which is 41% less than the worst performing congested dataset, Urban_4_Cong (median: 0.95 Mbps). This is unexpected since: (1) Tribal_4 has denser coverage and (2) our measurement setup was in close proximity to all the connected eNodeBs (this resulted in 12dBm higher RSRP than other rural locations). Similarly, if we exclude the outliers in Tribal_6 (AT&T and Sprint), the median throughput is 43% worse than Urban_4_Cong.

Our results demonstrate that, on average, the LTE networks we measured in rural regions perform worse than congested urban networks. This performance difference, and the absolute performance values, likely indicate the difference between having a stable teleconferencing session and an unusable service (e.g., Zoom recommends a minimum downstream bandwidth of 1.5 Mbps [40]). Urban baseline locations register a median throughput of 10.92 Mbps. In comparison, congested locations report $7\times$ less median throughput, while rural and tribal regions register $11\times$ lower throughput than baseline urban locations.

Latency: Figure 1(c) shows the average latency, measured as round-trip time (RTT), across all measurements in each location. Urban baseline datasets reveal a median latency of 64 ms, while in the congested networks, the average RTT more than doubles to a median value of 140 ms, again verifying our expectation of network congestion. Rural regions report a median latency of 193 ms, which translates to a 38% and 202% increase in round trip-times compared to congested and urban baseline datasets, respectively. Notably, Tribal_4 has an average latency of 147 ms despite close proximity of our measurement setup to the LTE base stations, and a location geographically closer to our ping server in Virginia, than the locations in California. Reasons for this extra latency are varied, and may include a less direct and/or slower path out of this region to a major Internet backbone. Overall, Tribal_2 and Urban_2_Cong exhibit the widest variability in latency measurements.

Packet Loss: The average loss rate across providers is

[1] The small difference in RSRP median value between the congested urban and baseline results may stem from multiple causes: a difference in weather conditions on measurement days, a change in transmission power of the eNodeBs due to utilization, and the fact that the Urban_4_Base collection location was approximately 30m away from the Urban_4_Cong location due to the closure of the original collection venue.

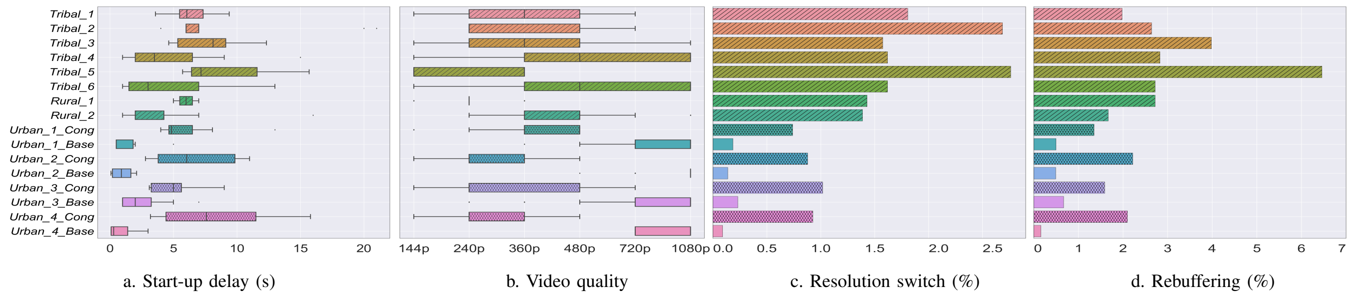


Fig. 2: Distribution of QoE metrics across different network conditions.

reported in Figure 1(d). While we observe variability, a birds-eye view clearly demonstrates that the baseline urban locations benefit from marginal loss rates (median is 0.23%). We observe more than 6× higher median loss rates in the congested urban datasets, again indicating heavy congestion. As a group, the congested urban locations experience the second best performance (median loss rate is 1.56%), and rural networks experience the highest average loss rates. We observe particularly high loss in Tribal_1, Tribal_2, Tribal_3 and Tribal_5 (median of 2.53% across all four locations). A performance outlier is Tribal_6; despite its comparable rural and tribal location, Tribal_6 reports a much lower loss rate of 0.72%.

Takeaway: Our analysis of QoS metrics reveals the wide gap in LTE performance across different regions and network conditions. Our results illustrate that the rural and tribal regions we study experience the poorest mobile broadband performance, performance that is typically even worse than heavily congested urban networks. This poor performance is consistent with prior findings [9], [41]. The chronic under-provisioning of LTE networks in rural and tribal regions, due to both sparser deployment and some combination of less efficient and lower speed backhaul, implies mobile broadband in these regions often cannot meet the minimum recommended QoS required for applications such as video streaming and video chat. While this poor Internet usability has a negative impact on local residents, this impact has been grossly amplified during the shelter-in-place orders of the COVID-19 pandemic, when schooling, work, telemedicine, and other critical activities have been moved online [4], [15]. Our measurements indicate that in many of these rural locations, despite the presence of mobile broadband, the quality of those networks is often too poor to support these now-essential video-based applications.

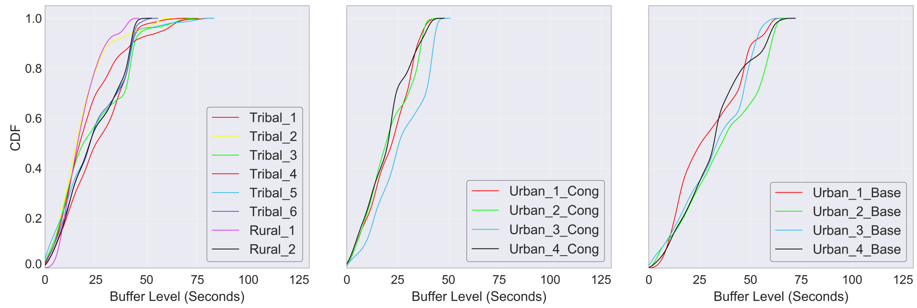
B. Video QoE Analysis

Next we characterize key video QoE indicators in different network conditions to reveal application-level performance differentials of video streaming. Similar to §IV-A, we report median values across all telecom providers unless there are samples that lie outside of one standard deviation (μ).

Start-up Delay: Figure 2(a) plots the start-up delay. In rural and tribal regions, the median start-up delay is 6.52s, while congested urban locations report 5.29s. Urban baselines have the lowest reported delay at 0.7s (median). We also note that the rural and tribal locations have far higher variability than congested datasets. For instance, the median range of start-up delay (i.e., difference between the max and the min values in a distribution) is 12.5s in these areas as opposed to 8.6s in congested networks. This behavior can be attributed to under-provisioned LTE networks in rural and tribal regions that are sensitive to user demand, even during normal operating hours, thereby resulting in large fluctuations. This result is consistent with observations in Figure 1(b); lower throughput coupled with higher packet loss would likely result in the inconsistencies in download time of the initial video segments. Baseline urban offers the least variability of 3.7s. In Figure 1(b), we saw that the AT&T network achieves higher throughput at Tribal_4 and Tribal_6. Consequently, we observe that video sessions in both Tribal_4 and Tribal_6 on AT&T network have 3× lower start-up delay than the other providers. We note that start-up delay does not convey any information about playback video resolution.

Video Quality: Figure 2(b) depicts playback resolution of the YouTube video, sampled at one second granularity. During our measurements, we ensure that the video resolution is set to auto. As a result, playback resolution and resolution switches are a direct result of network conditions and changes in congestion levels. While all of the urban baseline locations indicate near full-HD (1080p) rendering of the video, congested locations have a median resolution of 240p. One possible explanation could be the throttling of video quality by telecoms as part of their congestion mitigation schemes. Rural measurements show marginally better performance with a median resolution of 360p, but exhibit wider variability, ranging between 144p to 1080p. Video sessions with 1080p in rural regions are associated with the AT&T network, which is consistent with the results from Figure 1(b) where AT&T records distinctively higher throughput values.

Resolution Switching: Variability in video resolution is synonymous with quality switches, which is often perceived as a QoE performance degradation [42]. Figure 2(c) represents the number of samples that had a different resolution than the previous sample, as a percentage of total collected samples during the video session. We observe a median value of 1.64% in rural regions, as opposed to 0.92% in congested urban datasets. This value is nearly 6× smaller for baseline urban datasets, with a median value of 0.16%, as compared to urban congested measurements. Frequent resolution switches typically



a. Rural and tribal b. Congested urban c. Baseline urban

Fig. 3: Cumulative distribution of buffer size across different network conditions.

lead to lower user engagement [42]. This implies that the high percentage of switches in rural and tribal regions could lead to difficulty in user engagement with video streaming services, as needed during remote learning, work at home video conferences, and other vital applications.

Rebuffering: This metric represents the amount of time the video stalls or rebuffers during the playback expressed as a percentage of total playback time, shown in Figure 2(d). There is a higher median rebuffering ratio of 2.68% in rural and tribal regions, followed by congested urban regions at 1.85%. Baseline urban measurements report a more than 5× and 3× lower rebuffering ratio than rural/tribal and congested urban, respectively.

Buffer Size: Figure 3 shows the buffer size distribution captured during YouTube streaming sessions. For ease of comprehension, we separate the result graphs into the three categories. Here, a greater amount of buffered content is better, as it allows the application to smooth performance despite varying jitter. Baseline measurements report higher buffer levels with a median value of 44.3s (Figure 3(c)) while congested datasets report 27.7s (Figure 3(b)), a 34% decrease from baseline measurements. Rural regions have the lowest median buffer at 20.2s (Figure 3(a)), which is a reduction by 52% and 27% from baseline urban and congested urban measurements, respectively.

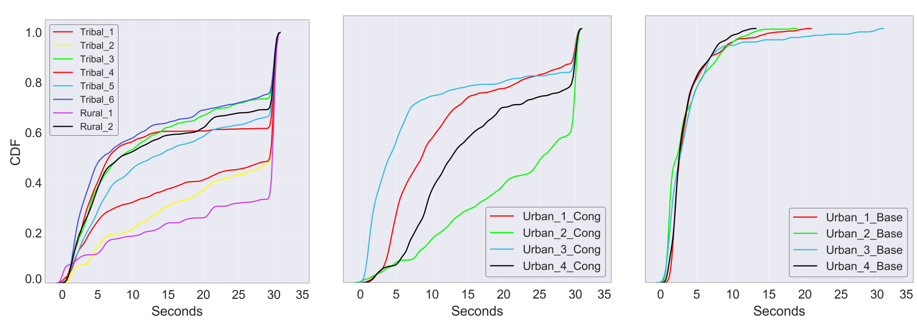
Takeaway: Our analysis of QoE metrics indicates that user experience suffers due to under-provisioned LTE networks in rural and tribal regions. The results reinforce our findings in §IV-A that LTE networks in these regions are likely to fail to provide a quality, or even usable, experience for video streaming. Unsurprisingly, in most cases rural networks underperform in comparison to congested LTE networks in urban regions, implying that *the worst case experience in an urban network is likely still better than the average case experience in a rural or tribal region*. However, the extent of performance degradation in rural and tribal areas as compared to other network conditions is remarkable and noteworthy.

C. Page Load Time

Web performance has long been crucial to the Internet ecosystem since a significant fraction of Internet content is consumed as Web pages. The end-user quality perception in the context of interactive data services is dominated by Web page loading times; the longer the wait, the lower the user satisfaction [43]. Studies have shown that *perceived* time for users accessing the Web can be exceedingly magnified

TABLE III: Webpage Load Timeouts.

Locations	PLT Timeout
Tribal_ 1	16.9%
Tribal_2	51.6%
Tribal_3	26.4%
Tribal_4	46.7%
Tribal_5	30.1%
Tribal_6	24.6%
Rural_1	66.1%
Rural_2	27.3%
Urban_1_Cong	17.6%
Urban_1_Base	0.0%
Urban_2_Cong	41.8%
Urban_2_Base	0.0%
Urban_3_Cong	17.3%
Urban_3_Base	0.88%
Urban_4_Cong	24.0%
Urban_4_Base	0.0%



a. Rural and tribal b. Congested urban c. Baseline urban

Fig. 4: Page load times of Tranco top 25 websites.

with respect to actual chronological time, thus degrading the *perceived* performance even further [44], [45]. Page load times are depicted in Figure 4. We bin the results into similar categories as in Figure 3. From our evaluation we learn that rural and tribal locations (shown in Figure 4(a)) produce the slowest load times with a median value of 9.75 seconds. This is 74% slower than congested dataset (median value is 5.6 seconds) and 2.7× lower than baseline urban measurements (median value is 2.67 seconds). Tribal_2 performs the worst in tribal and rural locations with a median load time of 13.08 seconds while Urban_2_Cong is the most under-performing dataset in congested urban at 7.28 seconds. In baseline urban, all of the locations exhibit similar load times within a margin of ±1 second difference. Our examination reveals a considerable fraction of Web pages fail to load within the timeout period of 30 seconds (shown in Table III) in rural, tribal and congested urban regions. We observe a median timeout value of 28.7% across tribal and rural areas, which is 38.7% higher than that reported in congested urban (median value is 20.8%). Rural_1 logs the highest timeout percentage with over 66% of Web pages failing to load; Urban_2_Cong reports the highest (41.8%) in the congested dataset. Urban baseline locations have faster load times and little to no timeouts for Web pages. Similar to section IV-B, we observe that Web browsing experience suffers more in rural and tribal regions than in urban regions (with or without atypical network utilization).

V. Related Work

Manual measurements are a common approach to calculate cellular coverage [46]. This includes methods such as war-driving [47], war-walking [48], and aerial systems [49], which usually require high operational expenditure. Mobile analytics companies [50], [51] contribute to measurement collection by crowd-sourcing measurements directly from end-user devices via standalone mobile apps [50], [52] or measurement SDKs [52]–[54] integrated into partner apps. However, these are limited in scope because crowd-sourced measurements do not have spatial uniformity. As a result, some of the desired measurement locations may not exist in these databases (possibly due to lack of adoption of the application or SDK by the local community). Further, outsourced databases typically carry a hefty licensing fee or are otherwise restricted [55]. While several public datasets consist of Internet performance measurements (e.g. [54], [56]), there is a lack of datasets that represent the variability in mobile broadband performance as a consequence of sparse deployments or network congestion. Many mobile network datasets focus on coverage [50], [57]; the FCC annually publishes its own broadband report [7]. Unfortunately, these broadband reports are widely known to be inaccurate [58]. Further, we have shown that the use of coverage maps alone are inadequate to infer actual usability. While several prior studies [59], [60] have focused on LTE performance analysis and traffic characterization, these studies do not compare performance across differing population densities and region types.

VI. Conclusion

Online learning, work-at-home, tele-medicine and other applications that already experienced regular usage have exploded in the post COVID-19 world, transitioning from conveniences to critical everyday applications. Web browsing and video streaming are necessary components of these applications, and as such the study of network performance for users in all regions is crucial. Through extensive measurements, we have revealed the sharp contrast in cellular performance between rural, tribal and urban locations; for instance, video QoE is at least 10× worse and Web browsing is more than 2× worse in the rural and tribal regions we studied than in urban locations with typical cellular load. While prior work and past surveys have reached similar conclusions, our study demonstrates and quantifies the extent to which network performance lags in rural and tribal communities. This suggests that users in under-served regions are far more likely to drop out of virtual engagement such as online lectures and e-learning. User disengagement will unfortunately lead directly to a greater digital gap than exists today [4], [15]. Broadband deployments that address these access and coverage quality gaps are urgently needed.

VII. Acknowledgements

Our work would not have been possible without the incredible support of Jerrold Baca. This work was funded through the National Science Foundation Smart & Connected Communities award NSF-1831698.

References

[1] Pew Research Center, "Mobile Fact Sheet," https://www.pewresearch.org/internet/fact-sheet/mobile/, June 2019.

[2] Hansi Lo Wang (NPR), "Native Americans On Tribal Land Are 'The Least Connected' To High-Speed Internet," https://www.npr.org/2018/12/06/673364305/native-americans-on-tribal-land-are-the-least-connected-to-high-speed-internet, December 2018.

[3] V. Adarsh, M. Nekrasov, U. Paul, A. Ermakov, A. Gupta, M. Vigil-Hayes, E. Zegura, and E. Belding, "Too Late for Playback: Estimation of Video Stream Quality in Rural and Urban Contexts," in *Proceedings of the Passive and Active Measurement Conference*, 2021, pp. 141–157.

[4] Amanda Holpuch (The Guardian), "US's Digital Divide 'is going to kill people' as COVID-19 exposes inequalities," https://www.theguardian.com/world/2020/apr/13/coronavirus-covid-19-exposes-cracks-us-digital-divide, April 2020, (Accessed on 05/10/2020).

[5] T. Mangla, E. Showalter, V. Adarsh, K. Jones, M. Vigil-Hayes, E. Belding, and E. Zegura, "A Tale of Three Datasets: Towards Characterizing Mobile Broadband Access in the United States," *Communications of the ACM*, 2021.

[6] Federal Communications Commission, "Coverage Maps Investigation," 2019. [Online]. Available: https://docs.fcc.gov/public/attachments/DOC-361165A1.pdf

[7] Federal Communications Commission, "Broadband Deployment Report," https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2019-broadband-deployment-report, May 2019.

[8] T. Norris, P. L. Vines, and E. M. Hoeffel, "The American Indian and Alaska Native Population: 2010," February 2012.

[9] BroadbandNow, "Internet Access in New Mexico," 2019. [Online]. Available: https://broadbandnow.com/New-Mexico

[10] Sandvine, "Mobile Phenomena Report 2020," https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/2020/Phenomena/Mobile%20Phenomena%20Report%201H%202020%2020200219.pdf, Feb 2020.

[11] Statista, "Penetration Rate of Subscription Video-on-demand (SVOD) Services Worldwide in the 1st Quarter of 2019," 2019. [Online]. Available: https://www.statista.com/statistics/813698/svod-reach-by-country/

[12] Rural Wireless Association, "Challenges Faced by Small Wireless Providers in Measuring LTE Coverage," 2019. [Online]. Available: https://ruralwireless.org/rwa-welcomes-fcc-investigation-into-violation-of-mobility-fund-phase-ii-mapping-rules/

[13] V. Adarsh, M. Nekrasov, E. Zegura, and E. Belding, "Packet-Level Overload Estimation in LTE Networks Using Passive Measurements," in *Proceedings of the Internet Measurement Conference*, 2019, pp. 158–164.

[14] U. Paul, A. Ermakov, M. Nekrasov, V. Adarsh, and E. Belding, "# Outage: Detecting Power and Communication Outages from Social Networks," in *Proceedings of The Web Conference 2020*, 2020, pp. 1819–1829.

[15] Grant Samms (Forbes), "As Cities Face COVID-19, The Digital Divide Becomes More Acute," https://www.forbes.com/sites/pikeresearch/2020/04/02/as-cities-face-covid-19-the-digital-divide-becomes-more-acute/#277c93e558c5, April 2020, (Accessed on 05/10/2020).

[16] 3GPP TS 136.214, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer - Measurements," April 2010.

[17] 3GPP TS 36.214 Release 12, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer Measurements," June 2016.

[18] N. Awad and I. Mkwawa, "The Impact Of The Reference Signal Received Power To Quality Of Experience For Video Streaming Over LTE Network," in *Proceedings of the Annual Conference on New Trends in Information Communications Technology Applications (NTICT)*, 2017.

[19] D. Aziz and R. Sigle, "Improvement of LTE Handover Performance through Interference Coordination," in *Proceedings of the 69th Vehicular Tech. Conf.* IEEE, 2009.

[20] E. Alimpertis, A. Markopoulou, C. Butts, and K. Psounis, "City-Wide Signal Strength Maps: Prediction with Random Forests," in *Proceedings of the The Web Conference*, ser. WWW, 2019, pp. 2536–2542.

[21] B. Lubek, "Network Monitor." [Online]. Available: https://github.com/caarmen/network-monitor

[22] Kali Tools, "Hping3 Package," 2019. [Online]. Available: https://tools.kali.org/information-gathering/hping3

[23] V. Adarsh, P. Schmitt, and E. Belding, "MPTCP Performance over Heterogenous Subpaths," in *Proceedings of the 28th International Conference on Computer Communication and Networks (ICCCN)*. IEEE, 2019, pp. 1–9.

[24] T. Hoßfeld, M. Seufert, C. Sieber, and T. Zinner, "Assessing Effect Sizes of Influence Factors towards a QoE Model for HTTP Adaptive Streaming," in *Sixth International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2014, pp. 111–116.

[25] I. Sodagar, "The MPEG-DASH Standard for Multimedia Streaming over the Internet," *IEEE Multimedia*, vol. 18, no. 4, pp. 62–67, 2011.

[26] A. Bouch, A. Kuchinsky, and N. Bhatti, "Quality is in the Eye of the Beholder: Meeting Users' Requirements for Internet Quality of Service," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2000, pp. 297–304.

[27] N. Bhatti, A. Bouch, and A. Kuchinsky, "Integrating User-perceived Quality into Web Server Design," *Computer Networks*, vol. 33, no. 1, pp. 1–16, 2000.

[28] F. Bronzino, P. Schmitt, S. Ayoubi, G. Martins, R. Teixeira, and N. Feamster, "Inferring Streaming Video Quality from Encrypted Traffic: Practical Models and Deployment Experience," *Proceedings of the ACM Meas. Anal. Comput. Syst.*, December 2019.

[29] M. Ghobadi, Y. Cheng, A. Jain, and M. Mathis, "Trickle: Rate Limiting YouTube Video Streaming," in *Proceedings of the USENIX Annual Technical Conference (ATC)*, 2012, pp. 191–196.

[30] F. Dobrian, V. Sekar, A. Awan, I. Stoica, D. Joseph, A. Ganjam, J. Zhan, and H. Zhang, "Understanding the Impact of Video Quality on User Engagement," in *Proceedings of the ACM SIGCOMM Conference*, 2011.

[31] S. S. Salunke, *Selenium Webdriver in Python: Learn with Examples*. CreateSpace Independent Publishing Platform, 2014.

[32] V. L. Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen, "Tranco: A Research-oriented Top Sites Ranking Hardened against Manipulation," *arXiv preprint arXiv:1806.01156*, 2018.

[33] W3C, "Navigation Timing," 2019. [Online]. Available: https://www.w3.org/TR/navigation-timing/#processing-model

[34] Statista, "Most Popular Video Streaming Services in the US," 2019. [Online]. Available: https://www.statista.com/statistics/910895/us-most-popular-video-streaming-services-by-reach/

[35] Selenium, "The Selenium Browser Automation Project." [Online]. Available: https://www.selenium.dev/documentation/en/

[36] YouTube, "YouTube Player API Reference for iframe Embeds," 2019. [Online]. Available: https://developers.google.com/youtube/iframe_api_reference

[37] V. Adarsh, M. Nekrasov, U. Paul, and E. M. Belding, "Estimation of Congestion from Cellular Walled Gardens using Passive Measurements," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2021.

[38] M. Ratcliffe, C. Burd, K. Holder, and A. Fields, "Defining Rural at the US Census Bureau," *American community survey and geography brief*, vol. 1, no. 8, 2016.

[39] CellMapper, "Crowd-sourced Cellular Tower and Coverage Mapping Service," https://www.cellmapper.net/map, 2019.

[40] Zoom, "Minimum System Requirements." [Online]. Available: https://support.zoom.us/hc/en-us/articles/201362023-System-requirements-for-Windows-macOS-and-Linux

[41] Federal Communications Commission, "Measuring Broadband America," https://www.fcc.gov/reports-research/reports/measuring-broadband-america/measuring-broadband-america-mobile-2013-2018, February 2018.

[42] V. Krishnamoorthi, N. Carlsson, E. Halepovic, and E. Petajan, "BUFFEST: Predicting Buffer Conditions and Real-time Requirements of HTTP (S) Adaptive Streaming Clients," in *Proceedings of the 8th ACM on Multimedia Systems Conference*, 2017, pp. 76–87.

[43] S. Egger, T. Hossfeld, R. Schatz, and M. Fiedler, "Waiting Times in Quality of Experience for Web based Services," in *Fourth International Workshop on Quality of Multimedia Experience*. IEEE, 2012, pp. 86–96.

[44] S. Grondin, "Timing and Time Perception: A Review of Recent Behavioral and Neuroscience Findings and Theoretical Directions," *Attention, Perception, & Psychophysics*, vol. 72, no. 3, pp. 561–582, 2010.

[45] R. Schatz, S. Egger, and A. Platzer, "Poor, Good Enough or Even Better? Bridging the Gap between Acceptability and QoE of Mobile Broadband Data Services," in *Proceedings of the IEEE International Conference on Communications (ICC)*. IEEE, 2011, pp. 1–6.

[46] J. Laiho, A. Wacker, and T. Novosad, *Radio Network Planning and Optimisation for UMTS*. Wiley Online Library, 2002, vol. 2.

[47] C. Hurley, R. Rogers, F. Thornton, and B. Baker, *Wardriving and Wireless Penetration Testing*. Syngress, 2007.

[48] A. W. T. Tsui, W.-C. Lin, W.-J. Chen, P. Huang, and H.-H. Chu, "Accuracy Performance Analysis Between War Driving and War Walking in Metropolitan Wi-Fi Localization," *IEEE Transactions on Mobile Computing*, vol. 9, no. 11, pp. 1551–1562, 2010.

[49] M. Nekrasov, V. Adarsh, U. Paul, E. Showalter, E. Zegura, M. Vigil-Hayes, and E. Belding, "Evaluating LTE Coverage and Quality from an Unmanned Aircraft System," in *Proceedings of the 16th IEEE International Conference on Mobile Ad-Hoc and Smart Systems*, 2019.

[50] OpenSignal, "Open Signal 3G and 4G LTE Cell Coverage Map," *línea]. Available: http://opensignal.com*, 2016.

[51] Root Metrics Inc., "Metro RootScore Reports," 2020. [Online]. Available: http://www.rootmetrics.com/en-US/home

[52] SamKnows, "Methodology and Technical Information Relating to the SamKnows Testing Platform," *Document Reference: SQ301-002-EN*, 2012.

[53] Tutela Inc., " Crowdsourced mobile data," 2020. [Online]. Available: https://www.tutela.com/

[54] Measurement Lab, "The M-Lab NDT Data Set," https://measurementlab.net/tests/ndt, (2009-02-11 – 2015-12-21).

[55] F. Wamser, N. Wehner, M. Seufert, P. Casas, and P. Tran-Gia, "YouTube QoE Monitoring with YoMoApp: A Web-Based Data Interface for Researchers," in *Proceedings of the Network Traffic Measurement and Analysis Conference*. IEEE, 2018, pp. 1–2.

[56] CAIDA, "CAIDA Data - Overview of Datasets, Monitors, and Reports," https://www.caida.org/data/overview/, 2019.

[57] Skyhook, "Skyhook Coverage Area," https://www.skyhook.com/coverage-map, 2019.

[58] Government Accountability Office, "Broadband Internet: FCC's Data Overstate Access on Tribal Lands," September 2018.

[59] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen, and O. Spatscheck, "An In-Depth Study of LTE: Effect of Network Protocol and Application Behavior on Performance," in *Proceedings of the ACM SIGCOMM Conference*, 2013, p. 363–374.

[60] A. Nikravesh, D. R. Choffnes, E. Katz-Bassett, Z. M. Mao, and M. Welsh, "Mobile Network Performance from User Devices: A Longitudinal, Multidimensional Analysis," in *Proceedings of the Passive and Active Measurement Conference*, 2014.